JMIR Formative Research

Impact Factor (2023): 2.0
Volume 6 (2022), Issue 7 ISSN 2561-326X Editor in Chief: Amaryllis Mavragani, PhDc

Contents

Original Papers

Short Papers

Protocol

Corrigenda and Addendas

Original Paper

An mHealth Platform for Augmenting Behavioral Health in Primary Care: Longitudinal Feasibility Study

Khatiya Chelidze Moon[1], MD; Michael Sobolev[2], PhD; Megan Grella[3], LCSW, MBA; George Alvarado[3], MD; Manish Sapra[3], MD; Trever Ball[3], PhD

[1]Zucker Hillside Hospital, Glen Oaks, NY, United States

[2]Cornell Tech, New York, NY, United States

[3]Northwell Health, Manhassett, NY, United States

Corresponding Author:
Khatiya Chelidze Moon, MD
Zucker Hillside Hospital
Kaufmann Building, Suite k204
75-59 263rd Street
Glen Oaks, NY, 11004
United States
Phone: 1 7184704367
Email: kmoon2@northwell.edu

Abstract

Background: The collaborative care model is a well-established system of behavioral health care within primary care settings. There is potential for mobile health (mHealth) technology to augment collaborative behavioral health care in primary care settings, thereby improving scalability, efficiency, and clinical outcomes.

Objective: We aimed to assess the feasibility of engaging with and the preliminary clinical outcomes of an mHealth platform that was used to augment an existing collaborative care program in primary care settings.

Methods: We performed a longitudinal, single-arm feasibility study of an mHealth platform that was used to augment collaborative care. A total of 3 behavioral health care managers, who were responsible for coordinating disease management in 6 primary care practices, encouraged participants to use a mobile app to augment the collaborative model of behavioral health care. The mHealth platform's functions included asynchronous chats with the behavioral health care managers, depression self-report assessments, and psychoeducational content. The primary outcome was the feasibility of engagement, which was based on the number and type of participant-generated actions that were completed in the app. The primary clinical end point was a comparison of the baseline and final assessments of the Patient Health Questionnaire-9.

Results: Of the 245 individuals who were referred by their primary care provider for behavioral health services, 89 (36.3%) consented to app-augmented behavioral health care. Only 12% (11/89) never engaged with the app during the study period. Across all participants, we observed a median engagement of 7 (IQR 12; mean 10.4; range 0-130) actions in the app (participants: n=78). The chat function was the most popular, followed by psychoeducational content and assessments. The subgroup analysis revealed no significant differences in app usage by age (P=.42) or sex (P=.84). The clinical improvement rate in our sample was 73% (32/44), although follow-up assessments were only available for 49% (44/89) of participants.

Conclusions: Our preliminary findings indicate the moderate feasibility of using mHealth technology to augment behavioral health care in primary care settings. The results of this study are applicable to improving the design and implementation of mobile apps in collaborative care.

(JMIR Form Res 2022;6(7):e36021) doi:10.2196/36021

KEYWORDS

collaborative care; mobile health; psychiatry; depression; virtual care; psychoeducation; mobile app; mobile phone

Introduction

Background

The reach of behavioral health services is insufficient for meeting the needs of the population [1,2]. The collaborative care model (CoCM) is a framework that attempts to meet this vast need for behavioral health services by embedding these services in primary care settings [3]. The CoCM is a system of outcome-driven, stepwise care for systematically identifying individuals who would benefit from behavioral health treatment and supporting primary care clinicians in their management. The model has been adapted by many health systems since its introduction in the 1990s and is considered best practice [4-6]. Unfortunately, there are challenges that limit the scalability of the CoCM, including financial and operational barriers [7,8]. Innovative approaches are needed to support these existing models of behavioral health care [9].

There is emerging evidence that digital and mobile health (mHealth) technologies have the potential to improve the reach of the CoCM [9-12]. Given the near ubiquity of smartphones with app capabilities, health systems are increasingly interested in understanding whether these tools can be harnessed to further extend collaborative care [13,14]. There are several meaningful ways that mobile apps could be used to augment collaborative care. The CoCM relies strongly on a measurement-based system of care in which validated clinical assessments (eg, the Patient Health Questionnaire-9 [PHQ-9]) are regularly collected to assess clinical responses [15]. mHealth platforms could help decrease care providers' workload by automating the collection of these measures [12,13]. Apps could also support clinical decision-making by collecting clinical information more frequently than what is currently possible [10]. Moreover, apps could facilitate more frequent communication between patients and care providers [16,17]. Finally, apps can act as repositories for educational materials and self-guided modules for reinforcing concepts that are learned in therapy and promoting patient engagement [13,18-20]. All of these factors have the potential to increase patient engagement in care and, eventually, result in improved clinical outcomes.

Despite the theoretical benefits of app-augmented collaborative care, relatively little is still known about the feasibility of app usage in collaborative care. It is unknown whether patients in collaborative care settings are likely to use apps, what features of mobile apps are the most beneficial in this setting, and what patient population(s) may be the most likely to benefit from app-augmented collaborative care [11]. For example, in the broader literature on mHealth, there is concern that older individuals may be less familiar with technology and may therefore be less inclined to engage with it [21]; however, this has not been systematically examined, to our knowledge, in the collaborative care setting. The limited existing studies suggest that overall, app usage among patients in collaborative care can be variable [9,12]. Understanding app usage is applicable to the optimization of mHealth platform interventions and their implementation in collaborative care [22]. If known, this information could lay the groundwork for improving the design and implementation of mobile apps in collaborative care.

Objectives

This study describes a feasibility study of the Valera Health mobile platform and app (Valera Health Inc), which was used to augment collaborative care within primary care practices in a large health care system. Our primary aim was to assess the feasibility of app usage, which was measured based on engagement. A secondary outcome was preliminary clinical improvement in depression scores.

Methods

Ethics Approval

The study methods were approved by the Institutional Review Board of Northwell Health (approval number: 20-0545-NH). Informed consent was not sought due to the retrospective nature of the study.

Study Overview

This was a retrospective review of a longitudinal, single-arm implementation initiative wherein individuals who were referred to the collaborative care program by their primary care providers (PCPs) were invited to participate in app-augmented collaborative care by the behavioral health care manager (BHCM). Individuals who were qualified to participate and agreed to do so were asked to download the Valera Health mobile app. Participants were told to use the app to complete in-app PHQ-9 measures, which were sent by the BHCM at preset monthly intervals; communicate with the BHCM through asynchronous chats as needed; and access the psychoeducational content in the app. Participants also experienced all usual collaborative care interventions, as described in the *Study Setting* section, including office visits and telephone contacts with the BHCM, short-term psychotherapy, care coordination, psychiatric case reviews, follow-ups with their PCPs as indicated, and the prescription of recommended psychiatric medications if indicated. PHQ-9 assessments were able to be completed through the app or on paper during office visits with the BHCM. For our primary outcome—feasibility—usage data on the number of user actions that were completed in the app were recorded by the app throughout the participation period. For our secondary outcome—clinical improvement—baseline and final PHQ-9 scores were compared.

Study Setting

The study was conducted in a large, primarily suburban, academic health care system with multiple affiliated primary care practices. The primary care practices that were involved in this study provide behavioral health services through a system that was modeled after the CoCM introduced in the Improving Mood—Providing Access to Collaborative Treatment trial [6]. Briefly, in this model, patients presenting for routine primary care are systematically screened for depression by their PCPs using the PHQ-9—a tool that has been validated for this purpose [23]. Patients who screen positive on the instrument and/or, in the opinion of the PCP, display clinical features that are concerning for a behavioral health disorder are referred to a BHCM who is physically embedded in the clinic. The BHCM maintains a registry of patients, tracks outcomes via serial PHQ-9 assessments, provides time-limited psychotherapy,

coordinates referrals to continued treatment and/or a higher level of care when necessary, and liaises with a psychiatrist who provides remote supervision to multiple BHCMs. Psychopharmacologic recommendations are relayed to the PCP, who remains the prescriber and clinician of record.

Recruitment

Recruitment was planned in 2 phases. The first phase lasted from November 2018 to June 2019 and included 1 primary care practice with 1 BHCM. A total of 5 additional practices and 2 BHCMs were added in phase 2, which lasted from November 2019 to March 2020. There were no differences in procedures between the two phases except for the number of clinics and BHCMs involved. Patients who were referred to the collaborative care program were invited to participate in app-augmented collaborative care by the BHCM during initial appointments. During recruitment at the initial visits, the BHCM guided the participants through the process of downloading and using the app, answered any initial questions, and provided written instructions on the use of the app. After the initial visits, the BHCM was available by phone to troubleshoot the app as needed. Recruitment was halted in March 2020 when the social distancing measures that were required to prevent the spread of COVID-19 in New York resulted in the remote provision of ambulatory behavioral health services.

Individuals who declined to participate or were excluded received usual collaborative care, including primary care and behavioral health services. Those who agreed to participate received usual collaborative care, as described above, with the Valera Health mobile app as augmentation.

Individuals were included if they were adults with a diagnosis of depression or anxiety. Individuals were also excluded if they did not speak English or had severe mental illnesses, suicidal or violent ideation, or substance abuse disorders. Also excluded were children and individuals who required a referral to a higher level of care or continued treatment after the completion of the program.

Participation flow is illustrated by Figure 1. Between November 2019 and March 2020, a total of 245 individuals were referred by their PCPs to the collaborative care program for behavioral health services. Further, 58% (n=142) of these patients were eligible for and were recruited to participate in our study using the Valera Health mobile app (Figure 2), and 62.7% (89/142) of recruited patients consented to participate; 34 consented during the first phase of piloting the Valera Health mobile app, and 55 consented during the second phase. The time required to train patients in the use of the app was a barrier to recruitment among a sizable minority of individuals in the target population (38/245, 15.5%; Figure 1). In addition, a portion of our eligible patient population was unable to participate due to technical barriers (22/142,15.5%; Figure 1). Further, 7 individuals declined to participate due to privacy concerns (Figure 1). Participants were mostly female (60/89, 67%) and middle-aged (mean 38.6, SD 14 years). Participants were enrolled in the study for an average of 22 weeks.

Figure 1. The flow of patient participation in the Valera Health mobile app pilot for behavioral health.

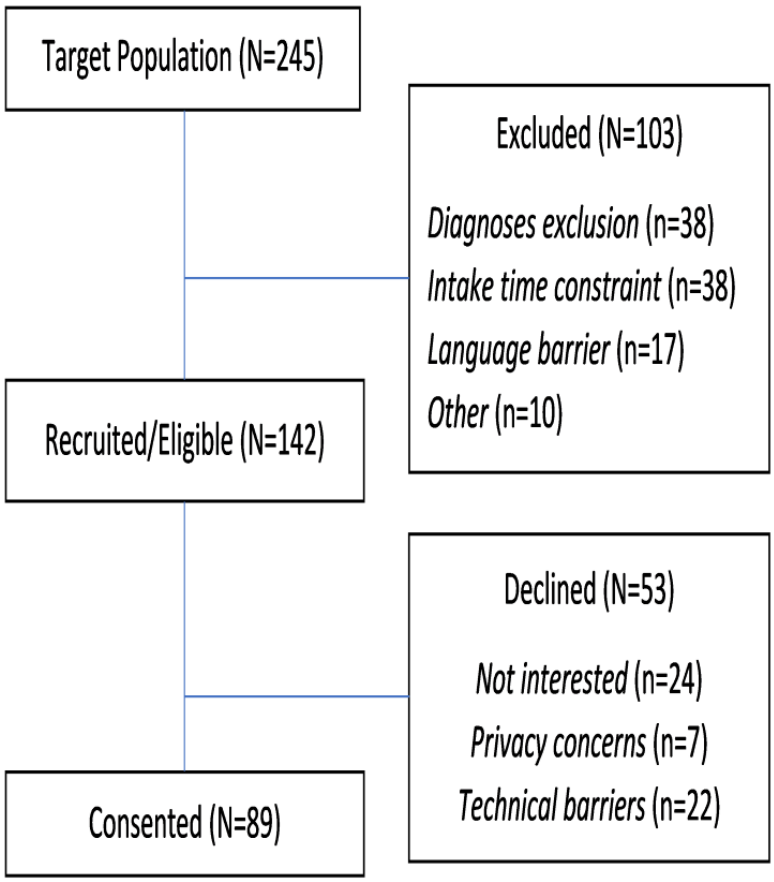
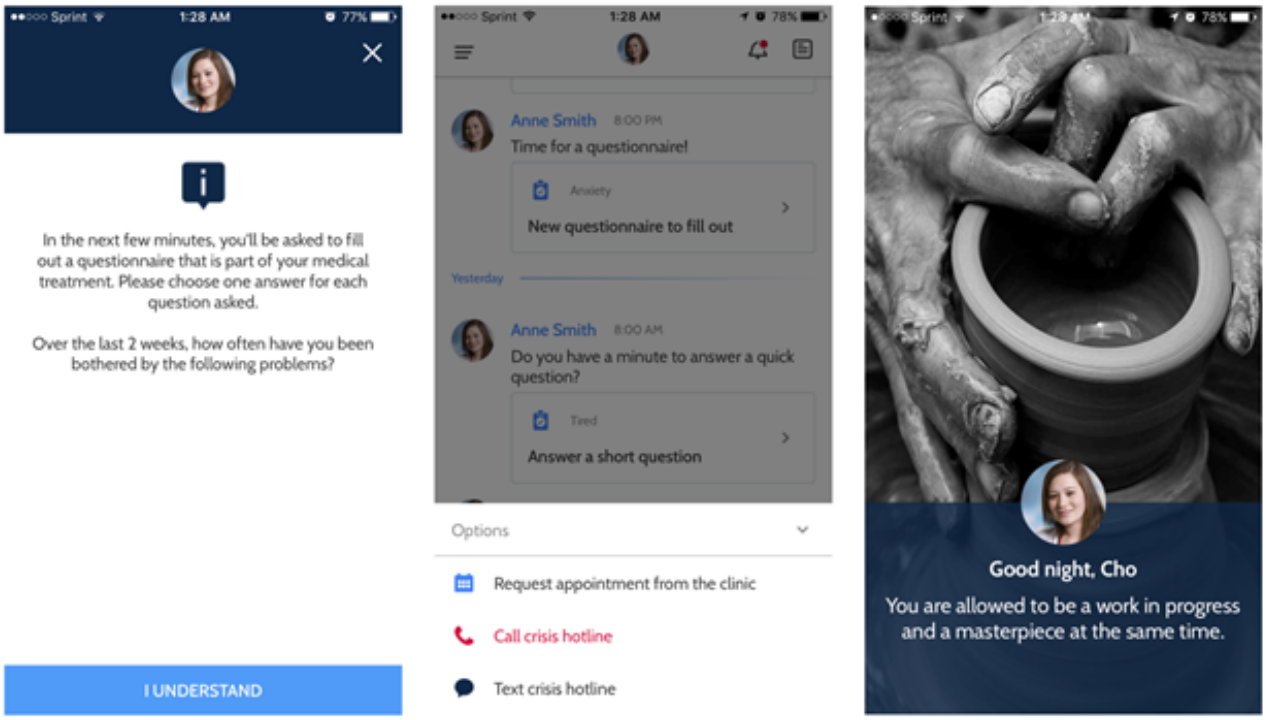


Figure 2. The patient user interface of and the experience on the Valera Health app.



Intervention

The Valera Health mobile app is an English-language secure platform with several functionalities. Figure 2 shows screenshots of the app. First, the app automatically sends PHQ-9 assessments to participants at monthly intervals that are preset by the BHCM. Second, the app allows for secure asynchronous messaging between participants and the BHCM. BHCMs typically responded to chat messages from participants within 1 business day. The app also contains psychoeducational content. This content includes written material about common behavioral health conditions such as depression and anxiety, education about treatments like medication and psychotherapy, instructional guides on topics such as mindfulness, and video and audio clips on these topics. The BHCM had the option of prompting the participants to access psychoeducational content that was relevant to the participants' care via the app. The Valera Health app was not integrated into the electronic health record (EHR). Relevant clinical information from the app, such as PHQ-9 scores, was documented into the EHR by the BHCM.

Outcomes

We assessed the feasibility of the mHealth intervention by investigating engagement with the app and improvement in clinical outcomes. The primary outcome—the feasibility of engagement—was assessed based on app usage, which was measured as the number of participant-generated actions completed in the app. Possible participant-generated actions were (1) the in-app completion of PHQ-9 assessments, (2) the sending of a chat message, or (3) the accessing of in-app psychoeducational material. App usage was monitored throughout the study by the Valera Health app. Engagement was stratified by age and sex. A secondary outcome was clinical improvement, which we defined as a final PHQ-9 score of less than 10 or a greater than 50% reduction in PHQ-9 scores. Baseline PHQ-9 scores were assessed during intake by the BHCM. The final PHQ-9 scores were the last ones recorded for the participants and were extracted from the EHR. Any PHQ-9 assessment, whether it was completed through the app or on paper during office visits with the BHCM, was considered in the analysis of clinical improvement.

Data Analysis

Data were analyzed by using Microsoft Excel, and an α of .05 was set as the a priori level of significance. Simple descriptive statistics of app usage were used to report on feasibility and engagement metrics. A Kruskal-Wallis equality-of-populations rank test of median differences in total mobile app actions was used to analyze differences in app usage between age and sex groups. For our secondary outcome—clinical improvement—a 2-tailed, 2-sample Welch *t* test with unequal variances was used to analyze the difference between baseline and follow-up PHQ-9 scores.

Results

Table 1 describes the brief demographic and engagement characteristics of participants. Only 12% (11/89) never engaged with the app during the trial. Across all participants, we observed a median engagement of 7 (IQR 12; mean 10.4; range 0-130) actions in the app. At least 1 action in the mobile app was completed by 87% (78/89) of participants. Psychoeducational content was reviewed by 75% (67/89) of participants (number of articles reviewed: median 2, IQR 4; range 0-18). Chat messages were sent by 62% (55/89) of participants. Participants sent a median of 1 (IQR 5; range 0-115) chat message used for either scheduling an appointment or reporting symptoms.

A baseline PHQ-9 score was reported for 97% (86/89) of study participants. However, follow-up PHQ-9 assessments were

available only for 49% (44/89) of participants. PHQ-9 scores improved from baseline to follow-up for 73% (32/44) of participants for whom we had baseline and follow-up PHQ-9 scores (n=44). The percentages of participants with improved PHQ-9 scores were not different by sex (*P*=.53) or age groups (18-35, 36-55, and ≥56 years: *P*=.90), although as previously noted, the sample of participants with recorded follow-up PHQ-9 scores was notably smaller than the sample of participants with baseline PHQ-9 scores.

Table 1. Demographic and engagement characteristics of behavioral health patients who participated in the Valera Health mobile app study (N=89).

	Value, n (%)	Age (years), mean (SD)	Baseline PHQ-9[a] score, mean (SD)[b]	Follow-up PHQ-9 score, mean (SD)[c]	*P* value[d]	Total mobile app actions, median (IQR)	*P* value[e]
Total	89 (100)	38.6 (14.5)	11.5 (5.3)	8.6 (4.6)	.002	7 (12)	N/A[f]
Sex							.84
Female	60 (67)	38.4 (14.6)	11.5 (5.3)	9.0 (4.8)	.03	6.5 (12)	
Male	29 (33)	39.1 (14.6)	11.7 (5.5)	7.5 (3.9)	.01	8 (11)	
Age (years)							.42
18-35	44 (51[g])	N/A	11.8 (5.1)	8.7 (4.0)	.01	8 (14)	
36-55	31 (36[g])	N/A	11.1 (4.9)	8.3 (5.9)	.13	6 (9)	
>55	12 (14[g])	N/A	11.7 (7.2)	8.9 (4.3)	.29	5.5 (11.5)	

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

[b]n=86.

[c]n=44.

[d]A 2-sample Welch *t* test with unequal variances between baseline and follow-up PHQ-9 scores.

[e]A Kruskal-Wallis equality-of-populations rank test of median differences in the total mobile app actions.

[f]N/A: not applicable.

[g]The denominator for this percentage is 87.

Discussion

Principal Findings

This study reports on the evaluation of one of the first implementations of mobile app–augmented care within a collaborative care program. The overarching purpose of this feasibility study was to understand whether patients in collaborative care are likely to participate in app-augmented care, what features of the app are used, and whether demographic differences exist among app users in this context. Overall, our results indicate the acceptability and feasibility of app usage; the overwhelming majority of participants (78/89, 87%) used the app at least once, and a modest median of 7 actions were completed in the app. Encouragingly, all of the features of the app were used at similar rates. In particular, the psychoeducational materials and chat feature were both popular functions, suggesting that the app may indeed act to reinforce the concepts that are learned during clinical encounters and can enhance communication between patients and care providers as postulated in the literature [13]. There was no significant difference in app usage by age group (*P*=.42). Finally, we were encouraged by the finding that privacy—a concern in the broader literature surrounding app usage—was an infrequent cause for declining to participate (n=7) in our study, further suggesting acceptability [24,25].

With regard to our secondary outcome, the preliminary clinical outcomes of app-augmented care in this study were encouraging, with 73% (32/44) of participants for whom follow-up PHQ-9 data were available experiencing improvements (n=44). Our findings on clinical outcomes are limited by the considerable drop-off in the number of participants who completed a follow-up PHQ-9 assessment. The reasons for such decreases are not completely known and include app attrition, which is consistent with previous literature showing that attrition is a challenge to the implementation of mobile technologies [12,26]. The missing data also introduced bias into the study, as the population of participants who did not complete follow-up assessments may not be random.

Implications

Our findings reveal several issues that deserve consideration and optimization prior to subsequent implementation efforts. First, despite the prevalence of smartphones, a nontrivial portion of our eligible patient population was unable to participate due to technical barriers (22/142, 15.5%). In addition, the time required to train patients in the use of the app was a barrier to recruitment among a sizable minority of individuals in the target population (38/245, 15.5%). To address these issues, previous literature has postulated the need for a digital health navigator—a new team member with expertise in digital and mobile strategies who can help educate patients on the use of these tools, thereby reducing the burden among staff who may lack this expertise and have insufficient time to address these topics during appointments [14,27,28]. Digital health navigators can be instrumental to training staff in rapidly developing competencies for mHealth [29]. The CoCM, which already operates in a framework of interdisciplinary collaboration, may

be uniquely suited to the adoption of the digital health navigator role in the future.

Second, future implementation efforts should carefully consider measures for mitigating the potential unintended negative consequences of app use. For example, the chat function could give participants the perception of continuous access to the clinician. Though most participants in our study used the chat a moderate amount of times, there was substantial variability, with some participants sending more than 100 messages. Care provider burnout is a concern in the face of such significant increases in patient communication. The future implementation of this and similar technologies would benefit from integration with existing health information systems and care provider workflows to better support care provider decision-making.

Finally, while our results indicate preliminary evidence that patients are willing to participate in app-augmented collaborative care, there exist many opportunities for the optimization of engagement. For example, our intervention allowed the BHCM to nudge patients to engage with psychoeducational materials. Higher engagement can potentially be achieved by personalizing technology to deliver the right content in the right amount and at the right time [11,30,31]. Future research should identify pathways for personalization and investigate its effects on engagement and clinical improvement.

Limitations

Our study has several limitations. First, qualitative data for contextualizing findings were not systematically collected as a part of this study. Second, the experiences of the BHCMs were also not systematically evaluated. This study examined app usage among participants, which was stratified by age and sex; however, other demographic factors of potential interest, such as race, ethnicity, insurance status, were not tracked. Finally, this was a feasibility study; thus, conclusions are limited by the lack of a control group and the unknown characteristics of nonparticipants. Future research should employ a rigorous clinical trial involving patient and/or clinic randomization to evaluate clinical effectiveness.

Conclusions

In conclusion, our preliminary findings indicate the moderate feasibility of using mHealth technology to augment behavioral health care in primary care settings. The results of this study are applicable to improving the design and implementation of mobile apps in collaborative care.

Acknowledgments

This pilot study was sponsored by a Delivery System Reform Incentive Payment grant from New York State.

Conflicts of Interest

None declared.

References

1. Butryn T, Bryant L, Marchionni C, Sholevar F. The shortage of psychiatrists and other mental health providers: Causes, current state, and potential solutions. Int J Acad Med 2017;3(1):5-9 [FREE Full text] [doi: 10.4103/IJAM.IJAM_49_17]
2. Thomas KC, Ellis AR, Konrad TR, Holzer CE, Morrissey JP. County-level estimates of mental health professional shortage in the United States. Psychiatr Serv 2009 Oct;60(10):1323-1328. [doi: 10.1176/ps.2009.60.10.1323] [Medline: 19797371]
3. Unützer J, Katon W, Callahan CM, Williams JWJ, Hunkeler E, Harpole L, IMPACT Investigators. Improving Mood-Promoting Access to Collaborative Treatment. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. JAMA 2002 Dec 11;288(22):2836-2845. [doi: 10.1001/jama.288.22.2836] [Medline: 12472325]
4. Dissemination of integrated care in adult primary care: The collaborative care model. American Psychiatric Association. 2016. URL: https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/Professional-Topics/Integrated-Care/APA-APM-Dissemination-Integrated-Care-Report.pdf [accessed 2022-05-26]
5. Bauer AM, Azzone V, Goldman HH, Alexander L, Unützer J, Coleman-Beattie B, et al. Implementation of collaborative depression management at community-based primary care clinics: an evaluation. Psychiatr Serv 2011 Sep;62(9):1047-1053. [doi: 10.1176/appi.ps.62.9.1047] [Medline: 21885583]
6. Blackmore M, Chung H, Ricketts S, Patel U. A cross sectional prospective comparison of collaborative care and co-location treatment for depressed, low income, diverse patients in primary care. In: Abstracts from the 2017 Society of General Internal Medicine Annual Meeting. 2017 Presented at: 2017 Society of General Internal Medicine Annual Meeting; April 19-22, 2017; Washington, DC p. S92-S93 URL: https://link.springer.com/article/10.1007/s11606-017-4028-8 [doi: 10.1007/s11606-017-4028-8]
7. Blasinsky M, Goldman HH, Unützer J. Project IMPACT: a report on barriers and facilitators to sustainability. Adm Policy Ment Health 2006 Nov;33(6):718-729. [doi: 10.1007/s10488-006-0086-7] [Medline: 16967339]
8. Whitebird RR, Solberg LI, Jaeckels N, Pietruszewski PB, Hadzic S, Unützer J, et al. Effective Implementation of collaborative care for depression: what is needed? Am J Manag Care 2014 Sep;20(9):699-707 [FREE Full text] [Medline: 25365745]
9. Carleton KE, Patel UB, Stein D, Mou D, Mallow A, Blackmore MA. Enhancing the scalability of the collaborative care model for depression using mobile technology. Transl Behav Med 2020 Aug 07;10(3):573-579. [doi: 10.1093/tbm/ibz146] [Medline: 32766866]

10. Hallgren KA, Bauer AM, Atkins DC. Digital technology and clinical decision making in depression treatment: Current findings and future opportunities. Depress Anxiety 2017 Jun;34(6):494-501 [FREE Full text] [doi: 10.1002/da.22640] [Medline: 28453916]
11. Moon K, Sobolev M, Kane JM. Digital and mobile health technology in collaborative behavioral health care: Scoping review. JMIR Ment Health 2022 Feb 16;9(2):e30810 [FREE Full text] [doi: 10.2196/30810] [Medline: 35171105]
12. Bauer AM, Iles-Shih M, Ghomi RH, Rue T, Grover T, Kincler N, et al. Acceptability of mHealth augmentation of Collaborative Care: A mixed methods pilot study. Gen Hosp Psychiatry 2018;51:22-29 [FREE Full text] [doi: 10.1016/j.genhosppsych.2017.11.010] [Medline: 29272712]
13. Raney L, Bergman D, Torous J, Hasselberg M. Digitally driven integrated primary care and behavioral health: How technology can expand access to effective treatment. Curr Psychiatry Rep 2017 Sep 30;19(11):86. [doi: 10.1007/s11920-017-0838-y] [Medline: 28965319]
14. Rodriguez-Villa E, Rauseo-Ricupero N, Camacho E, Wisniewski H, Keshavan M, Torous J. The digital clinic: Implementing technology and augmenting care for mental health. Gen Hosp Psychiatry 2020;66:59-66 [FREE Full text] [doi: 10.1016/j.genhosppsych.2020.06.009] [Medline: 32688094]
15. Bauer AM, Thielke SM, Katon W, Unützer J, Areán P. Aligning health information technologies with effective service delivery models to improve chronic disease care. Prev Med 2014 Sep;66:167-172 [FREE Full text] [doi: 10.1016/j.ypmed.2014.06.017] [Medline: 24963895]
16. Dinkel D, Caspari JH, Fok L, Notice M, Johnson DJ, Watanabe-Galloway S, et al. A qualitative exploration of the feasibility of incorporating depression apps into integrated primary care clinics. Transl Behav Med 2021 Sep 15;11(9):1708-1716. [doi: 10.1093/tbm/ibab075] [Medline: 34231855]
17. Marzano L, Bardill A, Fields B, Herd K, Veale D, Grey N, et al. The application of mHealth to mental health: opportunities and challenges. Lancet Psychiatry 2015 Oct;2(10):942-948. [doi: 10.1016/S2215-0366(15)00268-0] [Medline: 26462228]
18. Powell J, Hamborg T, Stallard N, Burls A, McSorley J, Bennett K, et al. Effectiveness of a web-based cognitive-behavioral tool to improve mental well-being in the general population: randomized controlled trial. J Med Internet Res 2012 Dec 31;15(1):e2 [FREE Full text] [doi: 10.2196/jmir.2240] [Medline: 23302475]
19. 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]
20. 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]
21. Kim BY, Lee J. Smart devices for older adults managing chronic disease: A scoping review. JMIR Mhealth Uhealth 2017 May 23;5(5):e69 [FREE Full text] [doi: 10.2196/mhealth.7141] [Medline: 28536089]
22. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. J Med Internet Res 2017 May 10;19(5):e153 [FREE Full text] [doi: 10.2196/jmir.7725] [Medline: 28490417]
23. 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]
24. Sim I. Mobile devices and health. N Engl J Med 2019 Sep 05;381(10):956-968. [doi: 10.1056/NEJMra1806949] [Medline: 31483966]
25. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Privacy and security in mobile health apps: a review and recommendations. J Med Syst 2015 Jan;39(1):181. [doi: 10.1007/s10916-014-0181-3] [Medline: 25486895]
26. 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]
27. Ben-Zeev D, Drake R, Marsch L. Clinical technology specialists. BMJ 2015 Feb 19;350:h945. [doi: 10.1136/bmj.h945] [Medline: 25697164]
28. Bhugra D, Tasman A, Pathare S, Priebe S, Smith S, Torous J, et al. The WPA-Lancet Psychiatry Commission on the future of psychiatry. Lancet Psychiatry 2017 Oct;4(10):775-818. [doi: 10.1016/S2215-0366(17)30333-4] [Medline: 28946952]
29. Hilty D, Chan S, Torous J, Luo J, Boland R. A framework for competencies for the use of mobile technologies in psychiatry and medicine: Scoping review. JMIR Mhealth Uhealth 2020 Feb 21;8(2):e12229 [FREE Full text] [doi: 10.2196/12229] [Medline: 32130153]
30. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462 [FREE Full text] [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
31. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To prompt or not to prompt? A microrandomized trial of time-varying push notifications to increase proximal engagement with a mobile health app. JMIR Mhealth Uhealth 2018 Nov 29;6(11):e10123 [FREE Full text] [doi: 10.2196/10123] [Medline: 30497999]

Abbreviations

BHCM: behavioral health care manager
CoCM: collaborative care model
EHR: electronic health record
mHealth: mobile health
PCP: primary care provider
PHQ-9: Patient Health Questionnaire-9

Edited by T Leung; submitted 28.12.21; peer-reviewed by J Bechtel, D Hilty; comments to author 30.01.22; revised version received 16.05.22; accepted 17.05.22; published 01.07.22.

Please cite as:
Moon KC, Sobolev M, Grella M, Alvarado G, Sapra M, Ball T
An mHealth Platform for Augmenting Behavioral Health in Primary Care: Longitudinal Feasibility Study
JMIR Form Res 2022;6(7):e36021
URL: https://formative.jmir.org/2022/7/e36021
doi:10.2196/36021
PMID:35776491

Original Paper

Exploring a Need for a Cardiometabolic Disease Staging System as a Computerized Clinical Decision Support Tool: Qualitative Study

Aizhan Karabukayeva[1*], MPH; Jami L Anderson[1*], MSc, MBS; Allyson G Hall[1*], PhD; Sue S Feldman[1,2*], MEd, PhD, RN; Tapan Mehta[3*], MSEE, PhD

[1]Department of Health Services Administration, School of Health Professions, University of Alabama at Birmingham, Birmingham, AL, United States
[2]Department of Graduate Medical Education, Heersink School of Medicine, University of Alabama at Birmingham, Birmingham, AL, United States
[3]Department of Family and Community Medicine, Heersink School of Medicine, University of Alabama at Birmingham, Birmingham, AL, United States
[*]all authors contributed equally

Corresponding Author:
Aizhan Karabukayeva, MPH
Department of Health Services Administration
School of Health Professions
University of Alabama at Birmingham
1716 9th Avenue South
Birmingham, AL, 35294
United States
Phone: 1 2055155146
Email: ak1984@uab.edu

Abstract

Background: Although cardiometabolic diseases are leading causes of morbidity and mortality in the United States, computerized tools for risk assessment of cardiometabolic disease are rarely integral components of primary care practice. Embedding cardiometabolic disease staging systems (CMDS) into computerized clinical decision support systems (CDSS) may assist with identifying and treating patients at greatest risk for developing cardiometabolic disease.

Objective: This study aimed to explore the current approach to medical management of obesity and the need for CMDS designed to aid medical management of people living with obesity, at risk of being obese, or diabetic at the point of care.

Methods: Using a general inductive approach, this qualitative research study was guided by an interpretive epistemology. The method included semistructured, in-depth interviews with primary care providers (PCPs) from university-based community health clinics. The literature informed the interview protocol and included questions on PCPs' experiences and the need for a tool to improve their ability to manage and prevent complications from overweight and obesity.

Results: PCPs (N=10) described their current approaches and emphasized behavioral treatments consisting of combined diet, physical activity, and behavior therapy as the first line of treatment for people who were overweight or obese. Results suggest that beneficial features of CDSS include (1) clinically relevant and customizable support, (2) provision of a comprehensive medical summary with trends, (3) availability of patient education materials and community resources, and (4) simplicity and ease of navigation.

Conclusions: Implementation of a CMDS via a CDSS could enable PCPs to conduct comprehensive cardiometabolic disease risk assessments, supporting clinical management of overweight, obesity, and diabetes. Results from this study provide unique insights to developers and researchers by identifying areas for design optimization, improved end user experience, and successful adoption of the CDSS.

(JMIR Form Res 2022;6(7):e37456) doi:10.2196/37456

KEYWORDS

cardiometabolic disease staging system; risk assessment; cardiometabolic disease; clinical decision support system; primary care; obesity; overweight; medical management

Introduction

Cardiometabolic diseases are leading causes of morbidity and mortality in the United States, including a wide array of diseases, typically beginning with insulin resistance and progressing later into a cluster of conditions that increase the risk of type 2 diabetes, stroke, and cardiovascular disease [1,2]. Being overweight (BMI ≥25 kg/m^2) is associated with double the risk of developing cardiometabolic multimorbidity, while having mild and severe obesity (BMI ≥30 kg/m^2) increases the risk 4 and 10 times, respectively [3]. However, current diagnostic categories that are based on standard BMI ranges defining overweight and obesity have high specificity but low sensitivity for identifying insulin resistance and cardiometabolic disease [4]. For example, with the current diagnostic categories, some individuals with overweight and obesity might not have cardiometabolic risk factors and may exhibit low rates of future diabetes and cardiovascular-related mortality; alternatively, some individuals who do not meet criteria for either metabolic syndrome or prediabetes exhibit risk of future diabetes [4]. Thus, risk assessments for cardiometabolic disease with greater sensitivity should be an integral component of medical practice, with tools to evaluate preventive and therapeutic options in patients at greatest risk for developing disease. Currently, there is no stratification of the population by level of obesity-related disease and mortality risk [5].

An estimated 42.5% of US adults aged 20 years and older are living with obesity, including 9.0% with severe obesity, and another 31.1% are overweight [6]. Because this group is at high risk of developing diabetes and other obesity-related complications, there is a need for risk stratification approaches to identify early those at highest risk and identify weight loss programs with appropriate treatment intensity. To provide appropriate medical management of obesity and facilitate the diabetes risk assessment of people with excess adiposity, a comprehensive staging system that establishes 5 stages of cardiometabolic disease risk—the cardiometabolic disease staging system (CMDS)—was developed [7,8]. This validated staging system is based on Adult Treatment Panel III metabolic syndrome risk factors and includes waist circumference, systolic and diastolic blood pressures, fasting and 2-hour blood glucose levels, triglycerides, and high-density lipoprotein cholesterol (HDL-C; Table 1) [4]. The purpose of this system is to help clinicians select treatment modality and intensity in the management of cardiometabolic diseases while balancing benefit and risk. Evidence demonstrates the CMDS has higher predictive and discriminative ability compared with other systems and relies on data typically collected during primary care visits; thus, it is more feasible to integrate into busy workflows of primary care providers (PCPs) [5].

Table 1. The cardiometabolic disease staging (CMDS) system.

Stage	Descriptor	Criteria
Stage 0	Metabolically healthy	No risk factors
Stage 1	One or two risk factors	Have 1 or 2 of the following risk factors: 1. High waist circumference (≥112 cm in men and ≥88 cm in women) 2. Elevated blood pressure (systolic ≥130 mm Hg and/or diastolic ≥85 mm Hg) or on antihypertensive medication 3. Reduced serum HDL-C[a] (<1.0 mmol/L or 40 mg/dL in men; <1.3 mmol/L or 50 mg/dL in women) or on medication 4. Elevated fasting serum triglycerides (≥1.7 mmol/L or 150 mg/dL) or on medication
Stage 2	Metabolic syndrome or prediabetes	Have only 1 of the following 3 conditions in isolation: 1. Metabolic syndrome based on 3 or more of 4 risk factors: high waist circumference, elevated blood pressure, reduced HDL-C, and elevated triglycerides 2. Impaired fasting glucose (IFG; fasting glucose ≥5.6 mmol/L or 100 mg/dL) 3. Impaired glucose tolerance (IGT; 2-h glucose ≥7.8 mmol/L or 140 mg/dL)
Stage 3	Metabolic syndrome + prediabetes	Have any 2 of the following 3 conditions: 1. Metabolic syndrome 2. IFG 3. IGT
Stage 4	T2DM[b] and/or CVD[c]	Have T2DM and/or CVD: 1. T2DM (fasting glucose ≥126 mg/dL or 2-h glucose ≥200 mg/dL or on antidiabetic therapy) 2. Active CVD (angina pectoris or status post a CVD event such as acute coronary artery syndrome, stent placement, coronary artery bypass, thrombotic stroke, nontraumatic amputation due to peripheral vascular disease)

[a]HDL-C: high-density lipoprotein cholesterol.

[b]T2DM: type 2 diabetes mellitus.

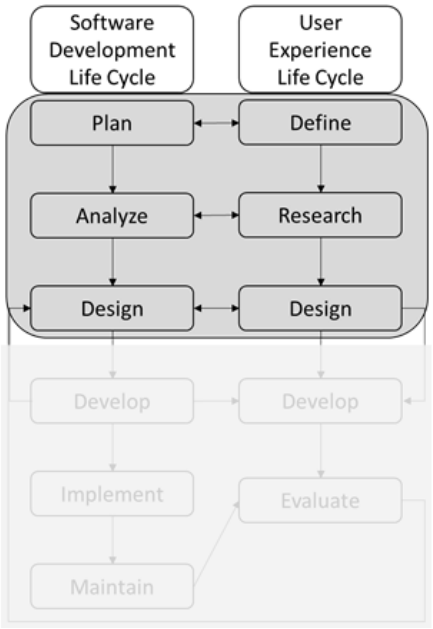
[c]CVD: cardiovascular disease.

Vigilance in the management of modifiable risk factors is critical, given that people with overweight and obesity are at increased cardiovascular risk. Primary care settings, as familiar and accessible clinical venues for patients, are well positioned to screen people with overweight and obesity and recommend appropriate weight loss treatment plans to prevent complications and weight progression. Many studies found that the largest weight losses were achieved with high-intensity counseling by PCPs and referral of interested individuals to appropriate interventions [9-13]. However, a study of a nationally representative sample of adults aged 35 years and older found that, despite more adults reported being screened for obesity (78.6%) and of those screened, nearly 40% had a BMI of 30 kg/m^2 or higher (39.2%), only slightly more than one-half (53.5%) of obese adults screened reported receiving counseling about weight management [14]. Furthermore, BMI is the most preferred screening tool, though literature indicates it could be a poor indicator of cardiovascular disease and overall mortality risk [7,15]. Research finds BMI is not a good index of visceral fat, which is the basis of metabolic disorders associated with increased cardiovascular risk, whereas waist circumference might be superior as a risk assessment tool [16]. PCP-indicated practice improvements, helpful in treating and managing overweight and obesity, include better tools for early identification of risk and preventive treatment for those with multiple risk factors [11].

Providing CMDS to PCPs via computerized clinical decision support systems (CDSS) may assist in stratifying the population by obesity-related disease risk and targeting those patients who are at greater risk for obesity-related complications. To the authors' knowledge, this would be the first electronic health record–integrated CDSS that would incorporate CMDS. Despite literature indicating CDSS may have a positive impact on provider performance and patient outcomes [17], evidence also indicates that CDSS rarely reach their full potential [18]. As with any innovation, user acceptance and integration within the clinical workflow are critical for successful uptake and routine use [19].

System analysis and design involve the process of planning, analyzing, designing, developing, implementing, and maintaining systems. A user-centered approach focusing on the user experience necessitates coordinated relationships between the system specialists, designers, and developers and the nonspecialists and users with outcomes knowledge. The system development life cycle, when combined with the user experience life cycle, allows for that coordination to occur and has been shown to lead to better system adoption [20]. Figure 1 illustrates our conceptual model for system analysis and design of the CMDS. This paper reports on the first 3 phases of each cycle: (1) plan and define, (2) analyze and research, and (3) design. As such, with the aim of involving users at key milestone stages of system development, this study explored the current approach to management of overweight and obesity and a need for the CMDS system at the point of care to facilitate specificity in treatment modalities.

Figure 1. Conceptual model for system analysis and design.



Methods

To ensure we adhered to qualitative reporting standards, we followed the 32-time consolidated criteria for reporting qualitative studies (COREQ) checklist (Multimedia Appendix 1).

Sampling

Participant recruitment used convenience sampling where the research team coordinated with the medical director for primary care of a large academic medical center in the southeastern United States. Recruitment emails to potential candidates indicated the study purpose and invited participation. The number of participants was determined to be sufficient when saturation was reached (N=10) [21]. All participants were

provided with a US $100 gift card. The present study is a foundation for ongoing research aimed at developing and implementing a CDSS based on the CMDS.

Data Collection

From August 2020 to January 2021, 10 semistructured interviews were conducted by 2 research team members, consisting of the principal investigator (TM; male) and a graduate research assistant (AK; female). Two senior female researchers (AH and SF) with training in qualitative interviewing provided guidance and supervision. The interviewers did not have prior relationships with the participants. Only the interviewers and participants were present during data collection. The duration of the interviews varied between 30 minutes and 45 minutes and were conducted via a collaborative, cloud-based videoconferencing service at a mutually agreed-upon time. Interviews were transcribed verbatim by a commercial transcription company. The interview guide was informed by the literature review and included questions designed to (1) understand how PCPs manage overweight and obesity and facilitate prevention and management of diabetes and cardiovascular disease risk during a standard primary care visit and (2) explore PCP needs for CMDS and preferences for a CDSS (Multimedia Appendix 2). Broad, open-ended questions along with permissive prompts were used to facilitate each semistructured interview. Prior to conducting interviews, the semistructured interview guide was pilot tested with several providers to ensure questions were clear, generated in-depth discussion, were acceptable to participants, and resulted in usable information. Feedback from pilot testing was used to modify the wording, content, and order of the interview questions.

Ethical Considerations

All investigations were conducted in conformity with ethical principles of research. Consent for participation and interview recording was obtained verbally before each interview. This study was determined to be exempt by the University of Alabama at Birmingham Institutional Review Board (IRB Protocol Number 300003559).

Data Analysis

Transcribed interviews were coded using an inductive thematic analysis approach with NVivo 12 Plus (QSR International, Melbourne, Australia). To increase reliability and reduce bias, all transcripts were coded by 2 team members (AK and JA) with expertise in thematic analysis [22,23]. The analysis consisted of 2 phases: codebook development and codebook refinement. First, during open coding, coders examined an initial set of transcripts for categories (processes or events that share an attribute) of information related to our research questions. The second phase of our analysis focused on comparing and applying our initial codes to both existing and new data generated from subsequent interviews. This constant comparative analysis [24] across data sets allowed merging and clarifying codes. Following the initial coding process, research team members (AK and JA) discussed questions and discrepancies until 95% agreement was reached. Then, coders identified key points and recurring categories and themes that were central to the experience described by the participants. The process consisted of both coders dividing the text into semantic segments, labelling the segments with codes, together examining the codes for overlap and redundancy, and aggregating these codes into broader categories and themes [25].

Results

Sample Characteristics and Suggestions

We recruited 10 PCPs (7 physicians and 3 certified registered nurse practitioners) with practice experience ranging from 3 years to 43 years, with a mean of 12.2 years. Out of 10 respondents, 4 were male, and 6 were female. The most common practice-based barriers included lack of time and knowledge of resources, including access to evidence-based medical models and affordable community options. Considering the results of this study, 4 factors emerged as important for consideration in the development of a CDSS for metabolic conditions: (1) clinically relevant and customizable information delivery, (2) provision of comprehensive medical summary with trends, (3) availability of patient education materials and community resources, and (4) simplicity and ease of navigation. Table 2 describes the key suggestions voiced by the PCPs for future design of the CDSS to be successfully adopted.

Table 2. Suggestions from primary care providers regarding preferrable clinical decision support system features.

Suggestions	Quotes
Speed of the information technology	"The other thing would be – does it run efficient? There are parts of Cerner that literally if you click the button, you're going be sitting there for 2 minutes just waiting, waiting, and waiting." [Primary care physician, male]
Synthesis of available information	"I think what would be good is if you had a piece of software that could extract that [lab] data out of the record. And then you could click on a button at the top of the record, and it said 'weight management'. If you click, it would have drop down algorithm and it was connected to the orders." [Primary care physician, male]
Fit in the workflow	"So, whatever you come up with has to be something that's integrated and uses the data that's there, and gives you immediate feedback. It can't be something that takes three minutes to enter the data." [Primary care physician, male]
User-friendly with minimalist design	"So, ideally something self-contained, within the same page gives me kind of risk information and recommendations based of that, especially if it could be set up such that off of that page, I could directly order things. That would be amazing." [Primary care physician, female]
Flexibility	"I think you definitely need to maintain the ability to customize or edit because, again, these are just sort of recommendations and sort of a part of the picture that the risk calculator gives you, but, you know, as long as you know, you could sort of edit to customize and individualize to a patient." [Primary care physician, female]
Justification of treatment based on guidelines	"If there was something to standardize [management of] obesity and would give you a quantifiable number that puts them at a higher risk factor. So, if there was something that took in more either genetic versus biological markers that could be influential, I think that would be very useful and something that we would definitely want to implement and make it more of a standardization and not just an extra research tool." [Primary care physician, male]

Current Management Practice of Obesity in PCP Clinics

Focus Not on Prevention But on Comorbidities

Almost all respondents reported that a significant portion of their patient populations was overweight, and they also noted that about 60% to 70% of patients had hypertension, diabetes, or other comorbidities. Even young populations presenting to primary care tended to have elevated BMIs or abnormal glucose levels. However, the respondents noted that they gave priority to management of the comorbidities rather than focusing on prevention and management of obesity. Respondents also noted they did not routinely use pharmacologic treatments for overweight or obesity but more to treat comorbidities, such as hypertension or elevated blood glucose levels.

> *About 80% of the patients I see are going to have one of the three: hypertension, diabetes, or obesity; and probably safe to say 60-70% definitely have all three. A lot of what I'm seeing is for blood pressure and diabetes management specifically.* [Certified registered nurse practitioner, female]

> *I try not to [prescribe medicine]. I have a fair number of patients that would like a pill to fix their weight problem. And sometimes they break me down and I do [prescribe medicine]. If I prescribe something, an appetite suppressant to help lose weight, it's under the premise that it's very short term, no more than three months.* [Internist, female]

> *Probably due to the acuity of our patient population, I feel like by the time people get to us specifically in the health system we're working under, there are usually a lot more problems that we're juggling, and obesity is always important, but it's probably like number 10 on the list of concern.* [Certified registered nurse practitioner, male]

BMI as a Main Diagnostic Measure

According to the respondents, BMI remains the primary tool for assessing obesity, as it is easy to access, is affordable to measure, and can conveniently be used to monitor weight changes. Additionally, participants responded that waist circumference measurement has not been integrated into routine practice. Patient risk factors associated with being identified or diagnosed as overweight or obese by their physician included higher BMI, family history, lifestyle, and habits. Respondents noted that they provided metabolic screening depending on patient's BMI, including blood glucose and blood lipids levels.

> *We don't measure it [waist circumference] in our clinic. We do have the BMI. So, the first point is BMI- this is all I look at because that's what I have available, and it's just a measure of numbers and the calculation. So, it's easy.* [Internist, female]

> *I look at their medical and family history. Like if they have diabetes in the family, then obviously that puts them at a higher risk automatically. Or if they have family members with hypertension. So, family history is very important for my understanding.* [Internist, female]

Reliance on Lifestyle Modifications

Most of the respondents' approaches to weight management were limited to assessing physical activity and assessing readiness for change, dietary habits, and expectations. The most common recommendations were to increase physical activity and dietary changes. Interestingly, half (5/10, 50%) of the

respondents noted they did not have any formalized treatment plan to manage overweight or obesity and did not follow specific treatment guidelines. In addition, there was limited use of external sources of weight management support, with only few patients being referred to weight loss clinics, mainly due to limited coverage of services by health insurance companies. External resources frequently included a nutritionist and a commercial weight loss program (eg, Weight Watchers).

> *I do not have sort of very specific treatment guidelines. I'm not saying I'd be opposed to that. I just have my own practice at this time. It really will depend, because I'm trying to gauge a person's willingness to change, so I will certainly ask some typical, open-ended questions about what have they tried in the past. It really becomes an individualized approach.* [Internist, male]

> *I would say exclusively exercise and diet. More recently, I [started] referring patients with BMI 30 or higher with multiple comorbidities to the weight loss management program.* [Primary care physician, female]

Lack of Knowledge About Referral Options in the Community

Respondents agreed that resources for intense lifestyle intervention and social support were important for the patients; however, respondents also noted the lack of knowledge about referral options in the community, including commercial-based programs. Because of the range in the socioeconomic status of their patient population, respondents expressed wanting point-of-care information about various affordable and convenient options that would be readily accessible and affordable for patient engagement.

> *I think there are a lot of resources out there but to be honest with you, I don't think that we really know where the resources are. Everyone, probably, has their own little list of resources that they use, I think.* [Primary care provider, male]

> *I don't think I necessarily have a good handle on it [the local resources]. I have certain things that I would say as a 'go to' that are probably out of date and missing a lot of some of the newer [resources].* [Certified registered nurse practitioner, female]

Lack of Patient Education Literature

One of the challenges voiced by respondents was lack of appropriate, "meet them where they are" weight management educational materials accessible for use at the point of care or after consultation.

> *Time during a visit is at a premium. In theory, our visits are 20 minutes, by the time the patient gets here, checks in, and triaged, I generally have about seven minutes out of 20 minutes to see a patient. Maybe a little bit more, sometimes a little less.* [Primary care physician, male]

> *The printed materials are not very good that we have available. They are not very helpful. That is why I don't give them out very often.* [Primary care physician, female]

Need for CMDS

Need for a Risk Stratification Tool Embedded Into the CDSS

Almost all respondents (9/10, 90%) expressed an interest in having a CDSS that would incorporate diabetes and cardiovascular disease risk assessment and, based on the risks, outline a treatment plan. As respondents noted, the advantage of using a CDSS would be providing "legitimate justification" for a treatment plan with an assumption that patients understand their risk and the reason for the proposed treatment. To the authors' knowledge, there is no decision support system available to assist providers in evidence-based weight loss treatment intensification. There is, however, a diabetes management protocol that has been developed but is not part of the electronic health records at this institution.

> *Currently, I have to pick up my phone, get on my coronary app and then put all the information. So, you could see where a tool like this that is incorporating the coronary risk score would be quite helpful built within Cerner. If it could even populate the data that we have with more recent blood pressure, that would be even more useful.* [Primary care physician, male]

> *Ideally, it would be something that I could just turn the computer monitor and show the patient, saying "Okay, well, this is why I'm recommending it. Your A1C is 5.9, up from 5.6 last year. Your cholesterol is up, your weight is up. So, this gives you a 17% chance of diabetes in the next two years. And these are the steps that we recommend...* [Primary care physician, male]

Diagnostic-Supported CDSS

Respondents noted a need for an CDSS that would consider diagnostics, such as relevant patient data and lab results. In addition, respondents indicated it would be useful to have access to clinically meaningful trends and track risk scores for complications. A majority of respondents manually calculate various risk scores, such as the 10-year Atherosclerotic Cardiovascular Disease (ASCVD) risk; therefore, embedding such a calculator in the CDSS could increase efficiency and reduce error.

> *If we had a good algorithm based on BMI and any potential risk factors that was easy to follow and implement, with good handouts and appropriate referral or community resources there, and if we could collate that information, I think it could be helpful.* [Primary care physician, male]

> *The main risk calculator that I use is the 10-year cardiovascular risk when I'm trying to decide if somebody should be on a statin. I just have on my phone and I just pull it up when I get their [patients'] labs back and plug in the numbers. Risk calculator*

would be helpful to have [embedded] in the computer as well. [Primary care physician, female]

Incorporating Evidence-Based Practice

Respondents thought that having a CDSS that incorporates evidence-based clinical guidelines for management of obesity or overweight, both medical and behavioral, and that provides intervention recommendations would standardize and streamline the care provided and interventions suggested to their patients. The general idea was that such a system might help assist with managing patients with required tests, follow-up appointments, and preventive care.

I wouldn't say that I could speak for the whole clinic, I may just not be getting something that everyone else is doing. But we would be very open to having a tool that brings a standardization and also makes sure you're not overlooking anything and following a best practice guideline on initial management, and then also routine follow up. [Internist, female]

I think what would be most beneficial is having suggested treatment plans. It would help me to know that I am on the right track if I had a treatment plan that was suggested based upon their [patient] other chronic co-morbidities and their current A1C results or current blood sugar trends. [Certified registered nurse practitioner, female]

Ability to Have Resources to Make Referrals and Educate Patients

Respondents expressed a strong interest in information about accessible and affordable resources in the local community for patients struggling with overweight or obesity. They believed that it would enhance patient engagement and motivate patients toward behavior change. In addition, several respondents suggested that, if the CDSS had the means to efficiently provide appropriate educational materials to patients, it could improve the patient's participation in their own care:

...having good patient-friendly handout material that was easy to attach in the patient portal in terms of a new diet, recommendations based on ADA or other more popular diets like the DASH (Mediterranean Diet) or other types of things based on the patient's history. I do think that would be helpful. [Primary care physician, female]

Maybe [CDSS] gives you an option that you can click on, like option A "Would this person be interested in nutritional counseling?", option B "Do you want to print this list of printouts to give them and present to them during your clinic visit?", or "Would they want a referral to weight loss clinic?" I mean, it would be awesome if we had some way to refer people to some community resource near them where they could be contacted and offered some kind of like exercise class or a way to get into a walk group or something. [Certified registered nurse practitioner, female]

Discussion

Principal Findings

In this study, we sought to understand and capture user requirements for a system that evaluates the stage and severity of cardiometabolic disease that would be incorporated into a CDSS. The idea to involve intended users early in the design process is well supported in the literature [26] and results in aligning user expectations with the resulting functionality [27]. Currently, the PCPs' approach to management of overweight and obesity largely focuses on treating comorbidities and counseling lifestyle modifications such as diet and exercise. There is limited use of medications to combat overweight and obesity. Although there are existing guidelines for obesity treatment [28] and related cardiometabolic conditions such as diabetes [29] and hypertension [30], our respondents were not consistently using them for diagnosis or treatment purposes.

Comparison With Prior Work

Findings from this study are consistent with a growing body of literature on how PCPs manage overweight and obesity, as well as on what CDSS features increase likelihood of its uptake. Turner et al [31] found, among a nationally representative sample of active health care providers, that (1) knowledge of physical activity and dietary guidelines was limited and (2) understanding of the appropriate initiation, intensity, and duration of pharmacotherapy was often inconsistent with evidence-based guidelines. Another study found that PCPs were least likely to say they would prescribe medication or refer a patient to counseling [32]. Regarding CDSS features, a systematic review by Groenhof et al [33] found that design and usability were important drivers behind the success, noting that information should be displayed all at once and at one glance. In addition, the lack of insight into the automated computation and source of information decreased user satisfaction. Further, the most recent systematic review by Kouri et al [34] identified important CDSS features that significantly predict uptake, such as averting the need for provider data entry by mining patient data from within electronic health record systems to inform CDSS.

Implication for Practice

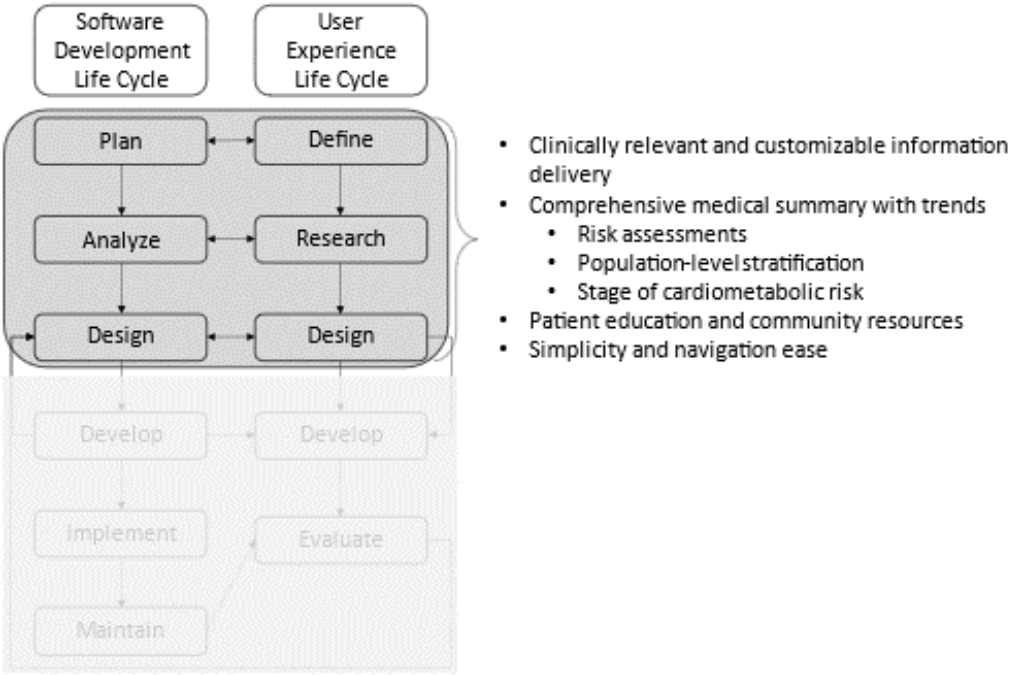
An important consideration would be providing easy access to the latest evidence-based clinical standards and protocols by embedding them in the CDSS. As a first step, CDSS could include measurement tools to perform a diagnostic evaluation based on evidence-based guidelines. If currently only BMI is considered, a more comprehensive evaluation must include additional measures such as measurement of waist circumference. Further, for the treatment, the physicians should be able to obtain clinical decision support by using CDSS to analyze pertinent information about the patient's current clinical condition, including information about medication, lab results, and treatment compliance. Given support for the CDSS among our sample, we propose a design of a CDSS that provides suggestions for treating the primary and augmenting medications with explanations. For the purposes of follow-up, the CDSS should have reminders to ensure the important considerations are not overlooked. Moreover, it could also recommend and

display when the patient should return for a visit. All entries should be automatically stored, providing electronic documentation and record keeping, thus providing access to complete patient information. Overall, information about a patient's demographic characteristics and other clinical records should be accessible by a single click.

Implications for Development

Results from this study were used to better understand user requirements within a parallel system analysis and design framework (see Figure 1), the importance of which was to ensure the voice of the user was adequately and accurately represented [20]. In this phase, we present the conceptual framework with the findings applied as high-level categories (see Figure 2). These categories and the details behind them as presented throughout this study will be used to inform the evaluation.

Figure 2. The conceptual framework with the findings as high-level categories.



Limitations

There are several limitations of this study. First, the 10 participants, predominantly non-Hispanic White (9/10, 90%), represented PCPs at a large academic center in the southeast. A small homogenous sample size could potentially limit the generalizability of our findings, and we recognize the need to add breadth and depth to this participant sample as development ensues. In addition, the strength of the study was our consideration of assuring intercoder reliability. Thus, we feel confident that we are correctly representing the voices of our participants.

Conclusion

Implementation of a CMDS system in the form of a CDSS could be used as a risk assessment tool that also provides risk-based and evidence-based treatment or program recommendations to better manage overweight and obesity and prevent diabetes. Results from this study provide unique insight to developers and researchers to identify areas for design optimization for improved end user experience to ensure successful adoption of the CDSS.

Acknowledgments

The study described was supported by a pilot and feasibility award to TM from the University of Alabama at Birmingham (UAB) Diabetes Research and funded by P30DK079626 from the National Institute of Diabetes and Digestive and Kidney Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Diabetes and Digestive and Kidney Diseases or the National Institutes of Health.

The authors wish to thank UAB Prime Care leadership and providers who participated and facilitated this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
COREQ checklist.
[PDF File (Adobe PDF File), 444 KB - formative_v6i7e37456_app1.pdf]

Multimedia Appendix 2
Interview guide.
[DOCX File , 23 KB - formative_v6i7e37456_app2.docx]

References

1. Johnson M. Diet and nutrition: implications to cardiometabolic health. J Cardiol and Cardiovasc Sciences 2019 Mar 1;3(2):4-9. [doi: 10.29245/2578-3025/2019/2.1168]
2. Tasic I, Lovic D. Hypertension and cardiometabolic disease. Front Biosci (Schol Ed) 2018 Jan 01;10(1):166-174 [FREE Full text] [doi: 10.2741/s506] [Medline: 28930524]
3. Kivimäki M, Kuosma E, Ferrie JE, Luukkonen R, Nyberg ST, Alfredsson L, et al. Overweight, obesity, and risk of cardiometabolic multimorbidity: pooled analysis of individual-level data for 120 813 adults from 16 cohort studies from the USA and Europe. The Lancet Public Health 2017 Jun;2(6):e277-e285. [doi: 10.1016/S2468-2667(17)30074-9]
4. Guo F, Moellering DR, Garvey WT. The progression of cardiometabolic disease: validation of a new cardiometabolic disease staging system applicable to obesity. Obesity (Silver Spring) 2014 Jan 05;22(1):110-118 [FREE Full text] [doi: 10.1002/oby.20585] [Medline: 23894121]
5. Ejima K, Xavier NA, Mehta T. Comparing the ability of two comprehensive clinical staging systems to predict mortality: EOSS and CMDS. Obesity (Silver Spring) 2020 Feb 06;28(2):353-361. [doi: 10.1002/oby.22656] [Medline: 31905265]
6. Fryar CD, Carroll MD, Afful J. Prevalence of Overweight, Obesity, and Severe Obesity Among Adults Aged 20 and Over: United States, 1960–1962 Through 2017–2018. National Center for Health Statistics. 2020 Dec. URL: https://www.cdc.gov/nchs/data/hestat/obesity-adult-17-18/overweight-obesity-adults-H.pdf [accessed 2022-06-24]
7. Guo F, Garvey WT. Development of a weighted cardiometabolic disease staging (CMDS) system for the prediction of future diabetes. J Clin Endocrinol Metab 2015 Oct;100(10):3871-3877 [FREE Full text] [doi: 10.1210/jc.2015-2691] [Medline: 26241327]
8. Wilkinson L, Yi N, Mehta T, Judd S, Garvey WT. Development and validation of a model for predicting incident type 2 diabetes using quantitative clinical data and a Bayesian logistic model: A nationwide cohort and modeling study. PLoS Med 2020 Aug 7;17(8):e1003232 [FREE Full text] [doi: 10.1371/journal.pmed.1003232] [Medline: 32764746]
9. Tronieri JS, Wadden TA, Chao AM, Tsai AG. Primary care interventions for obesity: review of the evidence. Curr Obes Rep 2019 Jun 19;8(2):128-136 [FREE Full text] [doi: 10.1007/s13679-019-00341-5] [Medline: 30888632]
10. Feldman SS, Cochran RA, Mehta T. Predictors of weight change: findings from an employee wellness program. Front Endocrinol (Lausanne) 2019 Feb 19;10:77 [FREE Full text] [doi: 10.3389/fendo.2019.00077] [Medline: 30837948]
11. Tucker CM, Williams JL, Wippold GM, Bilello LA, Morrissette TA, Good AJ, et al. Views of diverse primary care patients on the roles of healthcare providers and staff and the influence of other variables in their weight management. Clin Obes 2018 Feb 20;8(1):11-20 [FREE Full text] [doi: 10.1111/cob.12225] [Medline: 29052345]
12. Post RE, Mainous AG, Gregorie SH, Knoll ME, Diaz VA, Saxena SK. The influence of physician acknowledgment of patients' weight status on patient perceptions of overweight and obesity in the United States. Arch Intern Med 2011 Feb 28;171(4):316-321. [doi: 10.1001/archinternmed.2010.549] [Medline: 21357807]
13. Pool AC, Kraschnewski JL, Cover LA, Lehman EB, Stuckey HL, Hwang KO, et al. The impact of physician weight discussion on weight loss in US adults. Obes Res Clin Pract 2014 Mar;8(2):e131-e139 [FREE Full text] [doi: 10.1016/j.orcp.2013.03.003] [Medline: 24743008]
14. Shafer PR, Borsky A, Ngo-Metzger Q, Miller T, Meyers D. The practice gap: national estimates of screening and counseling for alcohol, tobacco, and obesity. Ann Fam Med 2019 Mar 11;17(2):161-163 [FREE Full text] [doi: 10.1370/afm.2363] [Medline: 30858260]
15. Powell-Wiley TM, Poirier P, Burke LE, Després J, Gordon-Larsen P, Lavie CJ, et al. Obesity and cardiovascular disease: a scientific statement from the American Heart Association. Circulation 2021 May 25;143(21):1. [doi: 10.1161/cir.0000000000000973]
16. Marinou K, Tousoulis D, Antonopoulos AS, Stefanadi E, Stefanadis C. Obesity and cardiovascular disease: from pathophysiology to risk stratification. Int J Cardiol 2010 Jan 07;138(1):3-8. [doi: 10.1016/j.ijcard.2009.03.135] [Medline: 19398137]
17. Taheri Moghadam S, Sadoughi F, Velayati F, Ehsanzadeh SJ, Poursharif S. The effects of clinical decision support system for prescribing medication on patient outcomes and physician practice performance: a systematic review and meta-analysis. BMC Med Inform Decis Mak 2021 Mar 10;21(1):98 [FREE Full text] [doi: 10.1186/s12911-020-01376-8] [Medline: 33691690]
18. Haynes RB, Wilczynski NL, Computerized Clinical Decision Support System (CCDSS) Systematic Review Team. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: methods of a

decision-maker-researcher partnership systematic review. Implement Sci 2010 Feb 05;5(1):12 [FREE Full text] [doi: 10.1186/1748-5908-5-12] [Medline: 20181104]
19. Doebbeling BN, Saleem J, Haggstrom D, Militello L, Flanagan M, Arbuckle N, et al. Integrating Clinical Decision Support into Workflow. Indiana University–Purdue University Indianapolis. 2011. URL: https://scholarworks.iupui.edu/handle/1805/6661 [accessed 2022-06-24]
20. Anderson J, Willson P, Peterson NJ, Murphy C, Kent TA. Prototype to practice: Developing and testing a clinical decision support system for secondary stroke prevention in a veterans healthcare facility. Comput Inform Nurs 2010;28(6):353-363. [doi: 10.1097/NCN.0b013e3181f69c5b] [Medline: 20978406]
21. Patton MQ. Q. In: Qualitative Research & Evaluation Methods: Integrating theory and practice. Newbury Park, CA: SAGE Publications Inc; 2014.
22. Burla L, Knierim B, Barth J, Liewald K, Duetz M, Abel T. From text to codings: intercoder reliability assessment in qualitative content analysis. Nurs Res 2008;57(2):113-117. [doi: 10.1097/01.NNR.0000313482.33917.7d] [Medline: 18347483]
23. Campbell JL, Quincy C, Osserman J, Pedersen OK. Coding in-depth semistructured interviews. Sociological Methods & Research 2013 Aug 21;42(3):294-320. [doi: 10.1177/0049124113500475]
24. Olson J, McAllister C, Grinnell L, Gehrke Walters K, Appunn F. Applying constant comparative method with multiple investigators and inter-coder reliability. TQR 2016 Jan 4:1. [doi: 10.46743/2160-3715/2016.2447]
25. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Newbury Park, CA: SAGE Publications Inc; 2012.
26. Shah SGS, Robinson I, AlShawi S. Developing medical device technologies from users' perspectives: A theoretical framework for involving users in the development process. Int J Technol Assess Health Care 2009 Oct 22;25(4):514-521. [doi: 10.1017/s0266462309990328]
27. Cresswell K, Morrison Z, Crowe S, Robertson A, Sheikh A. Anything but engaged: user involvement in the context of a national electronic health record implementation. Inform Prim Care 2011 Jul 01;19(4):191-206 [FREE Full text] [doi: 10.14236/jhi.v19i4.814] [Medline: 22828574]
28. Wharton S, Lau DC, Vallis M, Sharma AM, Biertho L, Campbell-Scherer D, et al. Obesity in adults: a clinical practice guideline. CMAJ 2020 Aug 04;192(31):E875-E891 [FREE Full text] [doi: 10.1503/cmaj.191707] [Medline: 32753461]
29. American Diabetes Association Professional Practice Committee. 1. Improving care and promoting health in populations: standards of medical care in diabetes-2022. Diabetes Care 2022 Jan 01;45(Suppl 1):S8-S16. [doi: 10.2337/dc22-S001] [Medline: 34964872]
30. Unger T, Borghi C, Charchar F, Khan NA, Poulter NR, Prabhakaran D, et al. 2020 International Society of Hypertension global hypertension practice guidelines. Hypertension 2020 Jun;75(6):1334-1357. [doi: 10.1161/hypertensionaha.120.15026]
31. Turner M, Jannah N, Kahan S, Gallagher C, Dietz W. Current knowledge of obesity treatment guidelines by health care professionals. Obesity (Silver Spring) 2018 Apr 23;26(4):665-671. [doi: 10.1002/oby.22142] [Medline: 29570250]
32. Zatz LY, Hersh E, Gudzune KA, Thorndike AN, N Goldenberg M, Bleich SN. Physicians' political party affiliation and clinical management of obesity. Clin Obes 2020 Oct 16;10(5):e12396. [doi: 10.1111/cob.12396] [Medline: 32803863]
33. Groenhof TKJ, Asselbergs FW, Groenwold RHH, Grobbee DE, Visseren FLJ, Bots ML, UCC-SMART study group. The effect of computerized decision support systems on cardiovascular risk factors: a systematic review and meta-analysis. BMC Med Inform Decis Mak 2019 Jun 10;19(1):108 [FREE Full text] [doi: 10.1186/s12911-019-0824-x] [Medline: 31182084]
34. Kouri A, Yamada J, Lam Shin Cheung J, Van de Velde S, Gupta S. Do providers use computerized clinical decision support systems? A systematic review and meta-regression of clinical decision support uptake. Implement Sci 2022 Mar 10;17(1):21 [FREE Full text] [doi: 10.1186/s13012-022-01199-3] [Medline: 35272667]

Abbreviations

ASCVD: Atherosclerotic Cardiovascular Disease
CDSS: computerized clinical decision support system
CMDS: cardiometabolic disease staging
COREQ: consolidated criteria for reporting qualitative studies
HDL-C: high-density lipoprotein cholesterol
PCP: primary care provider
UAB: University of Alabama at Birmingham

Edited by A Mavragani; submitted 21.02.22; peer-reviewed by OO Oyewole, S Sabarguna; comments to author 30.04.22; revised version received 02.06.22; accepted 16.06.22; published 01.07.22.

Please cite as:
Karabukayeva A, Anderson JL, Hall AG, Feldman SS, Mehta T
Exploring a Need for a Cardiometabolic Disease Staging System as a Computerized Clinical Decision Support Tool: Qualitative Study
JMIR Form Res 2022;6(7):e37456
URL: https://formative.jmir.org/2022/7/e37456
doi:10.2196/37456
PMID:35776499

Original Paper

Improvements in Depression Outcomes Following a Digital Cognitive Behavioral Therapy Intervention in a Polychronic Population: Retrospective Study

Aarathi Venkatesan[1], PhD; Benjamin Forster[1], BA; Prasanna Rao[1], BS, MS; Melissa Miller[1], BA, MS, MBA; Michael Scahill[1], MD, MBA

Vida Health, San Francisco, CA, United States

Corresponding Author:
Aarathi Venkatesan, PhD
Vida Health
100 Montgomery St, Ste 750
San Francisco, CA, 94104
United States
Phone: 1 14159891017
Email: aarathi.venkatesan@vida.com

Abstract

Background: Digital mental health interventions have shown promise in reducing barriers to effective care for depression. Depression and related mental disorders are known to be highly comorbid with common chronic physical conditions, such as obesity and type 2 diabetes. While some research has explored the interaction dynamics of treating populations living with both mental and physical disorders, very little is known about such dynamics in digital care.

Objective: We aimed to examine the effectiveness of a 12-week, therapist-supported, app-based cognitive behavioral therapy program in improving symptoms of depression and anxiety. The studied population included adults with a heavy burden of chronic physical disease, including obesity and type 2 diabetes.

Methods: A total of 1512 participants with at least moderate depression were enrolled. The treatment cohort consisted of 831 (54.96%) participants who completed a follow-up assessment. The program included structured lessons and tools (ie, exercises and practices) and offered one-on-one weekly video counseling sessions with a licensed therapist for 12 weeks and monthly sessions thereafter. The clinically validated 8-item Patient Health Questionnaire (PHQ-8) and the 7-item Generalized Anxiety Disorder scale (GAD-7) were used to assess depression and anxiety, respectively. Linear mixed-effects modeling was employed to examine changes in depression and anxiety over time. Given correlation among various measures of program usage, a composite variable for depth of usage was used to analyze the correlation between usage and changes in depressive symptoms. Body weight changes from baseline were assessed primarily with digitally connected scales.

Results: Out of 831 participants in the treatment cohort, 74.5% (n=619) showed a clinically significant reduction in depressive symptom severity after 12 weeks, where follow-up PHQ-8 scores had shifted downward by at least one diagnostic category. In total, 67.5% (n=561) of the participants showed a reliable improvement in PHQ-8 scores as measured by the reliable change index. There was an average reduction of 5.9 (SD 5.2) points (P<.001) between baseline and follow-up. Greater program usage was correlated with greater likelihood of reliable improvement in depressive symptoms (odds ratio 1.3, 95% CI 1.1-1.5; P=.002). An exploratory analysis of body weight changes with a multilevel, mixed-effect model suggested that reliable improvement in depressive symptoms at follow-up was associated with significantly greater weight loss at 9 months (β=–1.11, P=.002).

Conclusions: The results provide further support that digital interventions can support clinically meaningful improvements in depression. Some form of synergy in treatment of comorbid depression and obesity or diabetes could be studied in future research. The study was limited by postintervention participant attrition as well as the retrospective observational study design.

(JMIR Form Res 2022;6(7):e38005) doi:10.2196/38005

KEYWORDS

depression; anxiety; CBT; digital mental health intervention; cognitive behavioral therapy; digital health; obesity; diabetes; mental health

Introduction

In 2016, mental health disorders affected more than 1 billion people worldwide [1]. Among mental health disorders, clinical depression carries a lifetime risk of 15% to 18% [2] and has the strongest association with disability-adjusted life years, a global benchmark of disease burden [1,2]. Depression is also the psychiatric condition most strongly associated with suicide [3]. The relationships among clinical depression, well-being, and health care costs have been well documented, with costs of care for patients living with depression higher across spending categories [4,5].

It has been estimated from nationally representative surveys that the prevalence of depressive symptoms in the United States increased 3-fold during the COVID-19 pandemic, with the majority of those affected already having been at elevated risk [6]. Simultaneously, the pandemic has accelerated the adoption of digital technology to deliver health care, particularly for mental health, and highlighted the promise of remotely delivered health services when traditional models may not have been available or accessible [7].

Digital mental health interventions (DMHIs) are potentially scalable and effective treatment solutions for mental health. Although research evaluating the effectiveness of DMHIs notes significant heterogeneity in terms of study design, clinical model, and intervention characteristics, meta-analytic reviews of smartphone-based interventions have observed a clinically significant treatment effect compared to both wait-list and active controls for depression and anxiety symptoms [8-10].

There is considerable research consensus that cognitive behavioral therapy (CBT) is a highly effective therapeutic modality for the treatment of clinical depression and anxiety because of relatively short intervention durations with strong outcomes [11,12]. Specifically, CBT-based digital interventions were associated with a significantly greater reduction in depression compared to non-CBT therapeutic approaches (eg, mindfulness, psychoeducation, and mood monitoring) [9]. In the context of DMHIs, digital CBT interventions appear to be equivalent to face-to-face interventions in terms of treatment efficacy [13-16]. Furthermore, a recent assessment of digital CBT offerings that evaluated patient preference, cost savings, and clinical benefit concluded that a provider-guided digital CBT program was likely the most effective approach for depression and anxiety compared to self-guided and face-to-face interventions [17].

In addition to evaluating the efficacy of CBT-based digital interventions, there has been growing research interest in exploring their mechanisms, particularly in the role of program engagement in DMHIs. Recently, Chien et al [18] employed a machine learning–based approach to explore patterns of engagement with a digital CBT program and their relationship with improvements in depression and anxiety. While the study did identify five distinct classes of program engagement, from low to high engagement, all classes were all positively associated with clinical improvement. Higher engagement classes were associated with greater improvement in depressive symptoms, so a possible dose-response effect was suggested.

However, it has been observed that user engagement can be difficult to define and measure [19]. Where usage generally measures observable actions, engagement implies some subjective experience of the digital intervention with a focus on the quality of the experience [20,21]. Furthermore, as Torous et al [22] note, usage over time, a commonly used measure of engagement, cannot distinguish a user whose mental health needs have been met by an app with a single instance of usage from one who needs repeated access for support.

One app feature that has been previously operationalized as a measure of DMHI engagement is lesson completion [23]. Lessons are typically brief pieces of evidence-based content intended to support a skill or health habit. They can also serve a crucial function in mental health interventions, digital or otherwise. CBT, in particular, emphasizes the practice of skills in between sessions by way of “homework assignments” or lessons in order to reinforce practices such as cognitive reappraisal [24]. Kazantzis et al [25] observed a stronger treatment effect size for CBT-based therapeutic interventions that incorporated homework when compared to interventions without a homework component. Lesson completion has also been used as a measure of treatment adherence and has shown a positive association with improvements in depression and anxiety [17,25-29]. It has been noted that the ubiquity of smartphones and the flexibility of the technology enables access to extensive content, on demand or needs based, that is difficult to replicate in traditional delivery models [28,30].

Lastly, an often-overlooked dimension of clinical depression, particularly in the context of DMHIs, is its frequent co-occurrence with chronic health conditions, such as type 2 diabetes and obesity [31-33]. Among other aspects, depression is known to hamper self-care and medication adherence [33]. Cross-sectional research suggests a 1.18 times greater likelihood of depressive symptoms in individuals with obesity than in those without [34]. Similarly, the risk of prediabetes and related measures appears to be markedly elevated among those newly diagnosed with depression [35,36]. This relationship tends to be stronger among women [34,35]. In their meta-analytic review of collaborative care to treat depression and diabetes in tandem, Atlantis et al [31] observed significant improvements in both depression and glycemic management with tandem treatment. Another systematic review showed that psychological interventions tailored for people with diabetes were effective in improving both glycemic management (ie, hemoglobin A_{1c} [HbA_{1c}]) and elevated diabetes distress [37]. There are indications that the relationship between chronic conditions and depression are bidirectional, with an increased incidence of diabetes among those with diagnosed depression [38-40]. Many app-based health interventions have a singular health focus (eg, depression, weight, or diabetes management), making it more challenging to fully understand the impact and the influence of co-occurring conditions on treatment adherence and health outcomes [41].

In sum, research suggests that CBT-based DMHIs are feasible and effective solutions for depression and anxiety disorders [42-44]. There are some preliminary indications that the extent of program usage may influence treatment adherence and

outcomes [45]. There remain open research questions on how intervention type, mode of delivery, and mechanisms may affect treatment outcomes. There has also been less focus on DMHIs in populations living with co-occurring depression and chronic health conditions.

In this study, we evaluated the Vida CBT Program for moderate depression and anxiety in a polychronic adult population, that is, one with a high prevalence of chronic health conditions. Vida Health is a Health Insurance Portability and Accountability Act–compliant, app-based platform for management of both mental and physical health that combines tailored content with counseling by licensed therapists and other health education specialists. The platform is available directly to consumers or as a benefit from select employers and health plans. The primary objective of this study was to assess changes in depression following a 12-week digital CBT program: the Vida CBT Program. We hypothesized that participants who completed the program would show a reduction in depressive symptoms and that measures of program usage would be positively associated with these improvements. As part of a preliminary exploratory analysis, we also evaluated changes in weight among participants concurrently enrolled in a Vida physical health program. We suspected that improvements in depression would be positively associated with stronger weight loss.

Methods

Study Design

This study used a single-arm, retrospective design to evaluate changes in depression and anxiety following the Vida CBT Program.

Ethics Approval

The study protocol and informed consent statement was reviewed and approved by the Western Institutional Review Board Inc (protocol No. 20192591), an independent institutional review board. All data were fully anonymized prior to data analysis. Informed consent statements were sent to participants upon enrollment in the Vida CBT Program.

Measures

The 8-item Patient Health Questionnaire (PHQ-8) was used to assess severity of depressive symptoms. The following standard scoring cutoffs were applied to classify depressive symptom severity: 0 to 4 (asymptomatic or minimal), 5 to 9 (mild), 10 to 14 (moderate), 15 to 19 (moderately severe), and 20 or higher (severe). Anxiety symptoms were assessed using the 7-item Generalized Anxiety Disorder scale (GAD-7) using the following standard score cutoffs to classify anxiety symptom severity: 0 to 4 (asymptomatic or minimal), 5 to 10 (mild), 11 to 17 (moderate), and 18 or higher (severe). The PHQ-8 and GAD-7 are widely used in clinical settings and have shown robust reliability and validity [39,40]. The assessments were administered in the app automatically at program start, week 6, week 12, and every 3 months thereafter for up to 1 year. Participants were encouraged to complete the survey on the day of receipt but had the option to complete the assessment at any point during the 2 weeks following receipt, after which the survey would disappear until the next assessment time point. Therapists also had the option of sending the instrument to the participant at any point during the intervention as they deemed clinically appropriate.

Body weight was a secondary outcome measure in this study. Weight outcomes were either recorded via a connected wireless scale or self-reported by the participant using a logging tool available in the Vida Health app. Participants could either sync their personal wireless scale to the app or order a digitally connected scale via the app.

Study Sample and Recruitment

The study was open to adults, 18 years of age or older, who were fluent in English and had access to a smartphone or tablet. Participants were recruited between September 2019 and January 2021 using a combination of emails, mailers, and phone outreach efforts. The Vida Health app is available to individuals across the United States and can be downloaded from the Apple App Store or Google Play. Participants were drawn from several insurance plans and employers for whom Vida Health was offered as a covered benefit.

Upon downloading the app, participants completed a brief intake questionnaire that included name, contact information, basic demographics (ie, age, gender, height, and weight), and existing health conditions. Participants were offered a variety of health domains to focus on. Some participants chose to begin the Vida CBT Program directly. Others began in the Vida digital programs for weight loss or diabetes management described previously [46,47]. Participants received and completed an intake PHQ-8, and those with a score of 10 or greater (moderate depression) were included in the study. All participants also completed a baseline GAD-7 assessment. Eligible participants were paired with a licensed therapist based on their state of residence and preferred times for consultations. Therapists were mental health professionals working for Vida Health, licensed by their state's respective licensure board. Participants with severe depression (PHQ-8 score ≥20) or severe anxiety (GAD-7 score ≥15) were ineligible for the Vida CBT Program and were referred to alternative sources of care. Additional exclusion criteria included eating disorders, substance use disorder, suicidality, homicidality, acute posttraumatic stress disorder, and episodes of mania or psychosis. Participants who presented with any of the above symptoms during the intervention were referred for care outside of Vida Health. Participants who chose to enroll initially in weight loss or diabetes management programs and were then screened into the Vida CBT Program had the option to continue both programs simultaneously.

Therapeutic Approach and Intervention

A fundamental focus of CBT is to address maladaptive thinking patterns by understanding the associations among thoughts, emotions, and behaviors [48]. As part of the Vida CBT Program, participants received structured multimedia (ie, audio, video, or text) lessons, activities, and practices within the app. Based on core CBT principles, such as guided discovery and conscious re-evaluation, these lessons were designed to increase awareness of one's thinking patterns and support the practice of alternative, adaptive thoughts [24]. Ahead of their initial consultation with a Vida therapist, participants completed an informed consent

form for psychotherapy that detailed their rights to confidentiality and limits to confidentiality, including mandated reporting requirements as stipulated by the therapist's licensing board and state regulations. At the initial consultation, therapists performed a comprehensive biopsychosocial assessment that included a review of previous treatments and diagnoses, current presenting problem, and symptoms. Participants could communicate with their therapist using live video or audio consultations as well as with asynchronous messaging in the app.

Following the intake, therapists documented their initial diagnostic impressions and developed individualized treatment plans with short-term and long-term goals. Participants were offered weekly video or audio consultations with their therapist for the first 12 weeks and shorter monthly follow-up sessions thereafter for up to 1 year. Each therapist consultation comprised setting goals and homework activities for the upcoming week along with a review of strategies for further cultivating concepts and skills learned from earlier sessions. In between sessions, participants could complete assigned homework and use a thought tracker in the app. Lessons, once shared, remained available in the app for the participant to review and revisit concepts. The thought tracker feature allowed participants to record current thoughts using a Likert scale to assist in identifying, evaluating, and restructuring distorted thought patterns. In program weeks 10 through 12, therapists worked with participants to create a Wellness Recovery Action Plan intended to support maintenance of acquired skills and improved functioning and to prevent relapse [49]. Select screenshots from the program are shown in Figure 1.

Figure 1. Screenshots from the Vida CBT Program. CBT: cognitive behavioral therapy.

Participants who were concurrently enrolled in a program for chronic disease management also worked with an additional Vida provider (ie, a certified health coach or registered dietitian) who could collaborate on care with the relevant therapist. As in the Vida CBT Program, they were offered synchronous consultations weekly for up to 12 weeks and monthly thereafter for up to 1 year. All providers received extensive training on motivational interviewing, an approach that leverages improving perceived self-efficacy and autonomy to facilitate healthy behavior change [50]. Additionally, participants received app content covering topics such as nutrition, exercise, and medication adherence. All content was informed by evidence-based research and literature on health behavior change as described previously [46,47,51].

Statistical Plan

Change in depression, based on PHQ-8 scores, between baseline and follow-up was the primary dependent variable. We performed a paired, 2-tailed *t* test to assess if there was a significant change in PHQ-8 scores from baseline. We used a second paired *t* test to evaluate changes in anxiety scores among participants who scored in the moderate anxiety range at baseline (ie, GAD-7 score ≥11). Mean normalization was applied to all continuous predictors (eg, age). A Boolean variable was created for the presence of co-occurring anxiety (1 = baseline GAD-7 score ≥11). Gender was also coded as a binary variable (1 = female). Since we could not assume that all therapists were equally effective, all regression analyses were conducted using a cluster-robust approach with therapist as a cluster group variable [52]. The reliable change index (RCI) for depression and anxiety was also computed. RCI, a commonly used measure in psychometrics, is the ratio of the difference in pre-post assessment scores to the standard error of measurement [47,48]. An RCI score of 1.96 or higher (ie, 1 above the 95% CI) is regarded as an indication of reliable, statistically meaningful change [53].

To explore the interaction between program usage and changes in depression, we tabulated five program features that broadly encompass the program experience: number of therapist consultations, number of messages sent to the therapist, number of "core" lessons opened, number of thoughts logged, and total number of content pieces opened. Core lessons were those related to fundamental CBT concepts (eg, cognitive restructuring, behavioral activation, and techniques for addressing maladaptive thinking) [24]. Data exploration revealed a right skew for each of the usage factors, that is, a subpopulation of participants who used the app features quite extensively. In order to retain these heavy app users, but to limit their influence on downstream analyses, all usage factors were right-winsorized at the 99th percentile [54]. The treatment cohort showed a notable female predominance, a finding not unusual

in studies of mental health service use in the United States [55-57]. In this study, in order to adjust for this, the established technique of oversampling was employed, wherein 715 participants were drawn at random from the 116 participants who did not identify as female [58,59].

Finding, as expected, that features of program usage were correlated, we constructed a composite variable to limit collinearity. Participants were assigned an ordinal variable from 0 to 5, where 1 unit was assigned for each usage feature in which the participant was above the 25th percentile. Thus, someone with a depth-of-usage score of 5 had activity above the 25th percentile in each of the usage factors, while a score of 0 indicated usage within the 25th percentile across all usage features. We chose the 25th percentile as it empirically separated users into high and low usage categories without selecting only for the right tail of usage as alluded above. With reliable change as the binary dependent variable (1 = reliable improvement in depression symptoms), a cluster-robust logistic regression evaluated the relationship depth of usage and improvement in depression symptoms. Controls included therapist cluster, baseline PHQ-8 score, gender, presence of anxiety, and age.

A supplementary analysis evaluating changes in weight outcomes among participants concurrently enrolled in the Vida weight loss program was conducted. Analysis was restricted to participants who had been enrolled for at least 6 months and had logged weight at least twice in that period. A linear, mixed-effects model was used to address potential heterogeneity in the frequency and number of weight logs by participants and provider-level differences [60]. Percent change in weight from baseline was regressed on the following fixed factors: program time (in months), baseline BMI category (obesity or overweight), and reliable change (1 = reliable improvement in PHQ-8 scores; 0 = no reliable improvement) [61]. Participants and providers were specified in the model as random factors.

All data preparation and analyses were performed using Python (version 3.7.9; Python Software Foundation) and Stata (version 16.1; StataCorp LLC).

Results

Overview

A total of 1512 participants enrolled in the Vida CBT Program between September 2019 and January 2021. A schematic of the participant flow is presented in Figure 2. Of those, 54.96% (n=831) had a follow-up PHQ-8 assessment between weeks 6 and 12 and were considered part of the treatment cohort. Analyses evaluating changes in depression scores from baseline were restricted to the treatment cohort. The remainder of the study cohort, equal to 45.04% (n=681) of the participants, failed to complete a follow-up PHQ-8 during the assessment window. These participants were excluded from the primary analyses; however, we performed a supplementary intention-to-treat (ITT) analysis to evaluate overall changes in PHQ-8 scores from baseline across the entire cohort. Among those without a valid follow-up, 71.5% (487/681) did not have any assessment after their baseline. For these members, a baseline carryforward approach was employed. The remaining 28.5% (194/681) of the participants completed an assessment before finishing 6 program weeks. For these members, a last-value carryforward approach was applied.

Baseline characteristics of the study cohort are reported in Table 1. In addition to depression, the entire study cohort self-reported living with at least one chronic physical health condition (ie, type 2 diabetes, cardiovascular disease, or obesity). There were no significant differences in baseline PHQ-8 scores between the treatment and incomplete groups at baseline (t_{1510}=1.2, P=.22). There was a significantly lower rate of comorbid anxiety, defined as a GAD-7 score of 11 or higher, among participants in the treatment cohort (χ^2_1=19.5, P<.001). A 2-tailed chi-square analysis indicated that there were significantly more women in the treatment cohort compared to the program noncompleter group (χ^2_1=7.1, P=.01). There were also more participants concurrently enrolled in a health coaching program for a chronic condition in the treatment cohort (χ^2_1=171.1, P<.001). Lastly, we observed a significant average difference in age of 2.3 years between the treatment cohort (mean 48.5, SD 11.4 years) and noncompleters (mean 46.2, SD 12.4 years; t_{1510}=–3.8, P<.001).

Figure 2. A schematic of the participant flow. CBT: cognitive behavioral therapy; PHQ-8: 8-item Patient Health Questionnaire.

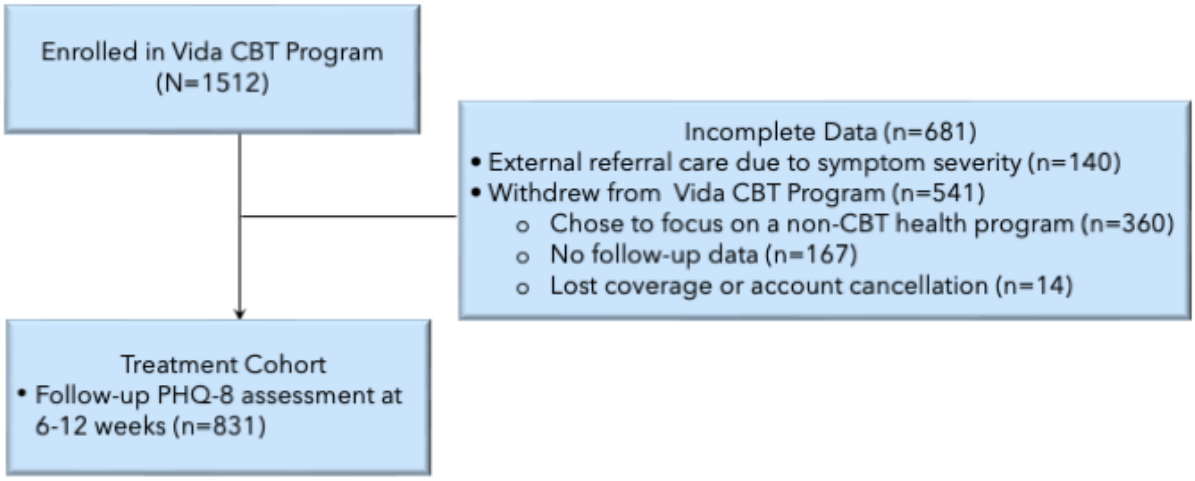


Table 1. Demographic characteristics of the treatment and intention-to-treat study cohorts.

Characteristic	Incomplete data (n=681)	Treatment (n=831)	Total (N=1512)
Participants (N=1512), n (%)	681 (45.0)	831 (55.0)	1512 (100)
Gender, n (%)			
Female	552 (81.1)	715 (86.0)	1267 (83.8)
Male	127 (18.7)	113 (13.6)	240 (15.9)
Not disclosed	2 (0.3)	3 (0.4)	5 (0.3)
Age in years, mean (SD)	46.2 (12.4)	48.5 (11.4)	47.5 (11.9)
Baseline PHQ-8[a] score, mean (SD)	14.6 (3.3)	14.4 (3.4)	14.5 (3.3)
Has anxiety (GAD-7[b] score ≥11), n (%)	467 (68.6)	478 (57.5)	945 (62.5)
Enrolled in physical health program, n (%)	360 (52.9)	697 (83.9)	1057 (69.9)
Chronic physical health condition, n (%)			
Obesity	554 (81.4)	723 (87.0)	1277 (84.5)
Cardiovascular disease	367 (53.9)	503 (60.5)	870 (57.5)
Diabetes	316 (46.4)	468 (56.3)	784 (51.9)

[a]PHQ-8: 8-item Patient Health Questionnaire; depressive symptom severity by score is classified as follows: 0 to 4 (asymptomatic or minimal), 5 to 9 (mild), 10 to 14 (moderate), 15 to 19 (moderately severe), and 20 or higher (severe).

[b]GAD-7: 7-item Generalized Anxiety Disorder scale; a score of ≥11 indicates moderate to severe anxiety.

Principal Results

Out of 831 participants in the treatment cohort, 74.5% (n=619) showed a clinically significant reduction in depressive symptom severity in 12 weeks, where follow-up PHQ-8 scores had shifted downward by at least one diagnostic category. A total of 67.5% (n=561) of the treatment cohort participants showed a reliable improvement in PHQ-8 scores as measured by the RCI. There was an average reduction of 5.9 (SD 5.2) points between baseline and follow-up (Table 2). A 2-tailed, paired *t* test revealed a significant reduction in scores between baseline and follow-up (t_{830}=32.9, *P*<.001). A cluster-robust linear regression examining the association between change in PHQ-8 scores and baseline scores, co-occurrence of anxiety, gender, age, and concurrent enrollment in physical health revealed a significant inverse relationship between baseline and change in PHQ-8 scores (β=–1.7, *P*<.001). That is, greater baseline depression severity was, unsurprisingly, associated with greater reduction of depression at follow-up. We observed that the co-occurrence of anxiety was associated with a smaller reduction in depression scores (β=1.33, *P*<.001). In supplementary analyses, a 2-tailed, paired *t* test including the entire ITT cohort, using carryforward as above, indicated a slightly attenuated but still significant reduction in PHQ-8 scores at follow-up (mean –3.6, SD 4.99; t_{1511}=29.1, *P*<.001).

Among the 478 participants with moderate anxiety or higher at baseline (ie, GAD-7 score ≥11), 89.7% (n=429) provided a follow-up assessment between weeks 6 and 12. A paired *t* test revealed a significant average reduction in anxiety scores of 6.1 (SD 5.4) points from baseline (t_{428}=22.96, *P*<.001; Table 2). Additionally, 57.7% (n=276) of these participants had a reliable improvement in anxiety scores from baseline. A cluster-robust linear regression showed that higher baseline GAD-7 scores

were associated with greater reductions in anxiety scores at follow-up (β=–3.5, *P*<.001). Age, gender, and baseline PHQ-8 scores were not significantly associated with anxiety score changes.

Table 2. Estimated marginal means of PHQ-8 and GAD-7 scores at baseline and follow-up.

Assessment type and time point	Score, estimated marginal mean (bootstrapped 95% CI)
PHQ-8[a]	
Baseline	14.4 (14.2-14.6)
Follow-up (12 weeks)	8.5 (8.1-8.8)
GAD-7[b]	
Baseline	14.8 (14.5-15.0)
Follow-up (12 weeks)	8.8 (8.3-9.3)

[a]PHQ-8: 8-item Patient Health Questionnaire; depressive symptom severity by score is classified as follows: 0 to 4 (asymptomatic or minimal), 5 to 9 (mild), 10 to 14 (moderate), 15 to 19 (moderately severe), and 20 or higher (severe).

[b]GAD-7: 7-item Generalized Anxiety Disorder scale; anxiety symptom severity by score is classified as follows: 0 to 4 (asymptomatic or minimal), 5 to 10 (mild), 11 to 17 (moderate), and 18 or higher (severe).

Program Usage Outcomes

Program usage was evaluated in the treatment cohort across a set of five program features: number of therapist consultations, number of messages sent to the therapist, number of times core lesson content was accessed, number of thoughts logged via the thought tracking tool, and number of program-related content cards viewed. As expected, there were notable correlations among these program usage factors, with Pearson coefficients ranging from 0.17 to 0.43 (Table 3).

In order to limit collinearity while evaluating the association between program usage and changes in depressive symptoms, we created a composite feature of overall depth of usage. For each usage feature, a score of 1 was assigned when activity for that specific feature was above the 25th percentile of the distribution. We then summed across each of the factors. Thus, depth-of-usage scores could range from 0 to 5, with a score of 5 indicating usage in excess of the 25th percentile of activity across the features, and a score of 0 indicating usage below the 25th percentile across the features. As noted above, the data were then resampled to adjust for the female predominance in the treatment cohort. From the resulting balanced cohort of 1430 participants, a cluster-robust logistic model controlling for age, gender, co-occurrence of anxiety (ie, baseline GAD-7 score ≥11), and baseline PHQ-8 score revealed a significant association between depth of usage and likelihood of reliable improvement in depression scores at follow-up (odds ratio [OR] 1.3, 95% CI 1.1-1.5; *P*=.002). In other words, greater usage across the platform was associated with improvement in depression symptom severity. The summary statistics for the features of program usage were as follows: mean number of therapist consultations was 7.04 (SD 3.4), mean number of messages sent to the therapist was 33.2 (SD 46.9), mean number of times core lesson content was accessed was 39.2 (SD 33), mean number of thoughts logged via the thought tracking tool was 13.6 (SD 16.5), and mean number of program-related content cards viewed was 8.35 (SD 12.4).

Table 3. Correlation analysis (Pearson r and 2-tailed *P* value) among the features of program usage (n=831).

Variable	Consults	Messages	Lessons opened	Thoughts logged	Content viewed
Consults					
r	1	0.29[a]	0.26[a]	0.26[a]	0.17[a]
P value	—[b]	<.001	<.001	<.001	<.001
Messages					
r	0.29[a]	1	0.33[a]	0.43[a]	0.31[a]
P value	<.001	—	<.001	<.001	<.001
Lessons opened					
r	0.26[a]	0.33[a]	1	0.44[a]	0.35[a]
P value	<.001	<.001	—	<.001	<.001
Thoughts logged					
r	0.26[a]	0.43[a]	0.44[a]	1	0.28[a]
P value	<.001	<.001	<.001	—	<.001
Content viewed					
r	0.17[a]	0.31[a]	0.35[a]	0.28[a]	1
P value	<.001	<.001	<.001	<.001	—

[a]The correlation is significant at a significance level of .05 (2-tailed).

[b]Not applicable.

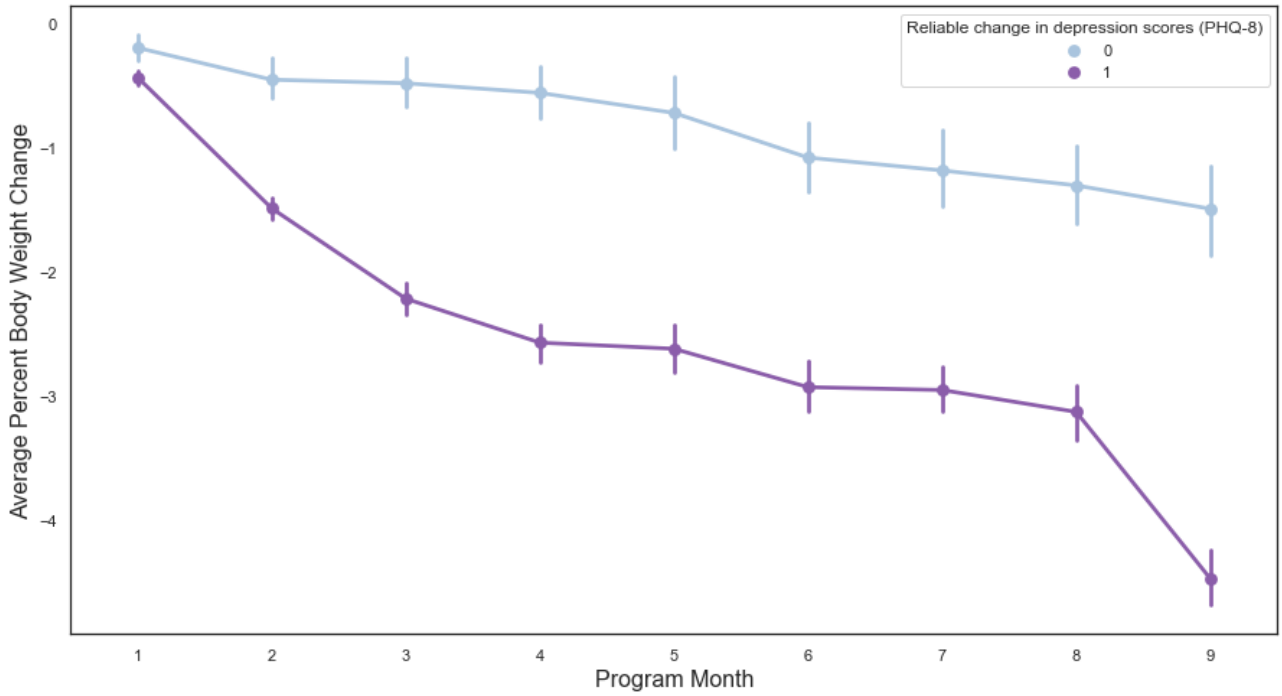
Exploratory Weight

We performed additional analyses to explore changes in body weight and diabetes management among participants simultaneously enrolled in the Vida CBT Program and a physical health program.

For body weight, in order to provide adequate time for meaningful clinical change, the analysis was restricted to the 595 participants with overweight or obesity (ie, BMI ≥25) who had been enrolled for at least 6 months and had logged weight at least twice in the period with the latest weight coming at least 2 months after the initial enrollment. Out of these participants, 86.6% (n=515) had a baseline BMI indicative of obesity and 13.4% (n=80) had a baseline BMI indicative of overweight. This population logged a total of 34,469 body weight values. Out of these entries, 83.17% (n=28,667) were logged using a wireless connected scale, and the remainder were entered manually by the participant. The majority of the participants (n=387, 65.0%) had logged their most recent weight in the ninth program month. The most recent weight was logged at least in the eighth program month for 78.8% (n=469) of the participants and at least in the sixth program month for 89.2% (n=531) of them.

To account for heterogeneity in the number and frequency of weight logs among participants, as well as to handle the nested structure of the data (ie, weight observations nested within participants who were nested within providers), a multilevel mixed-effects model was used to evaluate changes in body weight across participants. Fixed factors in the model were time (program month), baseline BMI category (obesity or overweight), and reliable change in depressive symptoms (1 = reliable improvement in PHQ-8 scores; 0 = no reliable improvement). Random factors included participants and providers. We observed a significant effect of program time on percent weight loss (β=–0.37, *P*<.001). Baseline weight had a small but significant inverse association with weight loss such that higher baseline weight was associated with greater percent weight loss (β=–0.01, *P*=.002). As shown in Figure 3, reliable improvement in depression symptoms at follow-up was associated with significantly greater weight loss at 9 months (β=–1.11, *P*=.002).

Figure 3. Estimated marginal means and SEs of average percent weight loss as a function of program month and reliable change in depression symptoms. A value of 1 represents reliable improvement in PHQ-8 scores and 0 represents no reliable improvement. PHQ-8: 8-item Patient Health Questionnaire.



Discussion

Principal Findings

The aim of this retrospective study was to evaluate the therapist-supported, digitally delivered Vida CBT Program for the treatment of moderate depression in an adult population. The treatment cohort of 831 participants with baseline PHQ-8 scores of 10 or greater were enrolled in the digital intervention and provided at least one follow-up assessment 6 to 12 weeks into the program. We observed a significant reduction in depression scores at follow-up (mean –5.9, SD 5.2), with 74.5% of participants shifting downward in symptom severity by at least one diagnostic category. A similar pattern of reduction was observed for anxiety scores among participants who had moderate anxiety (mean –6.1, SD 5.4). Supplementary ITT analysis that included the entire cohort of 1512 participants with baseline data or last-value carryforward showed a slightly attenuated but still significant reduction in PHQ-8 scores at follow-up.

Examining the relationship between program usage and improvement in depression symptoms revealed that participants who more extensively used features across the platform had a modestly greater likelihood of reliable improvement in depression scores at follow-up (OR 1.3, 95% CI 1.1-1.5). This limited dose-response relationship was seen in the context of highly right-skewed usage data overall.

Drawn largely from a medically complex adult population, cardiometabolic conditions, particularly diabetes and obesity, were highly prevalent in this population. Given the very high burden of metabolic disease in this population and the availability of interventions for these conditions on the Vida platform, although the study was not specifically designed for this, we wanted to explore these relationships in order to prepare for future research.

Our preliminary analyses showed significantly greater reduction in body weight among participants who had a reliable improvement in depression symptoms compared to those who did not show symptom improvement at follow-up. As Figure 3 illustrates, this trend was consistent throughout the program tenure and culminated in a mean body weight change of 4.5% (SD 6.3; reliable improvement in PHQ-8 scores) versus 1.5% (SD 6.1; no reliable improvement in PHQ-8 scores) at month 9 of the program.

As with much of the literature, it is impossible to infer causation from these results, but it opens a hypothesis for future research that there may be opportunities for synergy in the treatment of co-occurring mental and metabolic disorders [37,62]. Given the massive and overlapping burden of these diseases in the US population, if such synergy could be realized through efficient, digitally delivered interventions, it would certainly be welcome.

Comparison With Prior Work

The principal finding of improvements in depressive symptoms is consistent with the broader body of research on digitally delivered CBT programs for the treatment of mild to moderate depression and anxiety [9]. This finding reinforces, in a larger population, similar results previously published from the Vida CBT Program [23].

The relationship between program usage and outcomes in digital interventions, broadly, and DMHIs, in particular, remains a rich area of research. While this study was not designed to evaluate this, our results are consistent with several similar studies suggesting a limited dose-response phenomenon. That is, greater usage seems to correlate with improved outcomes but may not do so monotonically, as participants self-regulate their usage

to their needs, needs that are difficult for a researcher to observe with any meaningful precision [18,63-66].

Throughout the literature, interrelations among treatments for comorbid mental and physical disorders have been noted before. The direction of causality is unclear and, indeed, may be expected to run in both directions; that is, depression and anxiety may contribute to obesity and diabetes just as much as the reverse is true. In some cases, other factors may be driving both [62,67-69]. Preliminary research in this vein has, like this study, shown some hint of synergistic treatment effects in traditional care settings [70-72].

Limitations

This study had several important limitations. The lack of a control group and the retrospective design prevents drawing any causal inferences. Missing follow-up data was also a challenge, with data unavailable for 45.0% of the overall cohort (Table 1). We attempted to mitigate this limitation with an ITT analysis using baseline data or last-value carryforward as appropriate. While this still showed a significant relationship, the improvement in depression scores was weaker when accounting for those participants.

The gender imbalance across the study cohort was notable, with 83.8% females overall and 86.0% in the treatment group. At the overall level, this is a well-characterized phenomenon of care use, particularly for mental health, in North America and Europe. Despite seemingly comparable prevalence of mental distress by gender, ratios of 2 women for every man seeking care are not unusual in the literature [73,74]. The US National Institute of Mental Health has even sponsored campaigns to address barriers to males seeking mental health care, and there is literature on this phenomenon's underlying causes [56,75]. In this study, there was also a small but significant increase in the proportion of females in the treatment group. It may have been that whatever phenomenon inhibits males from seeking care, in general, also made males in this study less likely to follow up with treatment and assessment. This and the difference in age between the two groups could represent some self-selection of participants into the treatment group who were more likely to improve. While the ITT analysis attempts to adjust for this, some bias remains a possibility. More broadly, the limited uptake of the Vida intervention by males may limit its generalizability in a larger population.

This study was not specifically designed to investigate changes in body weight and glycemic control as related to improvements in depression and was underpowered to detect a change in HbA_{1c}. Future research is certainly needed to explore these potential relationships. Causality in any such relationship, should it bear out, could take any possible form and would need careful study to dissect.

Conclusions

This study provides further evidence that CBT-based DMHIs can be effective tools in treating symptoms of depression and anxiety. Improvements by whole diagnostic categories were common in this cohort. Usage of the platform correlated with improved outcomes in a pattern consistent with participants self-regulating usage to their individual needs. Furthermore, this was seen in a population with a very high prevalence of co-occurring physical disorders. More research is required to determine if there may be opportunities for synergistic treatment of mental and physical disorders through a similar modality.

Acknowledgments

We are deeply indebted to the hard work of all the therapists, dietitians, and health coaches who supported these participants. We are grateful to Chris Mosunic for helpful comments on the manuscript.

Authors' Contributions

AV and MS conceived and designed the study. MM led the intervention design. AV, PR, and BF acquired, analyzed, and interpreted the data. AV and MS drafted, structured, and edited the manuscript. All authors discussed the results and contributed to the final manuscript.

Conflicts of Interest

All authors are current or former employees of Vida Health and hold share options.

References

1. Rehm J, Shield KD. Global burden of disease and the impact of mental and addictive disorders. Curr Psychiatry Rep 2019 Feb 07;21(2):10. [doi: 10.1007/s11920-019-0997-0] [Medline: 30729322]
2. Bromet E, Andrade LH, Hwang I, Sampson NA, Alonso J, de Girolamo G, et al. Cross-national epidemiology of DSM-IV major depressive episode. BMC Med 2011 Jul 26;9:90 [FREE Full text] [doi: 10.1186/1741-7015-9-90] [Medline: 21791035]
3. Bachmann S. Epidemiology of suicide and the psychiatric perspective. Int J Environ Res Public Health 2018 Jul 06;15(7):1425 [FREE Full text] [doi: 10.3390/ijerph15071425] [Medline: 29986446]
4. Simon GE, VonKorff M, Barlow W. Health care costs of primary care patients with recognized depression. Arch Gen Psychiatry 1995 Oct;52(10):850-856. [doi: 10.1001/archpsyc.1995.03950220060012] [Medline: 7575105]
5. Kessler RC. The costs of depression. Psychiatr Clin North Am 2012 Mar;35(1):1-14 [FREE Full text] [doi: 10.1016/j.psc.2011.11.005] [Medline: 22370487]

6. Ettman CK, Abdalla SM, Cohen GH, Sampson L, Vivier PM, Galea S. Prevalence of depression symptoms in US adults before and during the COVID-19 pandemic. JAMA Netw Open 2020 Sep 01;3(9):e2019686 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.19686] [Medline: 32876685]
7. Doraiswamy S, Abraham A, Mamtani R, Cheema S. Use of telehealth during the COVID-19 pandemic: Scoping review. J Med Internet Res 2020 Dec 01;22(12):e24087 [FREE Full text] [doi: 10.2196/24087] [Medline: 33147166]
8. 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]
9. 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]
10. Philippe TJ, Sikder N, Jackson A, Koblanski ME, Liow E, Pilarinos A, et al. Digital health interventions for delivery of mental health care: Systematic and comprehensive meta-review. JMIR Ment Health 2022 May 12;9(5):e35159 [FREE Full text] [doi: 10.2196/35159] [Medline: 35551058]
11. Garrido S, Millington C, Cheers D, Boydell K, Schubert E, Meade T, et al. What works and what doesn't work? A systematic review of digital mental health interventions for depression and anxiety in young people. Front Psychiatry 2019;10:759 [FREE Full text] [doi: 10.3389/fpsyt.2019.00759] [Medline: 31798468]
12. Jonsson U, Bertilsson G, Allard P, Gyllensvärd H, Söderlund A, Tham A, et al. Psychological treatment of depression in people aged 65 years and over: A systematic review of efficacy, safety, and cost-effectiveness. PLoS One 2016;11(8):e0160859 [FREE Full text] [doi: 10.1371/journal.pone.0160859] [Medline: 27537217]
13. Luo C, Sanger N, Singhal N, Pattrick K, Shams I, Shahid H, et al. A comparison of electronically-delivered and face to face cognitive behavioural therapies in depressive disorders: A systematic review and meta-analysis. EClinicalMedicine 2020 Jul;24:100442 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100442] [Medline: 32775969]
14. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs face-to-face cognitive behavior therapy for psychiatric and somatic disorders: An updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
15. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: A meta-analysis. PLoS One 2010 Oct 13;5(10):e13196 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
16. Weisel KK, Fuhrmann LM, Berking M, Baumeister H, Cuijpers P, Ebert DD. Standalone smartphone apps for mental health-A systematic review and meta-analysis. NPJ Digit Med 2019;2:118 [FREE Full text] [doi: 10.1038/s41746-019-0188-8] [Medline: 31815193]
17. Health Quality Ontario. Internet-delivered cognitive behavioural therapy for major depression and anxiety disorders: A health technology assessment. Ont Health Technol Assess Ser 2019;19(6):1-199 [FREE Full text] [Medline: 30873251]
18. Chien I, Enrique A, Palacios J, Regan T, Keegan D, Carter D, et al. A machine learning approach to understanding patterns of engagement with internet-delivered mental health interventions. JAMA Netw Open 2020 Jul 01;3(7):e2010791 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.10791] [Medline: 32678450]
19. Ng MM, Firth J, Minen M, Torous J. User engagement in mental health apps: A review of measurement, reporting, and validity. Psychiatr Serv 2019 Jul 01;70(7):538-544 [FREE Full text] [doi: 10.1176/appi.ps.201800519] [Medline: 30914003]
20. O'Brien HL, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. J Am Soc Inf Sci Technol 2008 Apr;59(6):938-955. [doi: 10.1002/asi.20801]
21. 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]
22. Torous J, Michalak EE, O'Brien HL. Digital health and engagement-looking behind the measures and methods. JAMA Netw Open 2020 Jul 01;3(7):e2010918 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.10918] [Medline: 32678446]
23. Venkatesan A, Rahimi L, Kaur M, Mosunic C. Digital cognitive behavior therapy intervention for depression and anxiety: Retrospective study. JMIR Ment Health 2020 Aug 26;7(8):e21304 [FREE Full text] [doi: 10.2196/21304] [Medline: 32845246]
24. Beck JS. Cognitive Behavior Therapy: Basics and Beyond. 3rd edition. New York, NY: The Guilford Press; 2021.
25. Kazantzis N, Whittington C, Zelencich L, Kyrios M, Norton PJ, Hofmann SG. Quantity and quality of homework compliance: A meta-analysis of relations with outcome in cognitive behavior therapy. Behav Ther 2016 Sep;47(5):755-772. [doi: 10.1016/j.beth.2016.05.002] [Medline: 27816086]
26. Tang W, Kreindler D. Supporting homework compliance in cognitive behavioural therapy: Essential features of mobile apps. JMIR Ment Health 2017 Jun 08;4(2):e20 [FREE Full text] [doi: 10.2196/mental.5283] [Medline: 28596145]
27. Aguilera A, Ramos Z, Sistiva D, Wang Y, Alegria M. Homework completion via telephone and in-person cognitive behavioral therapy among Latinos. Cognit Ther Res 2018 Jun;42(3):340-347 [FREE Full text] [Medline: 29725144]

28. Callan JA, Dunbar Jacob J, Siegle GJ, Dey A, Thase ME, DeVito Dabbs A, et al. CBT MobileWork©: User-centered development and testing of a mobile mental health application for depression. Cognit Ther Res 2020 Sep 25;45(2):287-302. [doi: 10.1007/s10608-020-10159-4]
29. Kazantzis N, Brownfield NR, Mosely L, Usatoff AS, Flighty AJ. Homework in cognitive behavioral therapy: A systematic review of adherence assessment in anxiety and depression (2011-2016). Psychiatr Clin North Am 2017 Dec;40(4):625-639. [doi: 10.1016/j.psc.2017.08.001] [Medline: 29080590]
30. Callan JA, Wright J, Siegle GJ, Howland RH, Kepler BB. Use of computer and mobile technologies in the treatment of depression. Arch Psychiatr Nurs 2017 Jun;31(3):311-318. [doi: 10.1016/j.apnu.2016.10.002] [Medline: 28499574]
31. Atlantis E, Fahey P, Foster J. Collaborative care for comorbid depression and diabetes: A systematic review and meta-analysis. BMJ Open 2014 Apr 12;4(4):e004706 [FREE Full text] [doi: 10.1136/bmjopen-2013-004706] [Medline: 24727428]
32. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: Results from the World Health Surveys. Lancet 2007 Sep 08;370(9590):851-858. [doi: 10.1016/S0140-6736(07)61415-9] [Medline: 17826170]
33. Lin EHB, Katon W, Von Korff M, Rutter C, Simon GE, Oliver M, et al. Relationship of depression and diabetes self-care, medication adherence, and preventive care. Diabetes Care 2004 Sep;27(9):2154-2160. [doi: 10.2337/diacare.27.9.2154] [Medline: 15333477]
34. Carey M, Small H, Yoong SL, Boyes A, Bisquera A, Sanson-Fisher R. Prevalence of comorbid depression and obesity in general practice: A cross-sectional survey. Br J Gen Pract 2014 Mar;64(620):e122-e127 [FREE Full text] [doi: 10.3399/bjgp14X677482] [Medline: 24567650]
35. Markowitz S, Friedman MA, Arent SM. Understanding the relation between obesity and depression: Causal mechanisms and implications for treatment. Clin Psychol 2008 Mar;15(1):1-20. [doi: 10.1111/j.1468-2850.2008.00106.x]
36. Watson KT, Simard JF, Henderson VW, Nutkiewicz L, Lamers F, Nasca C, et al. Incident major depressive disorder predicted by three measures of insulin resistance: A Dutch cohort study. Am J Psychiatry 2021 Oct 01;178(10):914-920. [doi: 10.1176/appi.ajp.2021.20101479] [Medline: 34551583]
37. Schmidt CB, van Loon BJP, Vergouwen ACM, Snoek FJ, Honig A. Systematic review and meta-analysis of psychological interventions in people with diabetes and elevated diabetes-distress. Diabet Med 2018 Jun 13:1157-1172. [doi: 10.1111/dme.13709] [Medline: 29896760]
38. Mezuk B, Eaton WW, Albrecht S, Golden SH. Depression and type 2 diabetes over the lifespan: A meta-analysis. Diabetes Care 2008 Dec;31(12):2383-2390 [FREE Full text] [doi: 10.2337/dc08-0985] [Medline: 19033418]
39. Golden SH, Lazo M, Carnethon M, Bertoni AG, Schreiner PJ, Diez Roux AV, et al. Examining a bidirectional association between depressive symptoms and diabetes. JAMA 2008 Jun 18;299(23):2751-2759 [FREE Full text] [doi: 10.1001/jama.299.23.2751] [Medline: 18560002]
40. Aarts S, van den Akker M, van Boxtel MPJ, Jolles J, Winkens B, Metsemakers JFM. Diabetes mellitus type II as a risk factor for depression: A lower than expected risk in a general practice setting. Eur J Epidemiol 2009;24(10):641-648 [FREE Full text] [doi: 10.1007/s10654-009-9385-0] [Medline: 19718502]
41. Wang E, Zweig M. A defining moment for digital behavioral health: Four market trends. Rock Health. 2021 Mar 15. URL: https://rockhealth.com/insights/a-defining-moment-for-digital-behavioral-health-four-market-trends/ [accessed 2021-11-03]
42. Cuijpers P, Marks IM, van Straten A, Cavanagh K, Gega L, Andersson G. Computer-aided psychotherapy for anxiety disorders: A meta-analytic review. Cogn Behav Ther 2009;38(2):66-82. [doi: 10.1080/16506070802694776] [Medline: 20183688]
43. Ebert DD, Zarski A, Christensen H, Stikkelbroek Y, Cuijpers P, Berking M, et al. Internet and computer-based cognitive behavioral therapy for anxiety and depression in youth: A meta-analysis of randomized controlled outcome trials. PLoS One 2015;10(3):e0119895 [FREE Full text] [doi: 10.1371/journal.pone.0119895] [Medline: 25786025]
44. Richards D, Richardson T. Computer-based psychological treatments for depression: A systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
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. Silberman JM, Kaur M, Sletteland J, Venkatesan A. Outcomes in a digital weight management intervention with one-on-one health coaching. PLoS One 2020;15(4):e0232221 [FREE Full text] [doi: 10.1371/journal.pone.0232221] [Medline: 32353035]
47. Zimmermann G, Venkatesan A, Rawlings K, Scahill MD. Improved glycemic control with a digital health intervention in adults with type 2 diabetes: Retrospective study. JMIR Diabetes 2021 Jun 02;6(2):e28033 [FREE Full text] [doi: 10.2196/28033] [Medline: 34075880]
48. Fenn K, Byrne M. The key principles of cognitive behavioural therapy. InnovAiT 2013 Sep 06;6(9):579-585. [doi: 10.1177/1755738012471029]
49. Copeland ME. Wellness Recovery Action Plan. Updated edition. Sudbury, MA: Human Potential Press, Advocates for Human Potential, Inc; 2018:1-144.
50. Miller WR, Rollnick S. Motivational Interviewing: Helping People Change. 3rd edition. New York, NY: The Guilford Press; 2012.

51. Hood KK, Hilliard M, Piatt G, Ievers-Landis CE. Effective strategies for encouraging behavior change in people with diabetes. Diabetes Manag (Lond) 2015;5(6):499-510 [FREE Full text] [Medline: 30100925]
52. Cameron AC, Miller DL. A practitioner's guide to cluster-robust inference. J Hum Resour 2015 Apr 28;50(2):317-372. [doi: 10.3368/jhr.50.2.317]
53. Guhn M, Forer B, Zumbo BD. Reliable change index. In: Michalos AC, editor. Encyclopedia of Quality of Life and Well-Being Research. Dordrecht, the Netherlands: Springer; 2014:5459-5462.
54. Hastings C, Mosteller F, Tukey JW, Winsor CP. Low moments for small samples: A comparative study of order statistics. Ann Math Stat 1947 Sep;18(3):413-426. [doi: 10.1214/aoms/1177730388]
55. Sagar-Ouriaghli I, Godfrey E, Bridge L, Meade L, Brown JSL. Improving mental health service utilization among men: A systematic review and synthesis of behavior change techniques within interventions targeting help-seeking. Am J Mens Health 2019;13(3):1-18 [FREE Full text] [doi: 10.1177/1557988319857009] [Medline: 31184251]
56. Addis ME, Mahalik JR. Men, masculinity, and the contexts of help seeking. Am Psychol 2003 Jan;58(1):5-14. [doi: 10.1037/0003-066x.58.1.5] [Medline: 12674814]
57. Wang PS, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC. Twelve-month use of mental health services in the United States: Results from the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):629-640. [doi: 10.1001/archpsyc.62.6.629] [Medline: 15939840]
58. Deming WE. An essay on screening, or on two-phase sampling, applied to surveys of a community. Int Stat Rev 1977 Apr;45(1):29. [doi: 10.2307/1403001]
59. Vaughan R. Oversampling in health surveys: Why, when, and how? Am J Public Health 2017 Aug;107(8):1214-1215. [doi: 10.2105/AJPH.2017.303895] [Medline: 28657770]
60. Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics 1982 Dec;38(4):963-974. [Medline: 7168798]
61. Centers for Disease Control and Prevention. 2021 Jun 17. URL: https://www.cdc.gov/obesity/adult/defining.html [accessed 2022-03-03]
62. Renn BN, Feliciano L, Segal DL. The bidirectional relationship of depression and diabetes: A systematic review. Clin Psychol Rev 2011 Dec;31(8):1239-1246. [doi: 10.1016/j.cpr.2011.08.001] [Medline: 21963669]
63. Oser M, Wallace ML, Solano F, Szigethy EM. Guided digital cognitive behavioral program for anxiety in primary care: Propensity-matched controlled trial. JMIR Ment Health 2019 Apr 04;6(4):e11981 [FREE Full text] [doi: 10.2196/11981] [Medline: 30946022]
64. Donkin L, Hickie IB, Christensen H, Naismith SL, Neal B, Cockayne NL, et al. Rethinking the dose-response relationship between usage and outcome in an online intervention for depression: Randomized controlled trial. J Med Internet Res 2013 Oct 17;15(10):e231 [FREE Full text] [doi: 10.2196/jmir.2771] [Medline: 24135213]
65. Saul JE, Amato MS, Cha S, Graham AL. Engagement and attrition in internet smoking cessation interventions: Insights from a cross-sectional survey of "one-hit-wonders". Internet Interv 2016 Sep;5:23-29 [FREE Full text] [doi: 10.1016/j.invent.2016.07.001] [Medline: 30135803]
66. Ainsworth B, Steele M, Stuart B, Joseph J, Miller S, Morrison L, et al. Using an analysis of behavior change to inform effective digital intervention design: How did the PRIMIT website change hand hygiene behavior across 8993 users? Ann Behav Med 2017 Jun;51(3):423-431 [FREE Full text] [doi: 10.1007/s12160-016-9866-9] [Medline: 27909944]
67. Pouwer F, Schram MT, Iversen MM, Nouwen A, Holt RIG. How 25 years of psychosocial research has contributed to a better understanding of the links between depression and diabetes. Diabet Med 2020 Mar;37(3):383-392. [doi: 10.1111/dme.14227] [Medline: 31909844]
68. Nouwen A, Adriaanse MC, van Dam K, Iversen MM, Viechtbauer W, Peyrot M, European Depression in Diabetes (EDID) Research Consortium. Longitudinal associations between depression and diabetes complications: A systematic review and meta-analysis. Diabet Med 2019 Dec;36(12):1562-1572. [doi: 10.1111/dme.14054] [Medline: 31215077]
69. Harding KA, Pushpanathan ME, Whitworth SR, Nanthakumar S, Bucks RS, Skinner TC. Depression prevalence in type 2 diabetes is not related to diabetes-depression symptom overlap but is related to symptom dimensions within patient self-report measures: A meta-analysis. Diabet Med 2019 Dec;36(12):1600-1611. [doi: 10.1111/dme.14139] [Medline: 31532013]
70. Lutes LD, Cummings DM, Littlewood K, Solar C, Carraway M, Kirian K, et al. COMRADE: A randomized trial of an individually tailored integrated care intervention for uncontrolled type 2 diabetes with depression and/or distress in the rural southeastern US. Contemp Clin Trials 2018 Jul;70:8-14. [doi: 10.1016/j.cct.2018.04.007] [Medline: 29680319]
71. Goldstein LA, Mehling WE, Metzler TJ, Cohen BE, Barnes DE, Choucroun GJ, et al. Veterans group exercise: A randomized pilot trial of an integrative exercise program for veterans with posttraumatic stress. J Affect Disord 2018 Feb;227:345-352. [doi: 10.1016/j.jad.2017.11.002] [Medline: 29145076]
72. Brieler JA, Lustman PJ, Scherrer JF, Salas J, Schneider FD. Antidepressant medication use and glycaemic control in co-morbid type 2 diabetes and depression. Fam Pract 2016 Feb;33(1):30-36. [doi: 10.1093/fampra/cmv100] [Medline: 26743722]
73. Kessler RC, Demler O, Frank RG, Olfson M, Pincus HA, Walters EE, et al. Prevalence and treatment of mental disorders, 1990 to 2003. N Engl J Med 2005 Jun 16;352(24):2515-2523 [FREE Full text] [doi: 10.1056/NEJMsa043266] [Medline: 15958807]

74. Galdas PM, Cheater F, Marshall P. Men and health help-seeking behaviour: Literature review. J Adv Nurs 2005 Mar;49(6):616-623. [doi: 10.1111/j.1365-2648.2004.03331.x] [Medline: 15737222]
75. Rochlen AB, McKelley RA, Pituch KA. A preliminary examination of the "Real Men. Real Depression" campaign. Psychol Men Masc 2006 Jan;7(1):1-13. [doi: 10.1037/1524-9220.7.1.1]

Abbreviations

CBT: cognitive behavioral therapy
DMHI: digital mental health intervention
GAD-7: 7-item Generalized Anxiety Disorder scale
HbA_{1c}: hemoglobin A_{1c}
ITT: intention-to-treat
OR: odds ratio
PHQ-8: 8-item Patient Health Questionnaire
RCI: reliable change index

Edited by A Mavragani; submitted 15.03.22; peer-reviewed by R Pine, A Joseph, A Tannoubi; comments to author 09.05.22; revised version received 27.05.22; accepted 08.06.22; published 05.07.22.

Please cite as:
Venkatesan A, Forster B, Rao P, Miller M, Scahill M
Improvements in Depression Outcomes Following a Digital Cognitive Behavioral Therapy Intervention in a Polychronic Population: Retrospective Study
JMIR Form Res 2022;6(7):e38005
URL: https://formative.jmir.org/2022/7/e38005
doi:10.2196/38005
PMID:35788442

Original Paper

Optimization of the Chronic Kidney Disease–Peritoneal Dialysis App to Improve Care for Patients on Peritoneal Dialysis in Northeast Thailand: User-Centered Design Study

Eakalak Lukkanalikitkul[1], MD; Sawinee Kongpetch[1], MD; Wijittra Chotmongkol[1], MD; Michael G Morley[2], MSc, MD; Sirirat Anutrakulchai[3], MD, PhD; Chavis Srichan[4], PhD; Bandit Thinkhamrop[5], PhD; Theenatchar Chunghom[6], RN; Pongsai Wiangnon[6], RN; Wilaiphorn Thinkhamrop[5], MPH, DrPH; Katharine E Morley[7], MPH, MD

[1]Center of Excellence in Kidney Diseases, Srinagarind Hospital, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

[2]Ophthalmic Consultants of Boston, Harvard Medical School, Boston, MA, United States

[3]Division of Nephrology, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

[4]Department of Computer Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen, Thailand

[5]Data Management and Statistical Analysis Center, Faculty of Public Health, Khon Kaen University, Khon Kaen, Thailand

[6]Kidney Unit, Srinagarind Hospital, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

[7]Massachusetts General Hospital Center for Global Health, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Katharine E Morley, MPH, MD
Massachusetts General Hospital Center for Global Health
Harvard Medical School
125 Nashua St. Suite 722
Boston, MA, 02114
United States
Phone: 1 617 726 2000
Email: kemorley@mgh.harvard.edu

Abstract

Background: The prevalence of peritoneal dialysis (PD) in Thailand is increasing rapidly in part because of Thailand's *Peritoneal Dialysis First* policy. PD is a home-based renal replacement therapy in which patients with chronic kidney disease perform up to 4 exchanges of dialysate fluid per day in the peritoneal cavity. Overhydration is one of the most common complications in patients on PD and is associated with increased morbidity and mortality. To monitor hydration status, patients collect hydration metrics, including body weight, blood pressure, urine output, and ultrafiltration volume, from each dialysis cycle and enter this information into a PD logbook. This information is reviewed bimonthly at PD clinic appointments. The chronic kidney disease-PD (CKD-PD) app with near-field communication (NFC) and optical character recognition (OCR) was developed to automate hydration metric collection. The information was displayed in the app for self-monitoring and uploaded to a database for real-time monitoring by the PD clinic staff. Early detection and treatment of overhydration could potentially reduce the morbidity and mortality related to overhydration.

Objective: This study aims to identify usability issues and technology adoption barriers for the CKD-PD app with NFC and OCR and a monitoring system and to use this information to make rapid cycle improvements.

Methods: A multidisciplinary team of nephrologists, PD clinic nurses, computer programmers, and engineers trained and observed 2 groups of 5 participants in the use of the CKD-PD app with NFC and OCR and a monitoring system. The participants were observed using technology in their homes in 3 phases. The data collected included the Unified Theory of Acceptance and Use of Technology questionnaire, *think-aloud* observation, user ratings, completion of hydration metrics, and upload of hydration metrics to the central database. These results were used by the team between phases to improve the functionality and usefulness of the app.

Results: The CKD-PD app with NFC and OCR and a monitoring system underwent 3 rapid improvement cycles. Issues were identified regarding the usability of the NFC and OCR data collection, app stability, user interface, hydration metric calculation, and display. NFC and OCR improved hydration metric capture; however, issues remained with their usability. App stability and user interface issues were corrected, and hydration metrics were successfully uploaded by the end of phase 3. Participants' scores

on technology adoption decreased but were still high, and there was enthusiasm for the self-monitoring and clinical communication features.

Conclusions: Our rapid cycle process improvement methodology identified and resolved key barriers and usability issues for the CKD-PD app with NFC and OCR and a monitoring system. We believe that this methodology can be accomplished with limited training in data collection, statistical analysis, and funding.

(JMIR Form Res 2022;6(7):e37291) doi:10.2196/37291

KEYWORDS

peritoneal; dialysis; peritoneum; mobile health; mHealth; rapid cycle process improvement methodology; home monitoring; near-field communication; monitor; kidney; rapid cycle improvement; quality improvement; process improvement; methodology; nephrology; nephrologist; internal medicine; computer program; Unified Theory of Acceptance and Use of Technology; UTAUT; usability; interface; metric capture; barrier; renal; mobile phone

Introduction

Chronic Kidney Disease and Peritoneal Dialysis in Northeast Thailand

Chronic kidney disease (CKD) is a major health problem in Thailand because of its high prevalence, cost of treatment, significant morbidity and mortality, and substantial impact on the quality of life of patients and their families. From 2017 to 2018, the Chronic Kidney Disease Prevention in the Northeast of Thailand (CKDNET) project found that the prevalence of CKD was 27% in the rural provinces of northeast Thailand [1], primarily because of rising rates of diabetes, hypertension, and other primary renal diseases [2]. Despite this increased need, there is only 1 nephrologist for every 593,000 population compared with every 44,000 population in Bangkok [3]. In 2008, the Thai government adopted the *Peritoneal DialysisFirst* policy for renal replacement therapy under its universal health care coverage scheme, increasing access in low-resource settings, with approximately 21% of patients on peritoneal dialysis (PD) in northeast Thailand [4]. Patients on PD manage their PD at home manually or by using a PD cycler to deliver dialysate fluid through a catheter placed in the peritoneal cavity, where the fluid remains for several hours. They enter hydration metric data, including body weight, blood pressure, urine output, and ultrafiltration volume for each cycle, in handwritten notebooks for review by a nephrologist at bimonthly clinic appointments.

User Design and Evaluation of Successful Adoption of the CKD-PD App

In 2018, the Data Management and Statistical Analysis Center, Faculty of Public Health, Khon Kaen University, developed the CKD-PD mobile app to help nephrologists and patients on PD manage fluid status to prevent overhydration. It is a common complication in patients on PD. Overhydration increases morbidity because of PD-related peritonitis, stroke, congestive heart failure, major adverse cardiac events, and mortality [4-9]. The CKD-PD app has been of interest not only to nephrologists but also to the Thai Health Security Office as an intervention for improving the care of patients on PD. Early treatment of overhydration can decrease these related complications, thereby reducing hospitalization and health care expenses [10].

Many apps fail because of a lack of evaluation and removal of barriers to user adoption, fidelity of the technology, and design of the health care delivery system in which the app will be deployed [11,12]. Achieving rapid design and deployment of digital health interventions are 2 challenges facing the successful adoption and implementation of mobile health (mHealth) technologies [13]. This is especially true in low- and middle-income countries where mHealth interventions have been touted as solutions to a wide variety of health care challenges [14-16]; however, little is known about how they perform [17-19]. Before studying the effectiveness of the CKD-PD in managing overhydration in a real-world setting, a user design study was conducted using rapid cycle process improvement methods.

The objectives of this study were to (1) optimize the design and usability of the CKD-PD app and test the app in the context in which it is deployed using rapid cycle process improvement methods; (2) evaluate automatic data entry features using near-field communication (NFC) and optical character recognition (OCR) technology; and (3) identify and address the practical challenges that influence the successful adoption of the CKD-PD app and remote monitoring system in a real-world, low-resource setting.

Methods

Study Population

The study was conducted at Srinagarind Hospital, Khon Kaen University, between November 1, 2020, and April 30, 2021. Patients on PD for >3 months without a change in their PD prescription were invited to participate if they met the following inclusion criteria: aged ≥18 years, having access to a smartphone capable of running the CKD-PD app, and willing to allow research staff to observe their use of the CKD-PD app with NFC and OCR features and monitoring equipment in their home. Vulnerable populations specified as pregnant women, children, prisoners, individuals who were institutionalized, or those unable to participate in home data collection were excluded. Informed consent was obtained by the trained research staff. A total of 10 participants were enrolled and divided into group 1 (participants 1-5) and group 2 (participants 6-10). Baseline demographic data, including age, sex, education level, time on PD, and whether they used continuous ambulatory PD (CAPD) or automated PD (APD), were collected at enrollment. Participants were given prepaid cards to cover the cost of study related to internet or cellular data expenses.

CKD-PD App and Its Features

The CKD-PD app is available for free download in Android and iOS formats and is designed for daily hydration metric collection for patients on PD (Multimedia Appendix 1). Users can enter hydration metrics using manual input or voice recognition. The CKD-PD app graphically displays a patient's hydration metrics over time and uploads them to the CKDNET database stored in the Thai Care Cloud data repository [20], which is accessible to the PD clinic staff. Nephrologists can set individual hydration parameters for a patient so that they can self-monitor their hydration status. Alerts can be set to notify patients on PD and PD clinic staff when there are actionable hydration metric abnormalities, triggering a prompt review by a nephrologist and allowing the early detection and treatment of overhydration. Another key feature of the CKD-PD app is a direct link using LINE, a social messaging app widely used in Thailand, between the patient and PD clinic, facilitating communication about symptoms and changes in PD management.

A new prototype using NFC and OCR to automate the entry of hydration metric data into the CKD-PD app from measurement devices was developed. NFC uses radio frequency communication and wirelessly transmits data to an NFC-enabled device when it is placed within 4 cm. This offers a simple, low-cost solution that does not require an external power source or pairing, similar to Bluetooth. It is commonly used for mobile payments ad transit cards and, more recently, with subdermal glucose sensors [21-25]. NFC radio frequency tags with unique ID numbers can be added to medical devices such as scales. Users tap a card with an NFC receiver tag on the NFC-equipped scale and then tap it on the NFC card reader, and the data are transferred to the CKD-PD (Multimedia Appendix 2). The CKD-PD app is also equipped with OCR technology, which uses a smartphone camera to capture the digital output from the blood pressure machine and store it in the CKD-PD app.

Study Design

Overview

There were 3 user design phases separated by 2 improvement cycles. During the improvement cycle, the app improvement team analyzed the results from the previous phase and revised the CKD-PD app using the NFC, OCR, and monitoring system (Table 1). The app improvement team included study nephrologists, computer engineers and app developers, and the Data Management and Statistical Analysis Center Thai Care Cloud database team. The app improvement team met as needed to define technology and user design issues, develop solutions, and test modifications.

Table 1. Overview and timeline of research activities.

Research activity	Phase 1 (group 1)				IC[a] 1	Phase 2 (group 2)				IC 2	Phase 3 (groups 1 and 2)	
	Week				Week	Week				Week	Week	
	0	1-2	3	4	5-8	8	9-10	11	12	12-16	17-18	19-20
In-clinic training and observation	✓					✓						
UTAUT[b] survey	✓			✓		✓			✓		✓	✓
Home observation		✓					✓				✓	✓
Completion of hydration metrics			✓	✓				✓	✓			✓
Contact with PD[c] clinic			✓	✓				✓	✓			✓
Validation of hydration metrics				✓					✓			✓
Observation logbook											✓	

[a]IC: improvement cycle.

[b]UTAUT: Unified Theory of Acceptance and Use of Technology.

[c]PD: peritoneal dialysis.

Phase 1

Overview

Group 1 participants received training on the use of the CKD-PD app with the NFC and OCR and monitoring system during week 0. In-home observations were conducted once during weeks 1 to 4, followed by research activities 2 to 3 (see the following sections) during weeks 1 and 2. During weeks 3 and 4, the participants used the CKD-PD app with the NFC and OCR system at their homes and completed research activities 4 to 6 (see in the following sections).

Improvement Cycle 1

After completion of phase 1, the results were summarized and presented with the completion of hydration metrics and clinical contact results to the app improvement team. The app improvement team analyzed the results and modified the components of the system, including measurement devices, NFC, OCR, app design, and programming. The user processes were also adjusted.

Phase 2

Overview

Group 2 participants received training in the use of the CKD-PD app with the NFC and OCR and monitoring system during week

8. In-home observations were conducted during weeks 9 to 12 using the modified CKD-PD app with the NFC and OCR and monitoring system. Research activities 2 to 3 were conducted during weeks 9 and 10, followed by research activities 4 to 6 in weeks 11 and 12.

Improvement Cycle 2

After the completion of phase 2, the app improvement team reviewed the new results from participant observations, hydration metric completion, and clinic contacts. Additional issues were identified, and new solutions were developed and tested. Phase 3 was launched after the completion of the improvement cycle 2 modifications.

Phase 3

Overview

Participants from groups 1 and 2 performed research activities 2 to 6 again during weeks 17 to 20 using the modified CKD-PD app and monitoring system modified without NFC and OCR. This allowed for the evaluation of the usability and functionality of the CKD-PD app without NFC and OCR. In phase 3, participants were asked to rate the entry of body weight and dialysate volumes with NFC again based on their earlier experience for comparison with phases 1 and 2. In addition, research activity 7 (see the following sections) was performed to compare using the usual practice of entering their hydration metrics by hand into a logbook with the CKD-PD app.

Final Improvements

After completion of phase 3, the app improvement team made final improvements based on the results of the phase 3 home observation and hydration metric collection using the CKD-PD app and monitoring system without NFC and OCR.

Description of Study Activities

The details of the study activities are described in the following sections.

In-Clinic Training

Participants were trained to use the CKD-PD app with NFC and OCR and home monitoring equipment in the PD clinic by research staff and PD nurses, starting with the CKD-PD app downloaded onto the participant's mobile phone and account registration. Research assistants demonstrated the use of the CKD-PD app, followed by instructions on how to use the NFC and home monitoring equipment, how the uploaded hydration metrics can be monitored by the PD clinic, and how they can self-monitor at home on the app. A set of NFC hydration metric data collection equipment (body weight scale, NFC card, NFC card reader and connectors, and blood pressure machine) was prepared for each participant to use at home. They were trained to set up the equipment using a teaching video (Multimedia Appendix 2) followed by hands-on practice with a research assistant. During the study period, the participants were instructed to use the NFC data collection system and equipment, in addition to the standard method of recording their hydration metrics in a logbook.

Unified Theory of Acceptance and Use of Technology Survey

A structured interview questionnaire based on the Unified Theory of Acceptance and Use of Technology (UTAUT) model was completed at the beginning and end of each phase to collect perceptions of user technology acceptance and usability [26]. The questionnaire has 6 domains, each with 3 questions representing different factors affecting technology use and adoption. It was translated into the Thai language, and each question was scored on a 5-point Likert scale, with 1=strongly disagree and 5=strongly agree. The final score for each domain is the sum of the scores of the 3 questions for that domain and ranges from 3=strongly disagree to 15=strongly agree. The questionnaire was explained by the research staff and self-administered by the participants (Multimedia Appendix 3).

Home Observation

Research staff conducted the home observation of participants using the CKD-PD app with NFC and OCR using an observation guide (Multimedia Appendix 4), using the *think-aloud* method [27]. They asked the participants what they *liked and disliked* about each task and feature. Participants were asked to rate each feature as follows: 1=good, 2=neutral, and 3=not good. This information was recorded by research assistants using handwritten notes and summarized for use by the app improvement team.

Completion of Hydration Metrics

During the 2-week home use period, the number of times the participant successfully uploaded each of the required hydration metric values (body weight, blood pressure, and use of the CKD-PD app) was recorded. Participants using CAPD required 4 dialysate exchange volumes per day, whereas those using APD only required 1 each day; thus, the number of required hydration metrics varied among participants. Successful completion of hydration metrics included the entry of hydration metrics into the app and accurate upload to the CKDNET database.

Contact With PD Clinic

The number of times the participants contacted the PD clinic during all 3 phases was collected, along with the reason for contact.

Validation of Hydration Metrics

The hydration metrics collected by each participant were validated by comparing the values uploaded to the CKDNET database with the results recorded in their logbooks during the study period. Participants sent screenshots from the app, or the research staff reviewed the data on the participant's smartphone to confirm the correct data entry in the CKD-PD app. Successful validation of hydration metrics was defined by the completion of data entry by the participant and agreement between values recorded in their logbook and the uploaded results.

Observation of Logbook Use

Participants were observed while entering hydration metrics in their logbooks. They were asked what they *liked and disliked*

about each task and feature and rated each feature as 1=good, 2=neutral, and 3=not good.

Ethics Approval

This study was approved by the Ethics Committee for Human Research, Faculty of Medicine, Khon Kaen University, Thailand (project number HE621494), and the Mass General Brigham institutional review board (protocol number 2019P002648). All participants provided written informed consent in the Thai language.

Results

Participant Characteristics

Phase 1 participants characteristics had a mean age of 46 years (SD 10.3) and a mean time on PD of 5.5 years (SD 3.8); were female (2/5 40%); used the PD method of CAPD (3/5, 60%) or APD (2/5, 40%); and had educational backgrounds of high school (2/5, 40%), bachelor's degree (1/5, 20%), and postbachelor's degree (2/5, 40%).

Phase 2 participant characteristics had a mean age of 48 years (SD 15.7) and a mean time on PD of 1.3 years (SD 2.5); were female (3/5, 60%); used the PD method of CAPD (1/5, 20%) or APD (4/5, 80%); and had educational backgrounds of high school (2/5, 40%), bachelor's degree (2/5, 40%), and postbachelor's degree (1/5, 20%). One of the group 2 participants did not complete phase 3 because of sudden death.

In comparison, the characteristics of patients on PD at Srinagarind Hospital were as follows: mean age 49 years (SD 13); mean time on PD 3 years (SD 29.7); female (33/74, 45%); PD method CAPD (42/74, 57%) or APD (32/74, 43%); and educational background of primary school (37/74, 50%), high school (12/74, 16%), bachelor's degree (23/74, 31%), and postbachelor's degree (2/74, 3%).

UTAUT Survey

The UTAUT survey scores for each participant are presented in Table 2 as the total score (sum of all 6 domains) and the scores for each domain. The difference in the total score between the beginning of phase 1 or phase 2 and the end of phase 3 is also presented. Detailed results, including the scores at the beginning and end of all phases, are available in Multimedia Appendix 5.

Table 2. Difference in Unified Theory of Acceptance and Use of Technology total scores[a] by participant and domain between the beginning of phase 1 or 2 and end of phase 3 (detailed results in Multimedia Appendix 5).

Participant number[b]	1	2	3	4	5	6	7	8	9	10	Values, mean
Total score											
Phase 1 or 2	89	63	86	81	87	79	66	75	71	77	77
Phase 3	74	76	80	74	77	72	N/A[c]	80	79	72	76
Difference	−15	13	−6	−7	−10	−7	N/A	5	8	−5	−2.7
Difference by domains											
Performance expectancy	−2	3	1	1	−1	−1	N/A	1	1	0	0.33
Effort expectancy	−2	4	1	−1	1	−1	N/A	0	0	1	0.33
Social influence	−3	3	1	0	−2	−1	N/A	−1	1	−1	−0.67
Voluntariness	−3	−3	−5	−2	−3	−3	N/A	2	3	−4	−2
Intention to use	−2	3	−4	−4	−6	0	N/A	2	3	0	−0.9
Facilitating conditions	−3	3	0	−1	1	−1	N/A	1	0	−1	−0.1

[a]Scores for individual questions from 1 (strongly disagree) to 5 (strongly agree). The domain scores ranged from 5 to 15. The total score for all the domains ranged from 30 to 90.

[b]Participants 1 to 5: phase 1 and phase 3; participants 6 to 10: phase 2 and phase 3.

[c]N/A: not applicable; participant 7 expired before phase 3.

Of the 9 participants who completed phase 1 or 2 and phase 3, 6 (67%) individuals had a decrease in the total UTAUT score, ranging from −5 to −15 points. There were 33% (3/9) of individuals who had an increase in the total UTAUT score, ranging from 5 to 13 points. The mean score at the beginning of phases 1 or 2 was 77 (SD 8.8), with a range of 63 to 89. The mean score at the end of phase 3 was 76 (SD 3.2), with a range of 72 to 80. In group 1, the domains of voluntariness (−3) and intention to use (−2.6) showed the largest decrease in mean difference between the beginning of phase 1 and the end of phase 3, with all participants reporting a negative score in voluntariness and 80% (4/5) of participants reporting a negative score in intention to use. In group 2, there were small decreases in voluntariness (−0.5), social influence (−0.5), and facilitating conditions (−0.25) and an increase in intention to use (1.25) between phases 2 and 3 (Multimedia Appendix 5).

Participant Observation

The home observation results for each participant using the CKD-PD app with the NFC and OCR and monitoring system during the home observations are summarized by phase for all participants in Table 3, with detailed responses in Multimedia Appendix 6. Feature ratings are provided for each phase. The participants rated almost all features as *good*, except for some

data entry tasks using NFC, which were rated as *neutral*. Participants liked automatic data entry with NFC; however, technical issues such as slow transfer of time and difficulty tapping the card resulted in lower scores. Participants were asked to rate the entry of body weight and dialysate volume using NFC at the start of phase 3 based on their experience in phases 1 or 2 for comparison with manual data entry used in phase 3. The average rating for data entry with NFC was 2.3 for each task compared with 1.1 for manual data entry (Multimedia Appendix 6). Scores for all tasks using the handwritten logbook were nearly all higher than the CKD app with or without NFC and OCR, consistent with the preference for using the app.

Table 3. Feature and task ratings[a] from participant observation in phases 1, 2, and 3 (detailed results in Multimedia Appendix 6).

Feature	Phase 1[b], mean (range)	Phase 2[c], mean (range)	Phase 3[d], mean (range)	Phase 3[e], mean (range)
Open screens	1.1 (1-2)	1 (1)	1.6 (1-3)	N/A[f]
Enter body weight	1.1 (1-2)	1.2 (1-2)	1.1 (1-2)	1.4 (1-2)
Enter blood pressure	1.1 (1-2)	1.6 (1-3)	1.1 (1-2)	1.4 (1-3)
Enter dialysate	1.1 (1-2)	1.2 (1-2)	1.1 (1-2)	1.8 (1-3)
View metrics	1.3 (1-2)	1 (1)	1.1 (1-2)	1.7 (1-3)
Interpret metrics	1.1 (1-2)	1.2 (1-2)	1.1 (1-2)	N/A
Communications	1 (1)	1 (1)	1 (1)	1.2 (1-3)
User incentives	1.1 (1-2)	1.2 (1-2)	1.2 (1-3)	1.9 (1-3)

[a]Rating of task or feature: 1=good, 2=neutral, and 3=not good.

[b]Group 1 using the chronic kidney disease–peritoneal dialysis app with near-field communication or optical character recognition data entry; participants 1 to 5.

[c]Group 2 using the chronic kidney disease–peritoneal dialysis app with near-field communication or optical character recognition data entry; participants 6 to 10.

[d]Group 1 and 2 using the chronic kidney disease–peritoneal dialysis app with manual entry; participants 1 to 10, excluding 7.

[e]Group 1 and 2 using logbook; participants 1 to 10, excluding 7.

[f]N/A: not applicable.

In general, participants found that manual data entry was easier to perform than using the NFC and OCR system. The use of NFC and OCR presented multiple challenges. Some of these issues were solved by providing tools such as an extension device for tapping the card to eliminate the need to bend over and touch the card to the scale on the floor and training participants to wait long enough for data transfer. Some issues were more difficult to troubleshoot within the time frame of this study, such as (1) lack of space in participants' homes for NFC setup and access to power outlets; (2) internet instability resulting in slow data transfer; (3) size and design of the NFC-assembled weight scale, making it difficult for older or obese patients to stand on; and (4) lack of an alert to indicate that the NFC device was ready to place the NFC card for reading. Most patients also did not like to use the OCR function to enter the blood pressure readings because of poor image clarity and variations in ambient home lighting conditions.

Several important issues related to the functionality of the CKD-PD app were detected and corrected. The dialysate in and out volumes were reported by PD cycle, and the net daily ultrafiltration volume was not accurately calculated and displayed. There were issues with the display of hydration metrics, such as incorrect scale on the graph and a lack of previous results for comparison. During home observation of participants using the CKD-PD app with the NFC and OCR and monitoring system, participants expressed concerns about the usability of the app. An example was slowness when opening each icon, which was determined to be from opening multiple apps at once, and slow cellular or Wi-Fi network speeds were contributing factors. BMI was misinterpreted as *overweight* because of translation issues between the English and Thai languages. This was corrected by changing the wording so it would not be confused with *weight*. Participant height had to be entered daily to calculate the BMI. As this measurement did not change, the app was modified so that it automatically entered the participant height obtained daily. The original font was small and difficult to read. This was adjusted, and the readability of the screens was improved.

Completion and Validation of Hydration Metrics

The percentage completion of the hydration metrics for each participant was collected for all 3 phases (Table 4). The mean percentages of completion in phases 1 and 2 were 88% (SD 19) and 83% (SD 6.3), respectively, with a range of 57% to 100%. This was compared with the mean percentage of completion in phase 3 for all participants (68%, SD 18.6, range 18%-100%; Multimedia Appendix 7). The most common reasons for not collecting the hydration metrics in phases 1 and 2 were not knowing how to enter the data, forgetting to enter the data, problems with the system uploading the data to the CKDNET database, and switching from Android to iOS systems. The reasons for incomplete hydration metric collection in phase 3 (not using NFC) were forgetting to send the data to the CKDNET and switching from Android to iOS.

Table 4. Percentage completion of hydration metrics (detailed in Multimedia Appendix 7).

Participant	Phase 1[a]		Phase 2[b]		Phase 3[c]		Difference (%)[d]
	Total, N	Done, n (%)	Total, N	Done, n (%)	Total, N	Done, n (%)	
1	84	80 (95)	__[e]	—	42	19 (45)	−50
2	66	39 (59)	—	—	84	42 (50)	−9
3	42	42 (100)	—	—	42	26 (62)	−38
4	84	73 (87)	—	—	84	79 (94)	7
5	48	48 (100)	—	—	56	40 (71)	−29
6	—	—	42	38 (91)	42	42 (100)	9
7[f]	—	—	39	35 (90)	N/A[g]	N/A	N/A
8	—	—	42	37 (88)	42	39 (93)	5
9	—	—	30	24 (57)	84	15 (18)	−39
10	—	—	80	27 (90)	42	33 (79)	−11
Values, mean	65	56 (88)	46.6	32 (83)	57.6	37 (68)	−15

[a]Phase 1: participants 1 to 5.

[b]Phase 2: participants 6 to 10.

[c]Phase 3: participants 1 to 10.

[d]Difference in percentage completion between phase 1 or 2 and phase 3.

[e]Not available.

[f]Participant 7 expired before completion of the study.

[g]N/A: not applicable.

Of the 9 participants who entered hydration metric data in phase 1 or 2 and phase 3, there were 4 participants who had a greater than 29 percentage point decrease in their hydration metric completion. The other 5 participants had minimally decreased or increased hydration metric completion, ranging from −11% to 9%.

Validation of the hydration metrics improved from phases 1 or 2 and phase 3 (Table 5). In phases 1 and 2, the main issues were the upload of inaccurate body weight and ultrafiltration metric results to the CKDNET. This was because of problems with the NFC weight scale and app calculation of the ultrafiltration volume. In phase 3, these issues were resolved, the hydration metrics were fully validated with participants correctly entering the data in the CKD-PD app, and the results entered were accurately uploaded to the CKDNET.

Table 5. Validation of hydration metrics.

Participant	1	2	3	4	5	6	7	8	9	10	Values, n (%)[a]
Phase 1[b]											
Body weight	No[c]	No	No	No	No	—[d]	—	—	—	—	0 (0)
Blood pressure	Yes[e]	Yes	No	Yes	Yes	—	—	—	—	—	4 (80)
Dialysate in	Yes	No	Yes	No	Yes	—	—	—	—	—	3 (60)
Dialysate out	Yes	No	Yes	No	Yes	—	—	—	—	—	3 (60)
Ultrafiltration volume	Yes	No	Yes	No	No	—	—	—	—	—	2 (40)
Phase 2[f]											
Body weight	—	—	—	—	—	Yes	No	Yes	No	Yes	3 (60)
Blood pressure	—	—	—	—	—	Yes	Yes	Yes	Yes	Yes	5 (100)
Dialysate in	—	—	—	—	—	Yes	Yes	Yes	Yes	Yes	5 (100)
Dialysate out	—	—	—	—	—	Yes	Yes	Yes	Yes	No	4 (80)
Ultrafiltration volume	—	—	—	—	—	Yes	No	Yes	Yes	No	3 (60)
Phase 3[g]											
Body weight	Yes	ND[h]	ND	Yes	Yes	Yes	N/A[i]	Yes	Yes	Yes	7 (100)
Blood pressure	Yes	Yes	Yes	Yes	Yes	Yes	N/A	Yes	Yes	Yes	9 (100)
Dialysate in	Yes	Yes	Yes	Yes	Yes	Yes	N/A	Yes	Yes	Yes	9 (100)
Dialysate out	Yes	Yes	Yes	Yes	Yes	Yes	N/A	Yes	Yes	Yes	9 (100)
Ultrafiltration volume	Yes	Yes	Yes	Yes	Yes	Yes	N/A	Yes	Yes	ND	9 (100)

[a]Number and percentage of hydration metric values entered and uploaded with the correct results to Chronic Kidney Disease Prevention in the Northeast of Thailand.

[b]Phase 1: participants 1 to 5.

[c]Not uploaded accurately to Chronic Kidney Disease Prevention in the Northeast of Thailand.

[d]Not available.

[e]Uploaded accurately to Chronic Kidney Disease Prevention in the Northeast of Thailand.

[f]Phase 2: participants 6 to 10.

[g]Phase 3: participants 1 to 10.

[h]ND: not done by participant.

[i]N/A: not applicable; participant 7 expired before completion of the study.

Clinic Contacts

The number of times each participant contacted the PD clinic during each phase is presented in Table 6. All contacts were made using the Line messaging app by chat, audio, or video calls, except for1 in-person contact during phase 1. In phase 1, there were 104 clinic contacts compared with 13 and 22 in phases 2 and 3, respectively, and most were related to issues using the CKD-PD app. In phase 3, 59% (13/22) of the contacts were related to clinical concerns.

Table 6. Number and reason for contacting the peritoneal dialysis clinic by phase.

Phase	Total number of contacts, N	Clinical issue, n (%)	App issue, n (%)	Other, n (%)
Phase 1	104	0 (0)	99 (95.2)	5 (4.8)
Phase 2	13	3 (23.1)	10 (76.9)	0 (0)
Phase 3	22	13 (59.1)	9 (40.9)	0 (0)

Improvement Team Activities

After the completion of each phase, the research team summarized the findings for the app improvement team (Multimedia Appendix 8). In phase 1, troubleshooting issues with NFC and OCR were the focus of the improvement activities, along with app design issues. In phase 2, additional work was performed with the NFC data entry; however, more focus was given to resolving issues with how the hydration metric data were displayed and uploaded to the CKDNET database. In phase 3, the NFC prototype was not used as the participants preferred to input data manually into the CKD-PD

app, and the upload of the hydration metric data to the CKDNET database was not reliable using NFC. The app improvement team made final improvements to improve the stability of the data transfer from the CKD-PD app to the CKDNET database, added the total daily ultrafiltration volume to the CKDNET, and set parameters for overhydration alerts.

Discussion

Principal Findings

Overview

mHealth apps are touted as having great potential to transform the delivery of health care; however, the real-world development of mHealth apps in low- and middle-income countries reveals few success stories [12,17]. Often, a good app idea fails to succeed as the design does not meet the users' needs and is difficult to use. Our research describes how a rapid cycle process improvement strategy helped researchers understand the benefits and challenges of using an mHealth app for use by patients on PD in a low-resource setting and optimize its features. The process revealed valuable insights into the factors influencing user attitudes, identifying technological and design flaws, and addressing barriers to user adoption of the CKD-PD app and monitoring system. Moreover, we believe that this methodology can be accomplished without requiring significant training in data collection, statistical analysis, or funding. The general principles of the user design process are applicable to a wide variety of locations, contexts, and subject domains.

User Adoption

The total scores for all the participants for most domains were in the *agree* and *strongly agree* categories, demonstrating a strong willingness to adopt the technology at the beginning of the study. Over the course of the study, 67% (6/9) of participants who completed both assigned phases demonstrated decreased interest in using the technology, although the total scores remained in the *agree* range or above. The decreases in the total score were primarily driven by decreases in the domains of voluntariness and intention to use. We believe this is likely because of issues with the NFC and OCR data entry features and app speed and functionality. We acknowledge that the UTAUT questionnaire is not typically used to evaluate technology adoption over time; however, we found that it provided insights into how our participants' views changed over the course of using the CKD-PD app and monitoring system, although they had limited generalizability. The small sample size limited the statistical validity of our findings. However, we found the information useful as a *sentiment analysis* regarding user experience and provided insights into how different individuals may adopt the CKD-PD app.

CKD-PD App Issues

Design issues centered on simplicity and ease of use compared with a pencil and paper logbook. Simplifying the use of the app, for example, opening the app and the organization of the screens, readability of the fonts, input of the data, and graphical representation of trends, clearly needed to be refined based on user feedback. PD requires a substantial commitment of time to collect and record data from patients and their families throughout the day, whether in a logbook or an app. Our participants generally reacted positively to the self-monitoring features and quickly engaged with this new functionality but became frustrated when some features did not work well for them.

Technology Issues

Technology issues included problems with the NFC devices, unstable internet access, slow internet speed, app stability, uploading data to the CKD-NET cloud, and software issues causing it to freeze or not upload data properly. The participant observation revealed that NFC and OCR were good solutions for automatic data entry but that they need to be seamlessly integrated into the measurement devices; for example, the location of NFC on the scales on the floor was inconvenient, slow data transfer, and poor quality image capture with OCR. Despite these barriers, participants liked the concept, and with further development, we believe that NFC and OCR could be simple and effective methods of mitigating the barrier of data entry [28,29], as has been done with other devices such as glucose monitors.

Internet access, especially in low-income rural areas, is frequently slow or unstable, making the use of digital apps cumbersome. Some patients were concerned about the cost of sending or receiving data, although internet access was subsidized for the participants in this study. These will be challenging issues to remedy; however, if successful in reducing complications from overhydration, it may prove cost-effective in the long run to provide financial support for internet access for patients on PD, given the alternative costs of not using the app.

Programming issues resulting in app instability and data upload errors were critical problems that were identified. These were fixable by our software engineers and resulted in improved validation of the hydration metrics in phase 3. The completion of the hydration metrics dropped between phases 1 and 3. We believe this occurred as participants in phases 1 and 2 used NFC, which automatically uploaded the hydration metrics, and decreased after the team elected not to use NFC and OCR in phase 3 because of usability issues and inaccurate upload of the hydrate metrics to the CKDNET with NFC and OCR. The hydration metric accuracy improved with this change and resulted in 100% validation of the hydration metrics in phase 3. However, when manual entry was used in phase 3, participants had to remember to upload their data, resulting in incomplete hydration metrics. In the final improvements, a reminder alert addressed this issue.

Communication Issues

Contact with the PD clinic was much more frequent in phase 1 than in phase 2, although the participants in both phases were new users of the app. The study staff felt that this was because of differences in individual comfort when using the CKD-PD app. In phase 3, there were fewer contacts about the app, although there were 10 participants compared with 5 in phases 1 and 2. The contacts in phase 3 were about clinical concerns, suggesting increased confidence in self-monitoring features. In general, participants gave high ratings and positive comments

about the ease of communication with the PD clinic using the CKD-PD app.

Impact on Future Deployment and Development of CKD-PD

Our research findings provided critically useful information for the optimization of the CKD-PD app and monitoring system. We plan to use it in a randomized controlled trial to further evaluate the efficacy of CKD-PD in the early detection and treatment of overhydration.

We have developed training materials for CKD-PD app users based on the insights gained from this study and plan to share them with future CKD-PD app users. We have improved awareness of user adoption issues and realize that this will be an iterative process as more patients on PD use the app. We believe that with additional engineering and design work, NFC is a potential solution for automating data entry, and we plan to pursue the further development of these features.

Limitations

Our study has several limitations. The NFC system and CKD-PD app were initially tested together; however, the NFC issues were too complicated to solve within the time frame and budget of the study period. Through targeted observations and data collection, we were able to tease out NFC and data entry issues from the app design and cloud-based monitoring processes.

We identified potential sources of bias. Observation bias could have influenced our results as participant observations were conducted by research staff from the PD clinic. Patients on PD, cared for by nephrologists and PD nurses from their clinic, may be inclined to report feedback they thought the PD staff wanted to hear. Confirmation bias may also be a concern as research staff may be looking for patterns or use issues that confirm opinions they already hold about the CKD-PD app, NFC and OCR data entry, and the monitoring system. Our sample size was small; however, this is consistent with recommendations for user design studies, indicating that a few participants can uncover most usability issues [30]. We also acknowledge the potential for selection bias in our participants, given their higher educational level than the average population in rural northeast Thailand, although the participants were representative of the Srinagarind Hospital PD patient population. We selected participants who had baseline comfort with smartphones for the user design study, although it limits generalizability as they were more interested and engaged in the design process.

An additional weakness is that we did not conduct an in-depth investigation into why our participants responded to the UTAUT questionnaire as much as they did. Although it has been validated and used in Thai populations [26,31], we did not consider the impact of age, gender, education, and cultural influences on UTAUT scores [32,33]. We acknowledge that UTAUT results have limited generalizability.

Comparison With Prior Work

Our research included several recommended strategies for mHealth app development and evaluation [34,35]. Our iterative rapid cycle process improvement approach used a multidisciplinary team, including patients, computer engineers and programmers, nephrologists, and PD nurses, to evaluate the CKD-PD app. We conducted our study activities in the participants' homes and PD clinics where it would be used [36,37]. Our structured observation using the think-aloud method provided the app improvement team with a broad sense of what worked and what did not work, in addition to specific actionable feedback. Finally, our study participants were selected from the same study population as those who will use the CKD-PD app in northeast Thailand.

Our research team found several critical usability and functionality issues during the study. This experience reinforces the importance of these steps before an mHealth app can be successfully implemented [12,13,17]. Although our current NFC and OCR data entry features are not optimized, our results support the idea that app users want these features [27].

Conclusions

In conclusion, using rapid cycle progress improvement methodologies with a multidisciplinary team proved to be a useful strategy for optimizing the CKD-PD app and monitoring system. Our experience aligns with insights and recommendations regarding mHealth technology design and evaluation that using a multidisciplinary team in the context of the system in which it is used is essential for successful deployment. This process revealed critical issues that once addressed, position the CKD-PD app and monitoring system to achieve its potential to improve health outcomes.

Acknowledgments

The research reported in this publication was supported by the Fogarty International Center of the National Institutes of Health under award number R21TW010963. The content is the sole responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors would like to acknowledge the Chronic Kidney Disease Northeast Thailand research project staff for their assistance in organizing and assisting with home observations. The authors would like to thank Ms Thankyalak Rattanasawad; the staff of the Center of Excellence in Kidney Diseases, Srinagarind Hospital, Faculty of Medicine at Khon Kaen University; Dr Jessica Haberer and the Massachusetts General Hospital Center for Global Health; Dr Nathaphop Chaichaya, Data Management and Statistical Analysis Center, Faculty of Public Health, Khon Kaen University Thailand; the other staff from the Data Management and Statistical Analysis Center and the Faculty of Public Health at Khon Kaen University; and Pobporn Danvirutai from T. Robotics, Co Ltd, Bangkok, Thailand.

Authors' Contributions

SA, EL, KEM, MGM, and BT designed the study, including working with the Data Management and Statistical Analysis Center app development team. CS developed and oversaw the use of near-field communication or optical character recognition and integration with the chronic kidney disease–peritoneal dialysis app. EL, SK, WC, TC, and PW collected and analyzed the data and drafted the manuscript. SA, KEM, and MGM assisted with the data analysis and critically revised the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Chronic Kidney Disease–Peritoneal Dialysis app screenshots.
[PDF File (Adobe PDF File), 2613 KB - formative_v6i7e37291_app1.pdf]

Multimedia Appendix 2
Using Chronic Kidney Disease–Peritoneal Dialysis app (near-field communication version).
[MP4 File (MP4 Video), 15806 KB - formative_v6i7e37291_app2.mp4]

Multimedia Appendix 3
Unified Theory of Acceptance and Use of Technology questionnaire on adoption and use of the Chronic Kidney Disease–Peritoneal Dialysis app.
[PDF File (Adobe PDF File), 95 KB - formative_v6i7e37291_app3.pdf]

Multimedia Appendix 4
Participant observation guide.
[PDF File (Adobe PDF File), 42 KB - formative_v6i7e37291_app4.pdf]

Multimedia Appendix 5
Unified Theory of Acceptance and Use of Technology scores by participant and domain between phases 1 and 3 or phases 2 and 3.
[PDF File (Adobe PDF File), 174 KB - formative_v6i7e37291_app5.pdf]

Multimedia Appendix 6
Summary of participant observation phases 1, 2, and 3.
[PDF File (Adobe PDF File), 74 KB - formative_v6i7e37291_app6.pdf]

Multimedia Appendix 7
Detailed hydration metric completion.
[PDF File (Adobe PDF File), 78 KB - formative_v6i7e37291_app7.pdf]

Multimedia Appendix 8
Recommendations from the app improvement team.
[PDF File (Adobe PDF File), 55 KB - formative_v6i7e37291_app8.pdf]

References

1. Cha'on U, Wongtrangan K, Thinkhamrop B, Tatiyanupanwong S, Limwattananon C, Pongskul C, CKDNET group. CKDNET, a quality improvement project for prevention and reduction of chronic kidney disease in the Northeast Thailand. BMC Public Health 2020 Aug 27;20(1):1299 [FREE Full text] [doi: 10.1186/s12889-020-09387-w] [Medline: 32854662]
2. Anutrakulchai S, Mairiang P, Pongskul C, Thepsuthammarat K, Chan-On C, Thinkhamrop B. Mortality and treatment costs of hospitalized chronic kidney disease patients between the three major health insurance schemes in Thailand. BMC Health Serv Res 2016 Sep 29;16(1):528 [FREE Full text] [doi: 10.1186/s12913-016-1792-9] [Medline: 27686066]
3. Thailand Renal Replacement Therapy: Year 2020. Bangkok: The Nephrology Society of Thailand; 2020.
4. Jotterand Drepper V, Kihm LP, Kälble F, Diekmann C, Seckinger J, Sommerer C, et al. Overhydration is a strong predictor of mortality in peritoneal dialysis patients - independently of cardiac failure. PLoS One 2016;11(7):e0158741 [FREE Full text] [doi: 10.1371/journal.pone.0158741] [Medline: 27415758]

5. Guo Q, Lin J, Li J, Yi C, Mao H, Yang X, et al. The effect of fluid overload on clinical outcome in southern Chinese patients undergoing continuous ambulatory peritoneal dialysis. Perit Dial Int 2015 Dec;35(7):691-702 [FREE Full text] [doi: 10.3747/pdi.2014.00008] [Medline: 26152580]
6. Dao Bui Quy Q, Pham Ngoc Huy T, Nguyen Duc L, Pham Van M, Nguyen Huu D, Nguyen Duy T, et al. Overhydration and low serum prealbumin predict peritoneal dialysis-related peritonitis in continuous ambulatory peritoneal dialysis patients. BMC Nephrol 2020 Nov 25;21(1):512 [FREE Full text] [doi: 10.1186/s12882-020-02178-w] [Medline: 33238904]
7. Shu Y, Liu J, Zeng X, Hong HG, Li Y, Zhong H, et al. The effect of overhydration on mortality and technique failure among peritoneal dialysis patients: a systematic review and meta-analysis. Blood Purif 2018;46(4):350-358. [doi: 10.1159/000492148] [Medline: 30189422]
8. Ng JK, Kwan BC, Chow K, Pang W, Cheng PM, Leung C, et al. Asymptomatic fluid overload predicts survival and cardiovascular event in incident Chinese peritoneal dialysis patients. PLoS One 2018;13(8):e0202203 [FREE Full text] [doi: 10.1371/journal.pone.0202203] [Medline: 30102739]
9. Hong YA, Yoon HE, Choi BS, Shin SJ, Kim Y, Lee SY, et al. The effect of strict volume control assessed by repeated bioimpedance spectroscopy on cardiac function in peritoneal dialysis patients. Sci Rep 2019 Nov 27;9(1):17679 [FREE Full text] [doi: 10.1038/s41598-019-53792-0] [Medline: 31776362]
10. Kim Y, Biesen WV. Fluid overload in peritoneal dialysis patients. Semin Nephrol 2017 Jan;37(1):43-53. [doi: 10.1016/j.semnephrol.2016.10.006] [Medline: 28153194]
11. Kuehn BM. Is there an app to solve app overload? JAMA 2015 Apr 14;313(14):1405-1407. [doi: 10.1001/jama.2015.2381] [Medline: 25871657]
12. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
13. 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]
14. L'Engle K, Raney L, D'Adamo M. mHealth resources to strengthen health programs. Glob Health Sci Pract 2014 Feb;2(1):130-131 [FREE Full text] [doi: 10.9745/GHSP-D-14-00013] [Medline: 25276568]
15. Labrique AB, Vasudevan L, Kochi E, Fabricant R, Mehl G. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. Glob Health Sci Pract 2013 Aug;1(2):160-171 [FREE Full text] [doi: 10.9745/GHSP-D-13-00031] [Medline: 25276529]
16. Hoque MR, Rahman MS, Nipa NJ, Hasan MR. Mobile health interventions in developing countries: a systematic review. Health Informatics J 2020 Dec;26(4):2792-2810 [FREE Full text] [doi: 10.1177/1460458220937102] [Medline: 32691659]
17. Roess A. The promise, growth, and reality of mobile health - another data-free zone. N Engl J Med 2017 Nov 23;377(21):2010-2011. [doi: 10.1056/NEJMp1713180] [Medline: 29116869]
18. Arie S. Can mobile phones transform healthcare in low and middle income countries? BMJ 2015 Apr 22;350:h1975. [doi: 10.1136/bmj.h1975] [Medline: 25902967]
19. 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]
20. DAMASAC Data Management and Statistical Analysis Center. Faculty of Public Health, Khon Kaen University. URL: http://www.thaicarecloud.org/ [accessed 2022-02-10]
21. Coskun V, Ozdenizci B, Ok K. The survey on near field communication. Sensors (Basel) 2015 Jun 05;15(6):13348-13405 [FREE Full text] [doi: 10.3390/s150613348] [Medline: 26057043]
22. Morak J, Schreier G. Design and evaluation of Near Field Communication (NFC) technology based solutions for mHealth challenges. In: Mobile Health. Cham: Springer; 2015.
23. Morak J, Kumpusch H, Hayn D, Modre-Osprian R, Schreier G. Design and evaluation of a telemonitoring concept based on NFC-enabled mobile phones and sensor devices. IEEE Trans Inf Technol Biomed 2012 Jan;16(1):17-23. [doi: 10.1109/TITB.2011.2176498] [Medline: 22113811]
24. UA-767NFC Blood Pressure Monitor with NFC devices. A&D Medical. URL: https://www.andmedical.com.au/product-service/telemedicine-bluetooth-and-nfc-products/ua-767nfc-blood-pressure-monitor-with-nfc-devices [accessed 2022-02-10]
25. Anabtawi N, Freeman S, Ferzli R. A fully implantable, NFC enabled, continuous interstitial glucose monitor. IEEE EMBS Int Conf Biomed Health Inform 2016 Feb;2016:612-615 [FREE Full text] [doi: 10.1109/BHI.2016.7455973] [Medline: 28702512]
26. Kijsanayotin B, Pannarunothai S, Speedie SM. Factors influencing health information technology adoption in Thailand's community health centers: applying the UTAUT model. Int J Med Inform 2009 Jun;78(6):404-416. [doi: 10.1016/j.ijmedinf.2008.12.005] [Medline: 19196548]
27. Jaspers M, Steen T, Bos C, Geenen M. The think aloud method: a guide to user interface design. Intl J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]

28. Rajan JV, Moura J, Gourley G, Kiso K, Sizilio A, Cortez AM, et al. Understanding the barriers to successful adoption and use of a mobile health information system in a community health center in São Paulo, Brazil: a cohort study. BMC Med Inform Decis Mak 2016 Nov 17;16(1):146 [FREE Full text] [doi: 10.1186/s12911-016-0385-1] [Medline: 27855685]
29. Cornet VP, Toscos T, Bolchini D, Rohani Ghahari R, Ahmed R, Daley C, et al. Untold stories in user-centered design of mobile health: practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e17703 [FREE Full text] [doi: 10.2196/17703] [Medline: 32706745]
30. Nielson J, Landauer TK. A mathematical model of the finding of usability problems. In: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems. 1993 Presented at: INTERCHI93: Conference on Human Factors in Computing; Apr 24 - 29, 1993; Amsterdam The Netherlands. [doi: 10.1145/169059.169166]
31. Boontarig W, Chutimaskul W, Chongsuphajaisiddhi V, Papasratorn B. Factors influencing the Thai elderly intention to use smartphone for e-Health services. In: Proceedings of the 2012 IEEE Symposium on Humanities, Science and Engineering Research. 2012 Presented at: 2012 IEEE Symposium on Humanities, Science and Engineering Research; Jun 24-27, 2012; Kuala Lumpur, Malaysia. [doi: 10.1109/shuser.2012.6268881]
32. Nunes A, Limpo T, Castro SL. Acceptance of mobile health applications: examining key determinants and moderators. Front Psychol 2019;10:2791 [FREE Full text] [doi: 10.3389/fpsyg.2019.02791] [Medline: 31920836]
33. Dwivedi Y, Shareef M, Simintiras A, Lal B, Weerakkody V. A generalised adoption model for services: a cross-country comparison of mobile health (m-health). Government Inf Q 2016 Jan;33(1):174-187 [FREE Full text] [doi: 10.1016/j.giq.2015.06.003]
34. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
35. Wallis L, Blessing P, Dalwai M, Shin SD. Integrating mHealth at point of care in low- and middle-income settings: the system perspective. Glob Health Action 2017 Jun;10(sup3):1327686 [FREE Full text] [doi: 10.1080/16549716.2017.1327686] [Medline: 28838302]
36. Woulfe F, Fadahunsi KP, Smith S, Chirambo GB, Larsson E, Henn P, et al. Identification and evaluation of methodologies to assess the quality of mobile health apps in high-, low-, and middle-income countries: rapid review. JMIR Mhealth Uhealth 2021 Oct 12;9(10):e28384 [FREE Full text] [doi: 10.2196/28384] [Medline: 34636737]
37. Feroz A, Kadir MM, Saleem S. Health systems readiness for adopting mhealth interventions for addressing non-communicable diseases in low- and middle-income countries: a current debate. Glob Health Action 2018;11(1):1496887 [FREE Full text] [doi: 10.1080/16549716.2018.1496887] [Medline: 30040605]

Abbreviations

APD: automatic peritoneal dialysis
CAPD: continuous ambulatory peritoneal dialysis
CKD: chronic kidney disease
CKDNET: Chronic Kidney Disease Prevention in the Northeast of Thailand
CKD-PD: chronic kidney disease–peritoneal dialysis
mHealth: mobile health
NFC: near-field communication
OCR: optical character recognition
PD: peritoneal dialysis
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Mavragani; submitted 01.03.22; peer-reviewed by K Paik, R Marshall; comments to author 08.05.22; revised version received 26.05.22; accepted 29.05.22; published 06.07.22.

Please cite as:
Lukkanalikitkul E, Kongpetch S, Chotmongkol W, Morley MG, Anutrakulchai S, Srichan C, Thinkhamrop B, Chunghom T, Wiangnon P, Thinkhamrop W, Morley KE
Optimization of the Chronic Kidney Disease–Peritoneal Dialysis App to Improve Care for Patients on Peritoneal Dialysis in Northeast Thailand: User-Centered Design Study
JMIR Form Res 2022;6(7):e37291
URL: https://formative.jmir.org/2022/7/e37291
doi:10.2196/37291
PMID:35793137

Original Paper

Experimental Implementation of NSER Mobile App for Efficient Real-Time Sharing of Prehospital Patient Information With Emergency Departments: Interrupted Time-Series Analysis

Kiyomitsu Fukaguchi[1*], MD; Tadahiro Goto[2*], MD, MPH, PhD; Tadatsugu Yamamoto[1*], MD; Hiroshi Yamagami[1*], MD

[1]Department of Emergency Medicine, Shonan Kamakura General Hospital, Kamakura-shi, Kanagawa, Japan

[2]TXP Research, TXP Medical Co Ltd, Tokyo, Japan

[*]all authors contributed equally

Corresponding Author:
Kiyomitsu Fukaguchi, MD
Department of Emergency Medicine
Shonan Kamakura General Hospital
1370-1
Okamoto
Kamakura-shi, Kanagawa, 247-0072
Japan
Phone: 81 467 46 1717
Fax: 81 467 45 0190
Email: fukaskgh@gmail.com

Abstract

Background: With the aging society, the number of emergency transportations has been growing. Although it is important that a patient be immediately transported to an appropriate hospital for proper management, accurate diagnosis in the prehospital setting is challenging. However, at present, patient information is mainly communicated by telephone, which has a potential risk of communication errors such as mishearing. Sharing correct and detailed prehospital information with emergency departments (EDs) should facilitate optimal patient care and resource use. Therefore, the implementation of an app that provides on-site, real-time information to emergency physicians could be useful for early preparation, intervention, and effective use of medical and human resources.

Objective: In this paper, we aimed to examine whether the implementation of a mobile app for emergency medical service (EMS) would improve patient outcomes and reduce transportation time as well as communication time by phone (ie, phone-communication time).

Methods: We performed an interrupted time-series analysis (ITSA) on the data from a tertiary care hospital in Japan from July 2021 to October 2021 (8 weeks before and 8 weeks after the implementation period). We included all patients transported by EMS. Using the mobile app, EMS can send information on patient demographics, vital signs, medications, and photos of the scene to the ED. The outcome measure was inpatient mortality and transportation time, as well as phone-communication time, which was the time for EMS to negotiate with ED staffs for transport requests.

Results: During the study period, 1966 emergency transportations were made (n=1033, 53% patients during the preimplementation period and n=933, 47% patients after the implementation period). The ITSA did not reveal a significant decrease in patient mortality and transportation time before and after the implementation. However, the ITSA revealed a significant decrease in mean phone-communication time between pre- and postimplementation periods (from 216 to 171 seconds; −45 seconds; 95% CI −71 to −18 seconds). From the pre- to postimplementation period, the mean transportation time from EMS request to ED arrival decreased by 0.29 minutes (from 36.1 minutes to 35.9 minutes; 95% CI −2.20 to 1.60 minutes), without change in time trends. We also introduced cases where the app allowed EMS to share accurate and detailed prehospital information with the emergency department, resulting in timely intervention and reducing the burden on the ED.

Conclusions: The implementation of a mobile app for EMS was associated with reduced phone-communication time by 45 seconds (22%) without increasing mortality or overall transportation time despite the implementation of new methods in the real

clinical setting. In addition, real-time patient information sharing, such as the transfer of monitor images and photos of the accident site, could facilitate optimal patient care and resource use.

(JMIR Form Res 2022;6(7):e37301) doi:10.2196/37301

KEYWORDS

emergency department; emergency medical services; mobile apps; interrupted time series analysis; emergency; patient record; implementation; patient care; app; implement; medical informatics; clinical informatics; decision support; electronic health record; eHealth; digital health

Introduction

With the aging society, the number of emergency transportation has been growing [1]. Researchers used a nationwide database in Japan and reported that the annual emergency transportations increased from 4 million in 2000 to 6 million in 2020 [1]. The substantial increase in emergency transportation causes depletion of medical resources and emergency department (ED) overcrowding, resulting in poor patient outcomes [2,3]. Indeed, in Japan, approximately 10,000 emergency patients were turned away by at least 6 hospitals annually [4].

When a patient has stroke, acute myocardial infarction, or severe trauma, it is important to transport them to an appropriate hospital in order to provide appropriate management immediately [5-9]. Although accurate diagnosis in the prehospital setting is challenging, real-time sharing of detailed patient information should facilitate appropriate transportation and management [10,11]. However, at present, patient information is mainly communicated by telephone, which has the potential risk of communication errors such as mishearing. In addition, photos of the patient, patient monitors (eg, electrocardiogram), and the accident scene can contain important information that emergency physicians would want to know in advance. Therefore, the implementation of an app that provides on-site, real-time information to emergency physicians could be useful for early preparation, intervention, and effective use of medical and human resources [5-8,10-12].

Thus, in this study, we examined whether the implementation of a mobile app for emergency medical services (EMS) could improve patient outcomes. We hypothesized that patient mortality would be improved by shortening the transportation time and time for EMS to negotiate with ED staffs for transport requests.

Methods

Study Design and Settings

We performed an interrupted time-series analysis to examine the change in outcomes before and after the introduction of a mobile app for EMS. This study was conducted using data on EMS transportations by Kamakura City Fire Department to Shonan Kamakura General Hospital, a tertiary care hospital in Japan, from July 8, 2021, to October 27, 2021. Kamakura City is a city in Kanagawa Prefecture, Japan, with a population of 172,948 people and an average age of 48.8 years as of 2020. The city has a high ratio of the aging population, with 53,517 (31.1%) people aged ≥65 years. Kamakura City has 8 fire departments that make 29.9 emergency transportations per day and 10,896 per year [13]. The Emergency Department of Shonan Kamakura General Hospital had 43,506 ED visits by patients in 2020. Among these patients, 14,925 (34.3%) were transported to the emergency department by EMS, which was the highest annual number of emergency transportations of all hospitals in Japan that year. All patients transported by EMS were accepted [14]. The number of patients transported by the Kamakura City EMS during the study period (16 weeks) was estimated to be approximately 2000 based on the number of patients transported in the past.

Ethics Approval

This study protocol was approved by the Ethics Committee of Shonan Kamakura General Hospital (approval number: TGE01663-024; approval date: February 25, 2021). The need for informed consent was waived due to the retrospective nature of the study. For the publication of images and personal information, the patients' consent was obtained, and their consent was recorded in their medical charts. The institutional review board approved the use of patient information for this study.

Apps Used and Their Features

We implemented an app for EMS named NEXT Stage ER mobile (NSER mobile) on September 2, 2021. The NSER mobile is an app that was developed to reduce the complexity of information entry for EMS teams and to increase the efficiency of information sharing with EDs. The app is equipped with high-precision voice input and optical character recognition functions that enable timely information sharing by sending information to hospitals in the original text (eg, medications and past medical history) and in coded form (eg, International Statistical Classification of Diseases and Related Health Problems, 10th revision codes). We provide a description of the usage process of the NSER mobile by EMS and hospitals in Multimedia Appendices 1 and 2. All EMS staffs took 3-hour lectures to learn to use the app before implementation.

The developer of NSER mobile is TXP Medical Co Ltd. A free trial for this product is available through the website of TXP Medical [15]. The system's accuracy and validity are now being examined in a multicenter study (UMIN-CTR ID: UMIN000045775), but it is fundamentally the same as the NEXT Stage ER system, which is a well-validated ED information system [16]. In addition, in this study, we did not use table data from the app. All data are abstracted from the Kamakura City Fire Department and patient registry of ED patients of Shonan Kamakura General Hospital. Thus, the only data we extracted from the NEXT Stage ER mobile system for

this study are images (though more detailed data exist in the cloud server of the system).

Study Population

All patients transported to Shonan Kamakura General Hospital by the Kamakura City EMS during the study period were included in the study. Exclusion criteria were cases with missing data on call times between the hospital and EMS and EMS activity times (26/1992, 1.3%).

Outcome Measures

The primary outcome was inpatient mortality. Secondary outcomes were the overall transportation time (the time from the patient's call for an ambulance to arrival at the hospital), and phone-communication time (the time for EMS to negotiate with ED staffs for transport requests).

Statistical Analysis

Cases with missing data on the time spent in EMS and the emergency department (ED) were excluded from the analysis (26/1992, 1.3%). Using data that met the inclusion criteria, we performed an interrupted time-series analysis with a linear regression model to examine whether the implementation of the mobile app had an impact on the outcomes. An interrupted time-series analysis is a quasi-experimental design for evaluating the effectiveness of population-level health interventions implemented at a clearly defined point in time and is thus widely used to evaluate the effectiveness of interventions [17-19]. The time unit was weeks (ie, 8 time points before and 8 time points after implementation).

Additionally, we described specific cases where the immediate sharing of patient information and photos with NSER mobile led to rapid diagnosis and intervention.

All statistical analyses were conducted using R version 4.1.1 (R Foundation for Statistical Computing).

Results

Among the 1992 patients transported by EMS, we excluded 26 patients (1.3%) with missing data on the transport time or the phone communication time, and the remaining 1966 (98.7%) patients were eligible for this study. The mean age was 66.7 years (SD 25.4), and 49% (n=963) were men. Of the 1966 patients, 1033 (53%) were transported during the preimplementation period, and 933 (47%) were transported in the postimplementation period. The patient characteristics in the pre- and postimplementation periods were similar. Additionally, there were no significant differences in the region of transportation and the EMS team numbers (Table 1).

Table 1. Characteristics of patients transported by emergency medical services.

Patient characteristics and variables	Patients transported using the app (n=1033)	Patients transported in the usual way (n=933)	*P* value
Age (years), mean (SD)	65.8 (27.2)	68.1 (23.5)	.13
Age profile (years), n (%)			
≤18	98 (9.5)	48 (5.1)	
18-64	307 (29.7)	232 (24.9)	
65-84	324 (31.4)	343 (36.8)	
≥85	304 (29.4)	310 (33.2)	
Sex, n (%)			.80
Male	508 (49.2)	464 (49.7)	
Female	525 (50.8)	469 (50.3)	
Number of calls to the hospital, mean (SD)	1.07 (0.28)	1.05 (0.25)	.05
Region of emergency medical services, n (%)			.17
Kamakura	162 (15.7)	147 (15.8)	
Ofuna	176 (17.0)	181 (19.4)	
Fukasawa	203 (19.7)	155 (16.6)	
Tamanawa	139 (13.4)	128 (13.7)	
Koshigoe	121 (11.7)	109 (11.7)	
Imaizumi	74 (7.2)	82 (8.8)	
Shitirigahama	91 (8.8)	60 (6.4)	
Zyoumyouzi	67 (6.5)	71 (7.6)	
Level of consciousness (JCS[a]), n (%)			.70
0	524 (50.7)	443 (47.5)	
1	163 (15.8)	178 (19.1)	
2	143 (13.8)	123 (13.2)	
3	90 (8.7)	79 (8.5)	
10	47 (4.5)	45 (4.8)	
20	14 (1.4)	15 (1.6)	
30	5 (0.5)	4 (0.4)	
100	7 (0.7)	8 (0.9)	
200	7 (0.7)	4 (0.4)	
300	33 (2.9)	34 (3.6)	
Vital signs, mean (SD)			
Systolic blood pressure (mmHg)	133 (45.9)	135 (50.6)	.35
Diastolic blood pressure (mmHg)	79.4 (27.8)	79.6 (30.7)	.91
Pulse rate (per min)	89.1 (29.3)	84.5 (27.2)	<.001
Respiratory rate (per min)	21.2 (6.33)	20.8 (5.67)	.08
Saturation (%)	89.2 (24.5)	87.9 (26.8)	.25
Body temperature (°C)	35.0 (8.43)	34.4 (9.04)	.16
Classification of diseases[b], n (%)			<.001
Endogenous disease	748 (72.4)	680 (72.9)	
Trauma	270 (26.1)	234 (25.1)	
Cardiac arrest	15 (1.5)	19 (2.0)	

Patient characteristics and variables	Patients transported using the app (n=1033)	Patients transported in the usual way (n=933)	*P* value
Severity of illness at the ED[c] visit[b], n (%)			.70
Minor illness	375 (36.3)	351 (37.6)	
Moderate illness	567 (54.9)	492 (52.7)	
Serious illness	76 (7.3)	71 (7.6)	
Death	15 (1.5)	19 (2.0)	
Disposition at the ED, n (%)			.04
Discharge	657 (63.6)	549 (58.8)	
Admission	308 (29.8)	295 (31.6)	
Transfer to another hospital for admission	50 (4.8)	68 (7.3)	
Death	18 (1.7)	21 (2.3)	
Prognosis, n (%)			.56
In-hospital death	53 (5.1)	46 (4.9)	

[a]JCS: Japan Coma Scale.

[b]The severity of illness at the emergency department is classified as follows: minor illness—patient can return home after treatment; moderate illness—patient requires inpatient treatment, but the disease severity is low and can be managed in a general ward; serious illness—multiorgan failures such as respiratory or circulatory failure requiring monitoring, a ventilator, vasopressors such as catecholamines, and admission to an intensive care unit; death—cardiac arrest on arrival at the hospital.

[c]ED: emergency department.

Inpatient Mortality

Of the 1966 eligible patients, 53 (5.1%) died in hospital in the preimplementation period and 46 (4.9%) died in hospital in the postimplementation period. From the pre- to postimplementation period, the proportions of in-hospital deaths among patients who were transported to EDs during each period decreased by 5% (95% CI −11% to 1%), followed by a decreasing trend relative to preimplementation of −1% per week (95% CI −2% to 1%). There was no significant change in inpatient mortality. Figure 1 shows the proportions of in-hospital deaths among patients who were transported to emergency departments during each period. On September 2, 2021, EMS began transportation using the new app in place of the traditional method. The proportion of in-hospital deaths among the transported patients is plotted for 8-week periods before and after the implementation. For in-hospital mortality, the R^2 for the preimplementation model was 0.26, while the R^2 for the postimplementation model was 0.23. The results of interrupted time series analysis on inpatient mortality are shown in Table 2.

Figure 1. Interrupted time-series analysis of inpatient mortality.

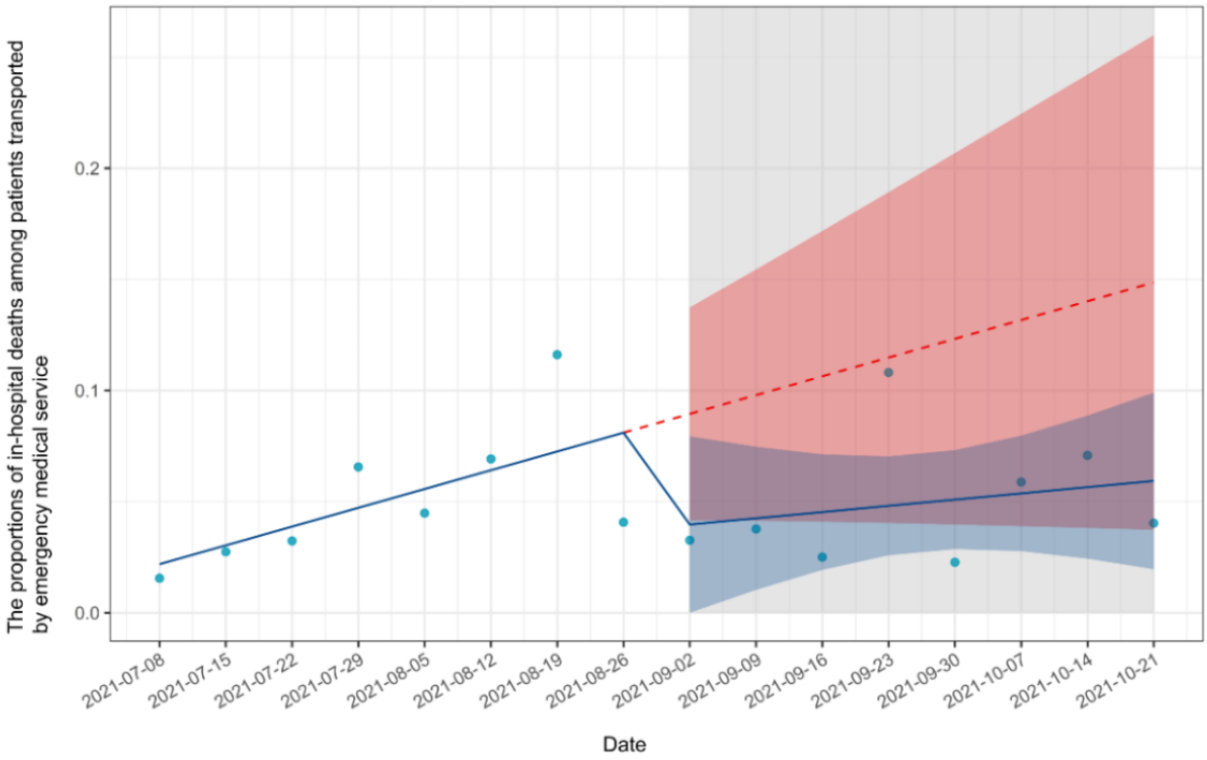


Table 2. Results of interrupted time-series analysis on inpatient mortality.

Time-series analysis	Estimate (95% CI)	*P* value
Trends in inpatient mortality before implementation	0.01 (0.00 to 0.02)	.07
Absolute change in the inpatient mortality before and after implementation	−0.05 (−0.11 to 0.01)	.11
Trends in inpatient mortality after implementation	0.00 (−0.01 to 0.01)	.50
Change in slope before and after implementation	−0.01 (−0.02 to 0.01)	.40

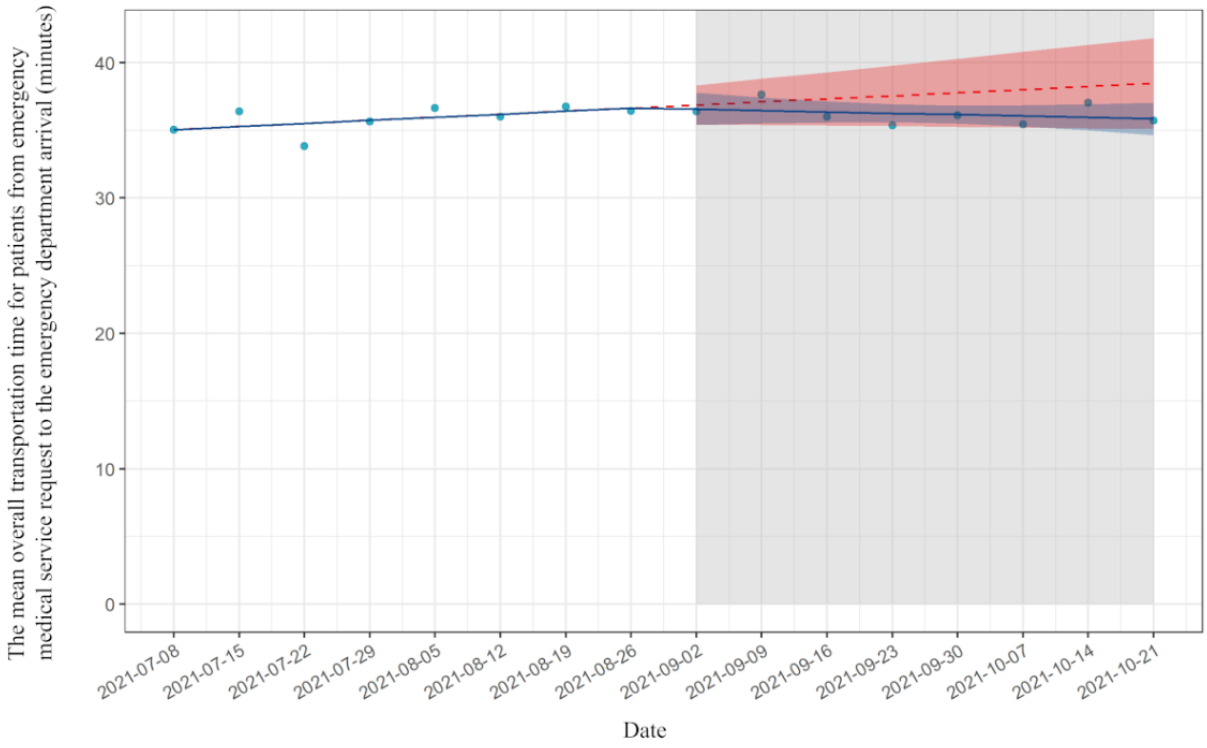
The Transportation Time From EMS Request to ED Arrival

The mean transportation time from EMS request to ED arrival was 35.9 minutes (SD 9.7 minutes) in the preimplementation period and 36.1 minutes (SD 8.5 minutes) in the postimplementation period. From the pre- to postimplementation period, the mean transportation time from EMS request to ED arrival decreased by 0.29 minutes (95% CI −2.20 to 1.60 minutes), followed by a decreasing trend relative to preimplementation of −0.33 minutes per week (95% CI −0.74 to 0.07; Table 3). Figure 2 shows the mean transportation time from the emergency call to arrival at the hospital. On September 2, 2021, EMS began transporting using the new app in place of the traditional method. The mean transportation time is plotted for 8-week periods before and after the implementation. For transportation time, the R^2 for the preimplementation model was 0.30, and the R^2 for the postimplementation model was 0.28.

Table 3. Results of interrupted time-series analysis on transportation time from emergency medical services to emergency department arrival.

Time-series analysis	Estimate (min), 95% CI	*P* value
Trends in mean transportation time before implementation	0.23 (−0.06 to 0.51)	.11
Absolute change in the transportation time before and after implementation	−0.29 (−2.20 to 1.60)	.70
Trends in mean transportation time after implementation	−0.10 (−0.39 to 0.18)	.40
Change in slope before and after implementation	−0.33 (−0.74 to 0.07)	.10

Figure 2. Interrupted time-series analysis on transportation time from emergency medical services to emergency department arrival.



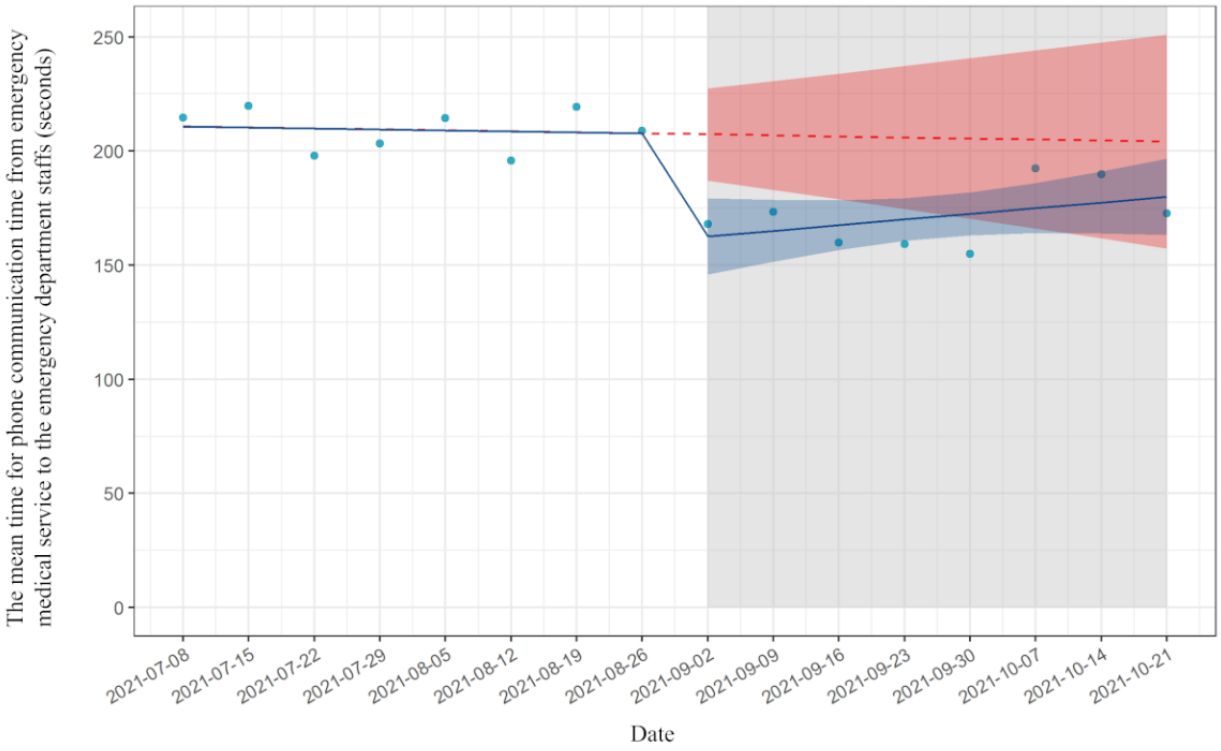
Phone-Communication Time Between EMS Teams and Hospital

The mean time of phone communication between EMS and ED staffs was 216 (SD 107) seconds in the preimplementation period and 171 (SD 120) seconds in the postimplementation period. From the pre- to postimplementation period, the phone-communication time decreased by 45 seconds (95% CI −71 to −18 seconds), followed by an increasing trend relative to preimplementation of +2.9 seconds per week (95% CI −2.7 to 8.6; Table 4). The mean phone-communication time between EMS and the hospital is shown in Figure 3. On September 2, 2021, the EMS started using the app for transportation in place of the conventional method. The mean phone-call time is plotted for 8-week periods before and after the implementation. For phone-communication time, the R^2 for the preimplementation model was 0.78, while the R^2 for the postimplementation model was 0.72.

Table 4. Results of interrupted time-series analysis on phone-communication time.

Time-series analysis	Estimate (s), 95% CI	*P* value
Trends in mean phone-communication time before implementation	−0.44 (−4.4 to 3.6)	.80
Absolute change in the phone-communication time before and after implementation	−45.0 (−71.0 to −18.4)	.003
Trends in mean phone-communication time after implementation	2.5 (−1.5 to 6.5)	.20
Change in slope before and after implementation	2.9 (−2.7 to 8.6)	.30

Figure 3. Interrupted time-series analysis on phone-communication time.



Specific Cases

We experienced several cases where the immediate sharing of patient information and photos using the NSER mobile has led to rapid diagnosis and intervention. For example, a 47-year-old man with a history of diabetes mellitus, who had visited the hospital the day before for chest pain, called for EMS again complaining of persistent chest pain. Upon EMS's arrival, his electrocardiography monitor showed ST-segment elevation in lead II (Multimedia Appendix 3). Before arriving at the hospital, he was diagnosed with ST-segment elevation acute myocardial infarction. Upon consulting with the cardiologist, he was able to start percutaneous coronary intervention within ten minutes after the patient's arrival at the ED. Multimedia Appendix 3 shows the monitor screen in the ambulance transporting a 47-year-old man who called EMS complaining of chest pain. Lead II showed ST-segment elevation, and the images were transmitted from EMS to the ED staff, leading to early diagnosis of acute myocardial infarction and percutaneous coronary intervention within 10 minutes of arrival at the hospital.

In another example, a 35-year-old man was riding a bicycle when he collided with a car traveling at 60 kilometers per hour. He had no memory when he was injured. While the mechanism of injury suggests a highly serious one, real-time information sharing (initial assessment of the patient, vital signs, photos of the injury scene, and damage to the bicycle) allowed us to determine that the injury was minor, with only a contusion on the left lower leg and the scalp. Consequently, we were able to reduce unnecessary preparation for the initial treatment, including surgery for damage control (Multimedia Appendices 4-6).

Discussion

Principal Findings

In this ITSA study, there was no substantial decrease in in-hospital mortality and overall transportation time between pre- and postimplementation periods. On the other hand, the implementation reduced phone-communication time by 22% without increasing mortality or overall transportation time despite the implementation of new methods in the real clinical setting.

The conventional communication between EMS and ED staff by telephone only poses substantial stress on EMS staff, exemplified by difficulty hearing and misidentification of information. Data related to patient demographics, vital signs, past medical history, and medications are likely misclassified due to such technical difficulties, and real-time information sharing could reduce such miscommunication. In addition, the app's feature to share visual information via optical character recognition in a timely manner was useful to ensure that ED staffs are fully prepared to receive patients. For ED staff, timely monitoring and understanding of the situation offered by the EMS were useful for maintaining a high quality of clinical practice.

Comparison With Prior Work

Appropriate use of medical apps could lead to a seamless transition of management from prehospital to post-ED arrival. As reported in this study, apps can be used to obtain information such as medical history and prehospital electrocardiograms in advance. From such information, physicians can prepare for urgent interventions (eg, catheterization) before the patient arrives at the hospital. There have been several reports on the

usefulness of apps that provide prehospital information in emergency medicine. A system called ORION (Osaka Emergency Information Research Intelligent Operation Network system), introduced in Osaka City in 2013, reduced the number of cases that are difficult to transport [12]. Another study reported that the communication-type medical apps can be accurately used remotely, and information can be shared with the stroke team to prepare for rapid treatment [5,20,21].

The NSER mobile app is a digitalization tool for EMS in the clinical setting, and there are no patients for whom the app cannot be used. Nevertheless, we think that the system is more suitable for patients who need emergency interventions (eg, cardiac catheterization) [2-5,7] rather than those with cardiac arrest (EMS may not have enough time to use the app). Due to the limited sample size, we could not analyze data after stratifying by these variables. Thus, we are conducting further study in different settings to examine the effectiveness of the app.

Strengths and Limitations

This study has several strengths, a few of which are as follows. First, there were no similar studies on prehospital information transfer apps aimed at improving the efficiency of emergency patient transport without assuming a specific disease. Second, the interrupted time-series analysis estimates the effect of intervention on a population and is a study design without a control group. Third, the advantage of reducing phone-communication time through real-time information sharing is noticeable especially when the EMS is consulting multiple hospitals to accept patients at the same time, not to mention that in many cases it is difficult to transport patients, particularly in urban areas in Japan. Indeed, according to statistics from the Ministry of Health, Labor, and Welfare in 2016, even for critically ill patients, there were 10,039 cases (2.3%) in which the number of consultations to medical institutions was ≥4 times and 22,104 cases (5.0%) in which the time spent on site was 30 minutes or longer [22]. In such cases, given the tough negotiation with hospitals, the reduction of phone-communication time while efficiently sharing prehospital information should reduce the burden on EMS. We believe that the findings from this study allow us to consider the substantial contribution and potential benefits of mobile apps to emergency medical care.

Our study has several limitations. First, although we performed a 3-hour lecture for using the app prior to implementation, users may not have been able to get accustomed to the app quickly enough in the clinical setting. Despite this, there was clear improvement in phone-communication time immediately after its implementation. Second, there was no control group in our ITSA design [23]. Nonetheless, our findings are likely robust, given that there were no interventions other than the implementation of the app that may have affected the outcome. While the COVID-19 pandemic may have affected the assumptions of the interrupted time series analysis, the implementation of personal protective equipment for EMS was initiated on March 2, 2020. Therefore, change in practice due to the pandemic may not have substantially affected the EMS during the study period (July 8, 2021, to October 27, 2021). Third, we did not have information on the time to intervention at the ED (eg, time to urgent catheterization) and ED overcrowding. Therefore, further studies are needed to examine the impact of the app on clinical practice. Fourth, in this study, we only evaluated the observed values for 8 weeks before and after the intervention. A recent simulation-based study on ITSA reported that 12 preintervention and 12 postintervention time points may be required for a moderate intervention effect sizes [24]. Lastly, there is limited generalizability of our findings because our study was a single-center, retrospective observational study in Japan with a small sample size. In addition, EMS systems are different across countries. The extrapolation of our findings to other settings should be done with caution, and therefore additional large-scale studies are warranted.

Future Directions

As shown in the 2 cases, the implementation of a mobile app for efficient real-time sharing of prehospital patient information has potential to reduce the time to intervention, resulting in better patient outcomes. In addition, in Japan, especially in the urban areas, there is the difficulty in determining the hospital for emergency patient [1]; however, for instance, the average number of hospitals that EMS phoned to transport patient was 1 in Kamakura city during this study period, so the decision to transport a patient did not take extra time. Therefore, the app may reduce the overall transportation time by decreasing the number of calling from EMS to hospitals.

Conclusions

The implementation of a mobile app for EMS reduced phone-communication time by 22% without increasing mortality or overall transportation time despite the implementation of new methods in the real clinical setting. Real-time patient information sharing, such as the transfer of monitor images and photos of the accident site, could facilitate optimal patient care and resource use.

Acknowledgments

This study was supported by the grant “2021 Survey and Research Project on First Aid” from Japan Emergency Medical Foundation. The study is still in progress, and the results presented in this study are preliminary.

Conflicts of Interest

TG is the Chief Scientific Officer of TXP Medical Co Ltd.

Multimedia Appendix 1

Workflow of an emergency medical services team using NEXT Stage ER mobile.
[PNG File , 303 KB - formative_v6i7e37301_app1.png]

Multimedia Appendix 2
An example of sharing trauma images, monitors, and clinical information from emergency medical services to the emergency department staff.
[PNG File , 355 KB - formative_v6i7e37301_app2.png]

Multimedia Appendix 3
Monitor screen image of a case of a 47-year-old man complaining of chest pain.
[PNG File , 1063 KB - formative_v6i7e37301_app3.png]

Multimedia Appendix 4
The damaged bicycle image of a case of a 35-year-old man with trauma.
[PNG File , 1410 KB - formative_v6i7e37301_app4.png]

Multimedia Appendix 5
Left lower extremity injuries of a case of a 35-year-old man with trauma.
[PNG File , 1435 KB - formative_v6i7e37301_app5.png]

Multimedia Appendix 6
Monitor screen image of a case of a 35-year-old man with trauma.
[PNG File , 1257 KB - formative_v6i7e37301_app6.png]

References

1. Summary of the 2020 edition of emergency and rescue activities. Ministry of Internal Affairs and Communications. URL: https://www.fdma.go.jp/pressrelease/houdou/items/sokuhouti.pdf [accessed 2022-05-03]
2. Derlet RW, Richards JR. Overcrowding in the nation's emergency departments: complex causes and disturbing effects. Ann Emerg Med 2000 Jan;35(1):63-68. [doi: 10.1016/s0196-0644(00)70105-3] [Medline: 10613941]
3. Di Somma S, Paladino L, Vaughan L, Lalle I, Magrini L, Magnanti M. Overcrowding in emergency department: an international issue. Intern Emerg Med 2015 Mar;10(2):171-175. [doi: 10.1007/s11739-014-1154-8] [Medline: 25446540]
4. Results of the survey on the status of medical institutions' acceptance of emergency transport during 2018. Fire and Disaster Management Agency. URL: https://www.fdma.go.jp/singi_kento/kento/items/post-48/03/sankou3.pdf [accessed 2022-05-03]
5. Wu Y, Chen F, Song H, Feng W, Sun J, Liu R, et al. Use of a smartphone platform to help with emergency management of acute ischemic stroke: observational study. JMIR Mhealth Uhealth 2021 Feb 09;9(2):e25488 [FREE Full text] [doi: 10.2196/25488] [Medline: 33560236]
6. Nogueira RG, Silva GS, Lima FO, Yeh Y, Fleming C, Branco D, et al. The FAST-ED App: a smartphone platform for the field triage of patients with stroke. Stroke 2017 May;48(5):1278-1284. [doi: 10.1161/STROKEAHA.116.016026] [Medline: 28411260]
7. Ortolani P, Marzocchi A, Marrozzini C, Palmerini T, Saia F, Baldazzi F, et al. Usefulness of prehospital triage in patients with cardiogenic shock complicating ST-elevation myocardial infarction treated with primary percutaneous coronary intervention. Am J Cardiol 2007 Sep 01;100(5):787-792. [doi: 10.1016/j.amjcard.2007.03.099] [Medline: 17719321]
8. Eder PA, Reime B, Wurmb T, Kippnich U, Shammas L, Rashid A. Prehospital telemedical emergency management of severely injured trauma patients. Methods Inf Med 2018 Nov;57(5-06):231-242. [doi: 10.1055/s-0039-1681089] [Medline: 30875702]
9. Synnot A, Karlsson A, Brichko L, Chee M, Fitzgerald M, Misra MC, Australia-India Trauma System Collaboration. Prehospital notification for major trauma patients requiring emergency hospital transport: A systematic review. J Evid Based Med 2017 Aug;10(3):212-221. [doi: 10.1111/jebm.12256] [Medline: 28467026]
10. Felzen M, Brokmann JC, Beckers SK, Czaplik M, Hirsch F, Tamm M, et al. Improved technical performance of a multifunctional prehospital telemedicine system between the research phase and the routine use phase - an observational study. J Telemed Telecare 2017 Apr;23(3):402-409. [doi: 10.1177/1357633X16644115] [Medline: 27080747]
11. Flynn D, Francis R, Robalino S, Lally J, Snooks H, Rodgers H, et al. A review of enhanced paramedic roles during and after hospital handover of stroke, myocardial infarction and trauma patients. BMC Emerg Med 2017 Feb 23;17(1):5 [FREE Full text] [doi: 10.1186/s12873-017-0118-5] [Medline: 28228127]
12. Katayama Y, Kitamura T, Kiyohara K, Iwami T, Kawamura T, Izawa J, et al. Improvements in patient acceptance by hospitals following the introduction of a smartphone app for the emergency medical service system: a population-based

before-and-after observational study in Osaka City, Japan. JMIR Mhealth Uhealth 2017 Sep 11;5(9):e134 [FREE Full text] [doi: 10.2196/mhealth.8296] [Medline: 28893725]
13. Statistical book of Kamakura City in 2020. Kamakura City. URL: https://www.city.kamakura.kanagawa.jp/soumu/toukei/kamakuratoukei/top2/documents/02kamakuranotoukei0428.pdf [accessed 2022-05-03]
14. Emergency Department. Shonan Kamakura General Hospital. URL: https://www.skgh.jp/department/er [accessed 2022-05-03]
15. Product Guide for Hospitals. TXP Medical. URL: https://txpmedical.jp/ [accessed 2022-05-03]
16. Goto T, Hara K, Hashimoto K, Soeno S, Shirakawa T, Sonoo T, et al. Validation of chief complaints, medical history, medications, and physician diagnoses structured with an integrated emergency department information system in Japan: the Next Stage ER system. Acute Med Surg 2020;7(1):e554 [FREE Full text] [doi: 10.1002/ams2.554] [Medline: 32884825]
17. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol 2017 Feb 01;46(1):348-355 [FREE Full text] [doi: 10.1093/ije/dyw098] [Medline: 27283160]
18. Ohbe H, Goto T, Matsui H, Fushimi K, Yasunaga H. Associations of government-issued intensive care unit admission criteria with clinical practices, outcomes, and intensive care unit bed occupancy. Ann Am Thorac Soc 2022 Jun;19(6):1013-1021. [doi: 10.1513/AnnalsATS.202107-844OC] [Medline: 34813412]
19. Sato D, Goto T, Uda K, Kumazawa R, Matsui H, Yasunaga H. Impact of national guidelines for antimicrobial stewardship to reduce antibiotic use in upper respiratory tract infection and gastroenteritis. Infect Control Hosp Epidemiol 2021 Mar;42(3):280-286. [doi: 10.1017/ice.2020.427] [Medline: 32959738]
20. Sakai K, Komatsu T, Iguchi Y, Takao H, Ishibashi T, Murayama Y. Reliability of smartphone for diffusion-weighted imaging-Alberta Stroke Program Early Computed Tomography Scores in acute ischemic stroke patients: diagnostic test accuracy study. J Med Internet Res 2020 Jun 09;22(6):e15893 [FREE Full text] [doi: 10.2196/15893] [Medline: 32515744]
21. Sakai K, Sato T, Komatsu T, Mitsumura H, Iguchi Y, Ishibashi T, et al. Communication-type smartphone application can contribute to reducing elapsed time to reperfusion therapy. Neurol Sci 2021 Nov;42(11):4563-4568 [FREE Full text] [doi: 10.1007/s10072-021-05132-2] [Medline: 33638012]
22. Current status and issues of the emergency medical system. Ministry of Health, Labour, and Welfare. URL: https://www.mhlw.go.jp/content/10802000/000328610.pdf [accessed 2022-05-03]
23. Lopez Bernal J, Cummins S, Gasparrini A. The use of controls in interrupted time series studies of public health interventions. Int J Epidemiol 2018 Dec 01;47(6):2082-2093. [doi: 10.1093/ije/dyy135] [Medline: 29982445]
24. Hawley S, Ali MS, Berencsi K, Judge A, Prieto-Alhambra D. Sample size and power considerations for ordinary least squares interrupted time series analysis: a simulation study. Clin Epidemiol 2019;11:197-205 [FREE Full text] [doi: 10.2147/CLEP.S176723] [Medline: 30881136]

Abbreviations

ED: emergency department
EMS: emergency medical services
ITSA: Interrupted time-series analysis
NSER mobile: NEXT Stage ER mobile

Edited by A Mavragani; submitted 15.02.22; peer-reviewed by Y Katayama, T Kitamura, Y Chu, D Sato; comments to author 19.04.22; revised version received 19.05.22; accepted 20.06.22; published 06.07.22.

Please cite as:
Fukaguchi K, Goto T, Yamamoto T, Yamagami H
Experimental Implementation of NSER Mobile App for Efficient Real-Time Sharing of Prehospital Patient Information With Emergency Departments: Interrupted Time-Series Analysis
JMIR Form Res 2022;6(7):e37301
URL: https://formative.jmir.org/2022/7/e37301
doi:10.2196/37301
PMID:35793142

Original Paper

Chinese Version of the Mobile Health App Usability Questionnaire: Translation, Adaptation, and Validation Study

Yi Shan[1], Prof Dr; Meng Ji[2], PhD; Wenxiu Xie[3], MPhil; Rongying Li[4], BM; Xiaobo Qian[5], BEng; Xiaomin Zhang[6], PhD; Tianyong Hao[5], Prof Dr

[1]School of Foreign Studies, Nantong University, Nantong, China
[2]School of Languages and Cultures, University of Sydney, Sydney, Australia
[3]Department of Computer Science, City University of Hong Kong, Hong Kong, China
[4]School of Artificial Intelligence, South China Normal University, Guangzhou, China
[5]School of Computer Science, South China Normal University, Guangzhou, China
[6]Department of Linguistics, Macquarie University, Sydney, Australia

Corresponding Author:
Yi Shan, Prof Dr
School of Foreign Studies
Nantong University
Number 9, Seyuan Road
Nantong, 226019
China
Phone: 86 15558121896
Email: victorsyhz@hotmail.com

Abstract

Background: The usability of mobile health (mHealth) apps needs to be effectively evaluated before they are officially approved to be used to deliver health interventions. To this end, the mHealth App Usability Questionnaire (MAUQ) has been designed and proved valid and reliable in assessing the usability of mHealth apps. However, this English questionnaire needs to be translated into other languages, adapted, and validated before being utilized to evaluate the usability of mHealth apps.

Objective: This study aims to improve, further adapt, and validate the Chinese version of the MAUQ (C-MAUQ; interactive for patients) on *Left-handed Doctor*, one of the most popular "reaching out to patients" interactive mHealth apps with chatbot function in China, to test the reliability and cross-cultural adaptability of the questionnaire.

Methods: The MAUQ (interactive for patients) has been translated into Chinese and validated for its reliability on *Good Doctor*, one of the most influential "reaching out to patients" mHealth apps without chatbot function in China. After asking for the researchers' approval to use this Chinese version, we adjusted and further adapted the C-MAUQ by checking it against the original English version and improving its comprehensibility, readability, idiomaticity, and cross-cultural adaptability. Following a trial survey completed by 50 respondents on *wenjuanxing*, the most popular online questionnaire platform in China, the improved version of the C-MAUQ (I-C-MAUQ) was finally used to evaluate the usability of *Left-handed Doctor* through an online questionnaire survey (answered by 322 participants) on *wenjuanxing*, to test its internal consistency, reliability, and validity.

Results: The I-C-MAUQ still retained the 21 items and 3 dimensions of the original MAUQ: 8 items for usability and satisfaction, 6 items for system information arrangement, and 7 items for efficiency. The translation problems in the C-MAUQ, including (1) redundancy, (2) incompleteness, (3) misuse of parts of speech, (4) choice of inappropriate words, (5) incomprehensibility, and (6) cultural difference–induced improper translation, were improved. As shown in the analysis of data obtained through the online survey, the I-C-MAUQ had a better internal consistency (ie, the correlation coefficient between the score of each item and the total score of the questionnaire determined within the range of 0.861-0.938; $P<.01$), reliability (Cronbach $\alpha=.988$), and validity (Kaiser–Meyer–Olkin=0.973), compared with the C-MAUQ. It was effectively used to test the usability of *Left-handed Doctor*, eliciting over 80% of informants' positive attitudes toward this mHealth app.

Conclusions: The I-C-MAUQ is highly reliable and valid for *Left-handed Doctor*, and suitable for testing the usability of interactive mHealth apps used by patients in China. This finding further confirms the cross-cultural validity, reliability, and adaptability of the MAUQ. We identified certain factors influencing the perceived usability of mHealth apps, including users' age, gender, education, profession, and possibly previous experience with mHealth apps and the chatbot function of such apps.

Most notably, we found a wider acceptance of this new technology among young Chinese female college students who were more engaged in the interaction with health care chatbots. The age-, gender-, and profession-induced preference for new digital health interventions in China aligns with the findings in other similar studies in America and Malaysia. This preference identifies areas for further research on the social, cultural, and gender adaptation of health technologies.

(JMIR Form Res 2022;6(7):e37933) doi:10.2196/37933

KEYWORDS

mHealth app; usability; Chinese version of MAUQ; improved translation; validity; stability; reliability; cross-cultural adaptability; mobile phone

Introduction

Background

Mobile health (mHealth) apps have been applied to deliver health interventions (eg, health education, health monitoring, recommendations on treatments) to alleviate the overburdened health systems in many countries. These apps can perform versatile tasks, including health management, behavior intervention, health data collection, self-diagnosis, disease management, medication management, rehabilitation, and acting as patient portals [1,2], improving medication compliance, saving time in diagnosis and treatment, and reducing medical costs [3-6]. Given these wide applications and diverse advantages, these apps need to be assessed for hidden expenses, heavy data entry burden, and interest loss [7] to ensure accurate data analysis before being put into use [8].

To effectively evaluate the usability of mHealth apps, different questionnaires were designed [9], among which the most popular are the System Usability Scale (SUS) and the Post-Study System Usability Questionnaire (PSSUQ) [10,11]. Although used to reliably measure certain usability aspects of mobile apps, the SUS and the PSSUQ, among others, failed to provide tailored information on the factors unique to mobile apps [10,12]. Zhou et al [9] developed and validated the mHealth App Usability Questionnaire (MAUQ), which was solely designed for assessing the usability of mHealth apps, attesting its reliability and validity. The MAUQ [9] was exclusively developed to evaluate the usability of mHealth apps. It has 4 versions designed to assess interactive or standalone mHealth apps among patients or health care providers. It shows a strong internal consistency, evidenced by the Cronbach α coefficients of its 3 dimensions (.895 for *ease of use and satisfaction*, .829 for *system information arrangement*, and .900 for *usefulness*) and the overall Cronbach α of .914. The items in the 3 dimensions are rated on a 7-point Likert scale from 1 (extremely strongly agree) to 7 (extremely strongly disagree). The usability of an app can be determined by calculating the total points and determining the average points of the responses to all statements: the closer the average is to 1, the higher the usability of the app [9].

Two more recent studies translated and adapted the MAUQ into Chinese [13] and Malay [14], respectively, finding that the Chinese and Malay versions exhibited high reliability and validity similar to those of the original English version [13,14]. The Chinese version of the MAUQ (C-MAUQ; interactive for patients) was testified to be reliable and valid, with content validity index of 0.952, Cronbach α of .912, value of test-retest reliability of 0.896, and value of the split-half reliability of 0.701 [14]. The Malay version of the MAUQ (standalone for patients) was proved to be reliable for evaluating the usability of the mHealth apps (Cronbach α=.946) [13]. Considering the painstaking efforts and considerable time and cost investment involved in developing new questionnaires [14], Marzuki et al [12] strongly recommended that established, accessible, and reliable questionnaires should be adapted, validated, and recorded cross-linguistically.

Left-handed Doctor is one of the most popular "reaching out to patients" [15] interactive mHealth apps in China. It integrates artificial intelligence technologies, such as deep learning, big data processing, semantic understanding, and interactive medical dialog with medicine and is committed to using artificial intelligence technology to expand the supply of high-quality medical resources. The *Left-handed Doctor* open platform provides solutions, such as smart hospitals, diagnostic robots for consultation rooms, intelligent online consultation, intelligent postdiagnosis management, and artificial intelligence internet hospitals. In combination with different application scenarios, it provides high-quality medical services for all parties, empowering the health care industry. Although it is popular among many people in China, no studies have empirically tested its usability using the C-MAUQ.

Objective

Informed by the MAUQ and its culturally adapted versions, this study aimed to testify further the reliability, validity, and cross-cultural adaptability of the MAUQ for its suitability to the mHealth app usability test. This was achieved by applying the improved version of the C-MAUQ (I-C-MAUQ) to *Left-handed Doctor*, one of the most popular "reaching out to patients" interactive mHealth apps with chatbot function in China. Two facts warrant this study: (1) the *Left-handed Doctor* app is different from the *Good Doctor* app: the former is empowered with the chatbot function, while the latter is not, and we thought that this difference would influence users' perceived usability of these apps; and (2) the informants differ from those in Mustafa et al [13] in terms of age, gender, education, and profession, and we believed that these differences would also impact users' perceived usability of these apps.

Methods

Overview

This study used the C-MAUQ [15] but made some improvements. The study was conducted from February 18 to March 8, 2022.

Improvement of the C-MAUQ

We first obtained the approval of the researchers [15] to use the C-MAUQ. Afterward, 2 translators (YS and MJ) independently adjusted this version by checking it against the original English version and improving its readability and idiomaticity. The I-C-MAUQ still retained the 21 items and 3 dimensions of the original MAUQ and the C-MAUQ: 8 items for usability and satisfaction, 6 items for system information arrangement, and 7 items for efficiency. Improper translations of all the 21 items in the C-MAUQ were modified through discussion among the whole research team.

Improvement of Cross-cultural Adaptation

The C-MAUQ has been adapted cross-culturally through experts' comments and a prediction test [15]. Based on this adaptation and drawing on Conway et al's translatability assessment (TA) [16], this study further adapted the C-MAUQ by inviting a group of bilingual translators and health educators to assess the comprehensibility of the content as well as the cultural relevance and appropriateness of each item. Subsequently, the revised version was subjected to a trial survey online, in which 50 college students participated to identify problems that needed to be resolved.

Informants and Online Survey

Participants were students of the School of Foreign Studies, Nantong University, China. Impacted by varying degrees of psychological problems that became increasingly serious during the repeated COVID-19 attacks, these students urgently needed mHealth apps for self-diagnosis and general health information to relieve their psychologically strained minds. The questionnaire was administered using the online questionnaire survey platform named *wenjuanxing* [17] on February 18, 2022, and the survey lasted until no additional questionnaire was submitted online for 2 consecutive days (March 4, 2022). Over this period, the survey was announced to the entire student body of over 1000 at the School of Foreign Studies, Nantong University, through emails and WeChat groups. Meanwhile, the candidate informants were requested to use the *Left-handed Doctor* app for 2 days to become familiar with it before answering the questionnaire. The majority of participants in this study were female, which is characteristic of all schools of foreign studies in China.

Data Collection

The survey was conducted through *wenjuanxing* [17], the most popular online questionnaire platform in China. Two categories of data were collected via online questionnaires: the demographic information of the participants and their ratings on the 21 items concerning the usability of *Left-handed Doctor*. The demographic data included the informants' age, gender, grade, and channel to obtain health care information. The usability test elicited data concerning the informants' ratings of the 21 items based on a 7-point Likert scoring system from 1 to 7 points (representing "strongly agree," "agree," "somewhat agree," "neither agree nor disagree," "somewhat disagree," "disagree," and "strongly disagree," respectively).

Data Analysis

Quantitative analyses were conducted using SPSS version 22.0 (IBM, Inc.). First, demographic data were presented in a table and briefly described as the background information of the analysis. Subsequently, item analysis, weight analysis, and Pearson correlation analysis were conducted, followed by the reliability, validity, test-retest reliability, and split-half reliability tests. Finally, the range, mean values, and SD of the collected usability data were calculated and described for each of the 21 items.

Ethics Considerations

This study was approved and supported by the Student Affairs Office and the Humanities and Social Sciences Office of Nantong University, which is authorized to provide such approval before collecting data from students.

Results

Improvement of the C-MAUQ

Both translators (YS and MJ) found items 1, 2, 5, 9, 11-14, 17-21 problematic after checking the C-MAUQ against the original English version independently. They modified these items independently, and then, through discussion, agreed on the corresponding revisions and the classification of translation problems, which were subjected to further amendments before a final consensus among the study researchers. The translation problems in the C-MAUQ were related to (1) redundancy (items 1, 2, and 18); (2) incompleteness (item 12); (3) misuse of parts of speech (items 5, 9, and 17); (4) choice of inappropriate words (items 5, 9, 14, and 18-21); and (5) incomprehensibility (items 9, 11, and 13).

Further Cross-cultural Adaptation

The I-C-MAUQ was further adapted cross-culturally through a panel meeting attended by a group of bilingual translators and health educators. This meeting identified and agreed on a common problem concerning inappropriate cultural adaptation of items 18-21. In English-speaking countries, a patient always visits the same doctor and addresses the doctor as "my health care provider." By contrast, in China, a patient usually sees different doctors when becoming ill and thus never uses "my" when referring to his/her "health care provider." Therefore, "my" was crossed out from these 4 items. No other problems were detected during the panel meeting. After the panel meeting, the comprehensibility, readability, idiomaticity, and cultural adaptability of the questionnaire content were further improved. Subsequently, the I-C-MAUQ version was validated in an online trial survey completed by 50 informants. The trial survey turned out to be successful (Cronbach α=.992), and so the I-C-MAUQ did not require further improvement. The I-C-MAUQ, together with the C-MAUQ and the MAUQ, is provided in Multimedia Appendix 1.

Informant Demographics

Multimedia Appendix 2 shows the informants' demographic information. A total of 322 responses were collected online, including 292 (90.7%) from female respondents. This can be explained by the fact that over 90% of students studying in the

School of Foreign Studies, Nantong University, are females. The age of the participants ranged from 18 to 33 years (mean 21.68, SD 2.30 years). The overwhelming majority (n=316, 98.1%) were aged between 18 and 26 years. The informants included freshman (n=64, 19.9%), sophomore (n=29, 9.0%), junior (n=88, 27.3%), senior (n=48, 14.9%), first-year postgraduate candidates (n=46, 14.3%), and second-year postgraduate candidates (n=47, 14.6%). The majority of the informants (n=306, 95.0%) obtained health care information by visiting a doctor; logging into the internet; and communicating with families, friends, and classmates. Only a minor percentage of participants (n=9, 2.8%) used mHealth apps to obtain health care information.

Questionnaire Item Analysis

The 21 items in the I-C-MAUQ were valid and appropriately designed (Table 1), as evidenced by the distinction between the high-score group (n=94) and the low-score group (n=149). Data below the 27% quantile belonged to the low-score group, and those above the 73% quantile belonged to the high-score group. There was a significant difference in each of the 21 items between the high-score group and the low-score group, with *P* value in each case being <.001 (ie, *P*<.01). This indicates that all 21 items could well be distinguished from one another and thus should all be retained in the final version of the questionnaire. Besides, all the 21 items were significant (Table 2), with critical values (CR) determined within the range of 14.751-19.449 and the *P* value (CR) calculated at <.001 (ie, *P*<.01). The correlation coefficient between the score of each item and the total score of the questionnaire was determined within the range of 0.861-0.938 (*P*<.01). Thus, all the 21 items were retained. According to the Pearson correlation values (Table S1 of Multimedia Appendix 3), all the 21 items were significantly and positively correlated, with the correlation coefficients ranging from 0.688 to 0.921 and *P*<.01.

Table 1. Item analysis.

Items[a]	Group, mean (SD)		*t* (critical values)	*P* value[b]
	Low-score group (*n*=149)	High-score group (*n*=94)		
1	1.58 (0.57)	3.65 (1.08)	17.031	<.001
2	1.52 (0.51)	3.34 (1.12)	14.751	<.001
3	1.56 (0.55)	3.54 (1.02)	17.251	<.001
4	1.58 (0.54)	3.66 (1.08)	17.352	<.001
5	1.58 (0.58)	3.51 (1.08)	15.961	<.001
6	1.57 (0.56)	3.53 (1.07)	16.348	<.001
7	1.58 (0.55)	3.85 (1.05)	19.449	<.001
8	1.56 (0.52)	3.62 (1.06)	17.555	<.001
9	1.61 (0.61)	3.60 (1.17)	15.225	<.001
10	1.53 (0.51)	3.61 (1.03)	18.186	<.001
11	1.52 (0.51)	3.49 (1.05)	16.905	<.001
12	1.58 (0.57)	3.46 (0.99)	16.724	<.001
13	1.52 (0.54)	3.40 (1.04)	16.262	<.001
14	1.56 (0.52)	3.65 (1.04)	18.038	<.001
15	1.55 (0.53)	3.55 (1.06)	16.993	<.001
16	1.56 (0.56)	3.56 (1.11)	16.178	<.001
17	1.56 (0.52)	3.51 (1.09)	16.242	<.001
18	1.55 (0.53)	3.54 (1.04)	17.182	<.001
19	1.55 (0.55)	3.50 (1.03)	16.832	<.001
20	1.67 (0.67)	3.68 (0.98)	17.527	<.001
21	1.56 (0.52)	3.57 (1.08)	16.862	<.001

[a]Items 1-21 represent the 21 items in the questionnaire.

[b]All *P* values <.01.

Table 2. Correlation between the 21 items and the overall score of the questionnaire.

Items	CR[a]	*P* value (CR)	COSQ[b]	*P* value[c] (COSQ)
1	17.031	<.001	0.874	<.001
2	14.751	<.001	0.885	<.001
3	17.251	<.001	0.902	<.001
4	17.352	<.001	0.907	<.001
5	15.961	<.001	0.861	<.001
6	16.348	<.001	0.883	<.001
7	19.449	<.001	0.890	<.001
8	17.555	<.001	0.921	<.001
9	15.225	<.001	0.879	<.001
10	18.186	<.001	0.925	<.001
11	16.905	<.001	0.938	<.001
12	16.724	<.001	0.923	<.001
13	16.262	<.001	0.906	<.001
14	18.038	<.001	0.923	<.001
15	16.993	<.001	0.914	<.001
16	16.178	<.001	0.879	<.001
17	16.242	<.001	0.910	<.001
18	17.182	<.001	0.912	<.001
19	16.832	<.001	0.896	<.001
20	17.527	<.001	0.869	<.001
21	16.862	<.001	0.905	<.001

[a]CR: critical value.

[b]COSQ: correlation with the overall score of the questionnaire.

[c]All *P* values <.01.

Weight of the 21 Items in the Questionnaire

Through the analytic hierarchy process, the weight of each of the 21 items in the questionnaire was determined. Based on the judgment matrix of the 21 items (Table S2 of Multimedia Appendix 3), the eigenvector and weight of each item were determined (Table 3). Drawing on the eigenvectors, the maximum eigenvalue (21.000) was worked out. According to the maximum eigenvalue, the CI (<0.001) was computed. According to Table 4, the random index (RI) of the judgment matrix was 1.6358. From the CI (<0.001) and the RI (1.6358), CR (<0.001) was finally calculated (Table 5). This CR value (<0.1) indicated that the judgment matrix passed the consistency test. Therefore, the weights of the 21 items in Table 3 were valid. These weight values meant that the 21 items were almost equally important in the questionnaire.

Table 3. Analytic hierarchy process analysis of the 21 items in the questionnaire[a].

Items	Eigenvectors	Weight (%)
1	1.018	4.846
2	0.954	4.541
3	0.990	4.712
4	1.010	4.808
5	1.020	4.858
6	1.007	4.795
7	1.071	5.099
8	0.990	4.712
9	1.015	4.833
10	0.992	4.725
11	0.956	4.554
12	0.980	4.668
13	0.956	4.554
14	1.015	4.833
15	0.994	4.731
16	0.983	4.681
17	0.996	4.744
18	1.004	4.782
19	0.995	4.738
20	1.056	5.029
21	0.999	4.757

[a]Maximum eigenvalue: 21.000; CI<0.001.

Table 4. RI[a] table of the judgment matrix.

Order	3	4	5	6	7	8	9	10	11	12	13	14	15	16
RI	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49	1.52	1.54	1.56	1.58	1.59	1.5943
Order	17	18	19	20	21	22	23	24	25	26	27	28	29	30
RI	1.6064	1.6133	1.6207	1.6292	1.6358	1.6403	1.6462	1.6497	1.6556	1.6587	1.6631	1.6670	1.6693	1.6724

[a]RI: random index.

Table 5. Consistency test of the weight of the 21 items.

Maximum eigenvalue	CI	RI[a]	Critical value	Result of test
21.000	<0.001	1.636	<0.001	Pass

[a]RI: random index.

Questionnaire Reliability and Validity

The statistics in Table 6 indicate the high reliability of the questionnaire. The corrected item-total correlation values of the 21 items all fell within 0.845-0.931, far exceeding 0.4. This meant that the 21 items were strongly correlated, and that they all had a high degree of reliability. Besides, the Cronbach α did not apparently increase when each of the 21 items was deleted, which implied that all items should be retained in the questionnaire. The overall Cronbach α (.988) for the 21 items was well above 0.9, indicating that the data collected for each item in the questionnaire were highly reliable. The values of test-retest reliability and split-half reliability were 0.918 and 0.828, respectively. Therefore, all the data were suitable for further analysis.

Table 6. Questionnaire reliability (and internal consistency).

Items	Corrected item-total correlation	Cronbach α if item deleted[a]
1	0.860	.988
2	0.873	.988
3	0.891	.988
4	0.897	.988
5	0.845	.988
6	0.870	.988
7	0.877	.988
8	0.912	.987
9	0.866	.988
10	0.917	.987
11	0.931	.987
12	0.915	.987
13	0.896	.988
14	0.914	.987
15	0.905	.987
16	0.866	.988
17	0.900	.987
18	0.902	.987
19	0.885	.988
20	0.855	.988
21	0.895	.988

[a]Cronbach α (standardized)=.988.

Table 7 reveals that the questionnaire is highly valid. The communalities for all 21 items ranged from 0.738 to 0.881, well above 0.4, indicating that the data can effectively be extracted from all these items. The Kaiser–Meyer–Olkin (KMO) value (0.973) was above 0.9, which showed that all the data concerning the 21 items could effectively be extracted. The percentage of variance (rotated) for factor 1 was 81.053%, considerably above 50%, meaning that all the data on all the items can validly be extracted.

Table 7. Questionnaire validity.

Items	Factor loadings (factor 1)	Communalities[a]
1	0.873[b]	0.762
2	0.885[b]	0.784
3	0.902[b]	0.813
4	0.907[b]	0.822
5	0.859[b]	0.738
6	0.882[b]	0.778
7	0.889[b]	0.790
8	0.921[b]	0.848
9	0.878[b]	0.771
10	0.925[b]	0.856
11	0.939[b]	0.881
12	0.924[b]	0.854
13	0.907[b]	0.823
14	0.923[b]	0.852
15	0.915[b]	0.837
16	0.880[b]	0.774
17	0.911[b]	0.830
18	0.912[b]	0.832
19	0.896[b]	0.803
20	0.868[b]	0.754
21	0.905[b]	0.819
Eigenvalues (initial)	17.021	N/A[c]
Variance (%) (initial)	81.053	N/A[c]
Cumulative variance (%) (initial)	81.053	N/A[c]
Eigenvalues (rotated)	17.021	N/A[c]
Variance (%) (rotated)	81.053	N/A[c]
Cumulative variance (%) (rotated)	81.053	N/A[c]
Kaiser–Meyer–Olkin	0.973	N/A[c]
Bartlett test of sphericity (chi-square); *df*	10873.765; 210	N/A[c]
P value	<.001	N/A[c]

[a]The communality is less than 0.4.

[b]The absolute value of loading is greater than 0.4.

[c]N/A: not applicable.

Usability of the *Left-handed Doctor* App

Table 8 presents the results of the descriptive analysis of the usability of *Left-handed Doctor*. The range, mean (SD), and median scores were based on the rating of each item (1=strongly agree; 2=agree; 3=somewhat agree; 4=neither agree nor disagree; 5=somewhat disagree; 6=disagree; and 7=strongly disagree). The mean scores of the 21 items were between 2.224 and 2.497, indicating that the respondents were inclined to agree

with the statements in all 21 items. In other words, they found the *Left-handed Doctor* app usable on the whole.

There were no significant differences (*P*=.35) in the mean scores concerning the 3 dimensions of usability and satisfaction (items 1-8), the arrangement of system information (items 9-14), and efficiency (items 15-21). This implied that the participants found the *Left-handed Doctor* app equally usable when it comes to the 3 dimensions.

Multimedia Appendix 4 shows the proportion of respondents falling into each of the 7 ratings of the 21 items. Over 60% (205/322, 63.7%; 223/322, 69.3%; 209/322, 64.9%; 206/322, 64.0%; 199/322, 61.8%; 206/322, 64.0%; 210/322, 65.2%; 203/322, 63.0%; 208/322, 64.6%; 219/322, 68.0%; 211/322, 65.5%; 218/322, 67.7%; 198/322, 61.5%; 208/322, 64.6%; 216/322, 67.1%; 207/322, 64.3%; 198/322, 61.5%; 203/322, 63.0%; 208/322, 64.6%, for items 1-6, 8-19, and 21, respectively) of informants *strongly agreed* or *agreed* with all items but items 7 (183/322, 56.8%) and 20 (187/322, 58.1%). More than 80% (267/322, 82.9%; 285/322, 88.5%; 277/322, 86.0%; 277/322, 86.0%; 271/322, 84.2%; 277/322, 86.0%; 259/322, 80.4%; 282/322, 87.6%; 270/322, 83.9%; 279/322, 86.6%; 287/322, 89.1%; 285/322, 88.5%; 288/322, 89.4%; 277/322, 86.0%; 280/322, 87.0%; 280/322, 87.0%; 282/322, 87.6%; 276/322, 85.7%; 276/322, 85.7%; 258/322, 80.1%; 277/322, 86.0%, for items 1-21, respectively) of participants *strongly agreed*, *agreed*, or *somewhat agreed* with all the 21 items. This meant that the vast majority of the participating students showed a positive attitude toward the usability of the *Left-handed Doctor* app.

Table 8. Descriptive analysis of the usability of the *Left-handed Doctor* app.

Item	Samples, n	Range	Mean (SD)	Median
1	322	1.000-7.000	2.373 (1.180)	2.000
2	322	1.000-7.000	2.224 (1.079)	2.000
3	322	1.000-7.000	2.307 (1.125)	2.000
4	322	1.000-7.000	2.354 (1.160)	2.000
5	322	1.000-7.000	2.379 (1.176)	2.000
6	322	1.000-7.000	2.348 (1.170)	2.000
7	322	1.000-7.000	2.497 (1.246)	2.000
8	322	1.000-7.000	2.307 (1.136)	2.000
9	322	1.000-7.000	2.366 (1.182)	2.000
10	322	1.000-7.000	2.314 (1.140)	2.000
11	322	1.000-7.000	2.230 (1.101)	2.000
12	322	1.000-7.000	2.286 (1.070)	2.000
13	322	1.000-7.000	2.230 (1.089)	2.000
14	322	1.000-7.000	2.366 (1.153)	2.000
15	322	1.000-7.000	2.317 (1.132)	2.000
16	322	1.000-7.000	2.292 (1.153)	2.000
17	322	1.000-7.000	2.323 (1.117)	2.000
18	322	1.000-7.000	2.342 (1.125)	2.000
19	322	1.000-7.000	2.320 (1.119)	2.000
20	322	1.000-7.000	2.463 (1.166)	2.000
21	322	1.000-7.000	2.329 (1.137)	2.000

Discussion

Principal Findings

Informed by Zhou et al [9] and Mustafa et al [13], the study improved the C-MAUQ translated, adapted, and validated in Zhao et al [14], and then used the I-C-MAUQ to test the usability of *Left-handed Doctor*, one of the most popular "reaching out to patients" interactive mHealth apps in China. The I-C-MAUQ had a better internal consistency (the correlation coefficient between the score of each item and the total score of the questionnaire ranging from 0.861 to 0.938; *P*<.001), reliability (Cronbach α=.988), validity (load factor ranging from 0.859 to 0.939, percentage of cumulative variance [rotated]=81.053%, KMO=0.973), test-retest reliability (0.918), and split-half reliability (0.828) than the C-MAUQ [14]. Such better performance of the I-C-MAUQ resulted from 4 factors: (1) better comprehensibility, readability, and cultural adaptation of the I-C-MAUQ; (2) different categories of participants in terms of age, gender, education, profession, and sample size; (3) different functions of the tested interactive mHealth apps used by patients (with vs without the chatbot function); and (4) respondents' experience with mHealth apps. Similarly, we found

that the reliability of the I-C-MAUQ was relatively higher than those reported in Mustafa et al [13] (Cronbach α=.946; corrected item-total correlation values between –0.057 and 0.868) and Zhou et al [9] (Cronbach α=.914). We once again attributed the reliability difference to the aforesaid 4 factors, which will be discussed in the following sections.

Cross-cultural Adaptation of the Translated Questionnaire

It is imperative to adapt questionnaires cross-culturally, but there is a lack of evidence for the best approaches to cross-cultural adaptation (CCA) [18]. The most adopted methods for CCA are Brislin's Translation Model [19], the use of panels or committees [20-26], and focus groups [27]. However, this study adopted another effective but a commonly neglected model: TA [16]. Drawing on the cross-cultural issues proposed in TA, we improved the C-MAUQ [15] by making further cultural and linguistic adaptations, solving the translation problems concerning redundancy, incompleteness, misuse of parts of speech, choice of inappropriate words, incomprehensibility, and relevance and appropriateness on the cultural, semantic, syntactic, and pragmatic facets. The newly adapted questionnaire was equivalent to the original questionnaire [18]. TA thus makes it possible to identify alternative versions for translation purposes, modify original versions to optimize subsequent translation efforts, and detect and discuss irrelevant or inappropriate items early [16]. Thus, TA needs to be adopted as an effective CCA method in prospective translation and adaptation of questionnaires.

Participant Differences in Age, Gender, Education, Profession, and Sample Size

Most (318/322, 98.8%) of the informants in this study were aged 18-28, compared with the majority (91.04%) of respondents aged 29-65 in Zhao et al [14], with just over half (52.3%) of the participants aged 18-28 and just below half (48.3%) aged 29-65 in Zhou et al [9], and with all (100%) those surveyed aged 22-25 in Mustafa et al [13]. We concluded that younger age potentially led to relatively positive ratings of questionnaire items and thus higher questionnaire reliability and internal consistency.

The proportions of male and female participants (30/322, 9.3% vs 291/322, 90.4%) were different from those (53.76% vs 46.24%) in Zhao et al [14], those (38.3% vs 61.7%) in Zhou et al [9], and those (8% vs 92%) in Mustafa et al [13]. Therefore, considerably higher percentages (292/322, 90.7%) of female respondents seemed to contribute to a higher degree of the questionnaire's internal consistency and reliability. This result showed that females were more interested in participating in surveys on the usability of mHealth apps and that more female users of mobile apps were keen on using mHealth apps for health care. This has been also testified by Zhou et al [9].

All informants in this study and Mustafa et al [13] were college students at the undergraduate or graduate level, but those in Zhao et al [14] and Zhou et al [9] had different levels of education: 33.24% and 67.2% held an undergraduate or above in Zhao et al [14] and Zhou et al [9], respectively. The overall higher level of respondent education may explain the relatively higher degree of questionnaire's internal consistency and reliability in our study and Mustafa et al [13], in comparison with that in Zhao et al [14] and Zhou et al [9]. However, the vast gap in participant education at or above the undergraduate level between Zhao et al [14] and Zhou et al [9] merely resulted in a considerably minor difference in questionnaire reliability (Cronbach α=.912 vs .914).

In terms of profession, being a student—100% (322/322) in this study and Mustafa et al [13], 31.4% in Zhou et al [9], and 1.56% in Zhao et al [14]—also likely impacted the questionnaire's internal consistency and reliability, with the rate of students participating positively proportional to the degree of reliability and internal consistency.

These findings concerning age, gender, education, and profession contradicted the result in Zhou et al [9], which asserted that the demographic factors (eg, age, gender, education, occupation) failed to significantly impact the answers to the individual statements or the overall score on the MAUQ.

The sample size was indeed not a contributing factor to the high internal consistency and reliability of the questionnaire. Zhao et al [14] recruited the largest number of participants (n=346) but reported the lowest internal consistency and reliability, whereas this study achieved the highest internal consistency and reliability of the questionnaire based on the data contributed by a similar number of informants (n=322), followed by a slightly lower internal consistency and reliability derived from the information provided by the smallest number of informants in Mustafa et al [13].

Respondents' Experience With mHealth Apps

The informants in Zhou et al [9] used mobile apps for an average of 6.64 years; 86.42% of participants in Zhao et al [14] used mHealth apps more than 3 times during the month before the survey. Only 2.8% (9/322) of respondents in this study resorted to mHealth apps for health care information, but they were requested to install the *Left-handed Doctor* app 2 weeks beforehand to become familiar with it. The informants in Mustafa et al [13] were also asked to do the same. Therefore, experience with mHealth apps did not seem to influence the users' perceived usability, and thus the internal consistency and reliability of the questionnaire adopted remained unaffected.

Interactive mHealth Apps for Patients Equipped With or Without the Chatbot Function

This study tested the usability of the I-C-MAUQ on the *Left-handed Doctor* app, which is empowered with the chatbot function. By contrast, Zhao et al [14] adopted the *Good Doctor* app, which was not equipped with the chatbot function. This difference in apps may somewhat explain the notable discrepancy in the questionnaire's internal consistency and reliability between this study (Cronbach α=.988) and that by Zhao et al [14] (Cronbach α=.912). The mHealth apps used in Krebs and Duncan [7] and Mustafa et al [13] did not have the chatbot function. Thus, further research needs to be conducted to pinpoint the impact of this function on the usability of mHealth apps.

Implications

It is worth adapting established and appropriate questionnaires with recorded validity because designing a new one is effort-, time-, and cost-consuming [12]. Proper translation and adaptation and TA [16,28] are essential to ensure equivalence between the original questionnaire and the translated version. Cultural and linguistic sensitivity is a prerequisite for ironing out the translation problems resulting from cultural and linguistic differences and making the translated questionnaire culturally relevant and appropriate. Therefore, qualified translators highly proficient in the source and target languages and health educators or practitioners need to make joint efforts to complete this challenging task.

Validation is crucial for ensuring the equivalence between the original version and the translated one. Content validity index has been used to quantify the questionnaire validity in some studies [9,13,15,29,30]. It has been widely used because of its simple measurement, accessibility, power to provide details for each item, and indication of item modification or deletion [30].

Limitations

This study has several limitations. First, the convenient sampling of college students from a single university made it challenging to generalize the findings to the whole population in China. The recruitment of only healthy students also made the generalization of the results less convincing. Finally, the sample size was not sufficiently large to guarantee the generalization of findings.

Conclusions

The I-C-MAUQ is highly reliable and valid for the *Left-handed Doctor* app, and thus suitable for testing the usability of interactive mHealth apps used by patients in China. This finding is in line with the study by Marzuki et al [12], further confirming the cross-cultural validity, reliability, and adaptability of the MAUQ. We identified certain factors that influence the perceived usability of mHealth apps, including users' age, gender, education, profession, and possibly previous experience with mHealth apps as well as the chatbot function of such apps. Most notably, we found a wider acceptance of this new technology among young Chinese female college students who were more engaged in the interaction with health care chatbots. The age-, gender- and profession-induced preference for new digital health interventions in China aligns with the findings from other similar studies in the United States [9] and Malaysia [13]. This preference identifies areas for further research on the social, cultural, and gender adaptation of health technologies.

Conflicts of Interest

None declared.

Multimedia Appendix 1
I-C-MAUQ, together with the C-MAUQ and the MAUQ. See also [31].
[DOCX File , 62 KB - formative_v6i7e37933_app1.docx]

Multimedia Appendix 2
Informants' demographics.
[DOCX File , 16 KB - formative_v6i7e37933_app2.docx]

Multimedia Appendix 3
Additional tables.
[DOCX File , 32 KB - formative_v6i7e37933_app3.docx]

Multimedia Appendix 4
Frequency analysis of the usability of the *Left-handed Doctor* app.
[DOCX File , 29 KB - formative_v6i7e37933_app4.docx]

References

1. Kao C, Liebovitz DM. Consumer Mobile Health Apps: Current State, Barriers, and Future Directions. PM R 2017 May;9(5S):S106-S115. [doi: 10.1016/j.pmrj.2017.02.018] [Medline: 28527495]
2. Roess A. The Promise, Growth, and Reality of Mobile Health - Another Data-free Zone. N Engl J Med 2017 Nov 23;377(21):2010-2011. [doi: 10.1056/NEJMp1713180] [Medline: 29116869]
3. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Perceptions and experiences of heart failure patients and clinicians on the use of mobile phone-based telemonitoring. J Med Internet Res 2012 Feb 10;14(1):e25 [FREE Full text] [doi: 10.2196/jmir.1912] [Medline: 22328237]
4. Fairman AD, Dicianno BE, Datt N, Garver A, Parmanto B, McCue M. Outcomes of Clinicians, Caregivers, Family Members and Adults with Spina Bifida Regarding Receptivity to use of the iMHere mHealth Solution to Promote Wellness. Int J Telerehabil 2013;5(1):3-16 [FREE Full text] [doi: 10.5195/ijt.2013.6116] [Medline: 25945209]

5. Parmanto B, Pramana G, Yu DX, Fairman AD, Dicianno BE, McCue MP. iMHere: A Novel mHealth System for Supporting Self-Care in Management of Complex and Chronic Conditions. JMIR Mhealth Uhealth 2013 Jul 11;1(2):e10 [FREE Full text] [doi: 10.2196/mhealth.2391] [Medline: 25100682]
6. Pfammatter A, Spring B, Saligram N, Davé R, Gowda A, Blais L, et al. mHealth Intervention to Improve Diabetes Risk Behaviors in India: A Prospective, Parallel Group Cohort Study. J Med Internet Res 2016 Aug 05;18(8):e207 [FREE Full text] [doi: 10.2196/jmir.5712] [Medline: 27496271]
7. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101. [doi: 10.2196/mhealth.4924] [Medline: 26537656]
8. World Health Organization (WHO). Process of Translation and Adaptation of Instruments. Geneva, Switzerland: World Health Organization; 2020.
9. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500 [FREE Full text] [doi: 10.2196/11500] [Medline: 30973342]
10. Brooke J. SUS: A quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland AL, editors. Usability Evaluation in Industry. London, UK: Taylor and Francis; 1996:189-194.
11. Lewis JR. Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. International Journal of Human-Computer Interaction 2002 Sep;14(3-4):463-488. [doi: 10.1080/10447318.2002.9669130]
12. Mohamad Marzuki MF, Yaacob NA, Yaacob NM. Translation, Cross-Cultural Adaptation, and Validation of the Malay Version of the System Usability Scale Questionnaire for the Assessment of Mobile Apps. JMIR Hum Factors 2018 May 14;5(2):e10308 [FREE Full text] [doi: 10.2196/10308] [Medline: 29759955]
13. Mustafa N, Safii NS, Jaffar A, Sani NS, Mohamad MI, Abd Rahman AH, et al. Malay Version of the mHealth App Usability Questionnaire (M-MAUQ): Translation, Adaptation, and Validation Study. JMIR Mhealth Uhealth 2021 Feb 04;9(2):e24457 [FREE Full text] [doi: 10.2196/24457] [Medline: 33538704]
14. Zhao S, Cao Y, Cao H, Liu K, Lv X, Zhang J, et al. Chinese Version of the mHealth App Usability Questionnaire: Cross-Cultural Adaptation and Validation. Front. Psychol 2022 Feb 2;13:813309. [doi: 10.3389/fpsyg.2022.813309]
15. What Types of Telemedicine Can I Choose From? evisit. URL: https://evisit.com/resources/what-are-the-types-of-telemedicine/ [accessed 2022-06-07]
16. Conway K, Acquadro C, Patrick DL. Usefulness of translatability assessment: results from a retrospective study. Qual Life Res 2014 May 22;23(4):1199-1210. [doi: 10.1007/s11136-013-0572-9] [Medline: 24264803]
17. wenjuanxing. wjx. URL: https://www.wjx.cn/ [accessed 2022-06-07]
18. Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. J Clin Epidemiol 2015 Apr;68(4):435-441. [doi: 10.1016/j.jclinepi.2014.11.021] [Medline: 25698408]
19. Brislin RW. Comparative Research Methodology: Cross-Cultural Studies. International Journal of Psychology 1976 Jan;11(3):215-229. [doi: 10.1080/00207597608247359]
20. Hambleton R. Guidelines for adapting educational and psychological tests: a progress report. Eur J Psychol Assess 1994;10(3):229-244.
21. Hambleton RK. The Next Generation of the ITC Test Translation and Adaptation Guidelines. European Journal of Psychological Assessment 2001 Sep;17(3):164-172. [doi: 10.1027//1015-5759.17.3.164]
22. Erkut S. Developing Multiple Language Versions of Instruments for Intercultural Research. Child Dev Perspect 2010 Apr 01;4(1):19-24 [FREE Full text] [doi: 10.1111/j.1750-8606.2009.00111.x] [Medline: 21423824]
23. Hunt SM. Cross-cultural issues in the use of socio-medical indicators. Health Policy 1986 Jan;6(2):149-158. [doi: 10.1016/0168-8510(86)90004-7]
24. Hunt SM, Alonso J, Bucquet D, Niero M, Wiklund I, McKenna S. Cross-cultural adaptation of health measures. Health Policy 1991 Sep;19(1):33-44. [doi: 10.1016/0168-8510(91)90072-6]
25. Swaine-Verdier A, Doward LC, Hagell P, Thorsen H, McKenna SP. Adapting Quality of Life Instruments. Value in Health 2004 Sep;7:S27-S30. [doi: 10.1111/j.1524-4733.2004.7s107.x]
26. McKenna SP, Doward LC. The translation and cultural adaptation of patient-reported outcome measures. Value Health 2005 Mar;8(2):89-91 [FREE Full text] [doi: 10.1111/j.1524-4733.2005.08203.x] [Medline: 15804316]
27. Arredondo E, Mendelson T, Holub C, Espinoza N, Marshall S. Cultural adaptation of physical activity self-report instruments. J Phys Act Health 2012 Jan;9 Suppl 1(Suppl 1):S37-S43 [FREE Full text] [doi: 10.1123/jpah.9.s1.s37] [Medline: 22287446]
28. Streiner D, Norman G, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use (5th Edition). Oxford, UK: Oxford University Press; Jan 2015.
29. 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. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
30. Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi-Majd H, Nikanfar A. Design and Implementation Content Validity Study: Development of an instrument for measuring Patient-Centered Communication. J Caring Sci 2015 Jun;4(2):165-178 [FREE Full text] [doi: 10.15171/jcs.2015.017] [Medline: 26161370]

31. Hao K. The pandemic is emptying call centers. AI chatbots are swooping in. MIT Technology Review. 2020 May 14. URL: https://www.technologyreview.com/2020/05/14/1001716/ai-chatbots-take-call-center-jobs-during-coronaviruspandemic [accessed 2022-06-07]

Abbreviations

CCA: cross-cultural adaptation
CITC: the corrected item-total correlation
C-MAUQ: the Chinese version of the MAUQ
CR: critical value
I-C-MAUQ: the improved C-MAUQ
MAUQ: mHealth App Usability Questionnaire
mHealth: mobile health
PSSUQ: Post-Study System Usability Questionnaire
RI: random index
SUS: System Usability Scale
TA: translatability assessment

Edited by A Mavragani; submitted 11.03.22; peer-reviewed by O Pavliuk, I Dronyuk, L HongMin; comments to author 23.05.22; revised version received 27.05.22; accepted 27.05.22; published 06.07.22.

Please cite as:
Shan Y, Ji M, Xie W, Li R, Qian X, Zhang X, Hao T
Chinese Version of the Mobile Health App Usability Questionnaire: Translation, Adaptation, and Validation Study
JMIR Form Res 2022;6(7):e37933
URL: https://formative.jmir.org/2022/7/e37933
doi:10.2196/37933
PMID:35793132

Original Paper

Perspectives of Patients and Therapists on Social Media and Digital Data Use in Mental Health Therapy: Thematic Analysis

Lauren Southwick[1*], MPH; Rebecca Suh[1*], BA; Elissa Kranzler[2], PhD; Megan Bradley[1], BA; Raina M Merchant[1,3], MD

[1]University of Pennsylvania, Philadelphia, PA, United States

[2]Fors Marsh Group, Washington DC, DC, United States

[3]Department of Emergency Medicine, Perelman School of Medicine, Philadelphia, PA, United States

[*]these authors contributed equally

Corresponding Author:
Lauren Southwick, MPH
University of Pennsylvania
3600 Civic Center Blvd
Philadelphia, PA, 19014
United States
Phone: 1 9145826995
Email: lauren.southwick@pennmedicine.upenn.edu

Abstract

Background: Incorporating insights from social media into the patient-provider encounter is increasingly being explored in health care settings. Less is known about the utility of these data in mental health therapy.

Objective: This study aims to prospectively investigate and characterize how social media and digital data are used in mental health therapy from both the patient and mental health therapist perspective.

Methods: Patients enrolled in mental health therapy and mental health therapists were interviewed using a semistructured interview guide. All interviews were transcribed and coded using a deductive framework analysis. Themes and subthemes were identified. Participants completed a sociodemographic survey, while mental health therapists also completed a behavioral norms and elicitation survey.

Results: Seventeen participants, that is, 8 (48%) mental health therapists and 9 (52%) patients were interviewed. Overall, participants identified 4 themes and 9 subthemes. Themes were current data collection practices, social media and digital data in therapy, advantages of social media and digital data in therapy, and disadvantages of social media and digital data in therapy. Most subthemes were related to the advantages and disadvantages of incorporating digital data in mental health therapy. Advantage subthemes included convenience, objective, builds rapport, and user-friendliness while disadvantage subthemes were nonreflective, ethically ambiguous, and nongeneralizable. The mental health therapists' behavioral norms and elicitation survey found that injunctive and descriptive normative beliefs mapped onto 2 advantage subthemes: convenience and objectivity.

Conclusions: This qualitative pilot study established the advantages and disadvantages of social media and digital data use in mental health therapy. Patients and therapists highlighted similar concerns and uses. This study indicated that overall, both patients and therapists are interested in and are comfortable to use and discuss social media and digital data in mental health therapy.

(JMIR Form Res 2022;6(7):e32103) doi:10.2196/32103

KEYWORDS

social media; digital health; digital data; mental health therapy; mobile phone

Introduction

Background

Adults in the United States are frequent users of social media platforms such as Facebook, Instagram, and Twitter [1,2]. Such platforms have a profound impact on everyday life and provide new opportunities for understanding behavioral, social, and environmental determinants of mental health and well-being [3-5]. Social media data (eg, Facebook wall posts) and digital data (eg, search engine use, step data, smartphone metadata) are increasingly used to support patient mental health care [6]. Previous mental health research has demonstrated that social

media data can reveal and predict risk for mental health conditions such as depression, loneliness, suicide ideation, posttraumatic stress, schizophrenia, and bipolar disorder [6-12]. Furthermore, prior research has demonstrated that data from these digital platforms can provide critical information not readily attained through in-person or remote health care encounters to help therapists identify, address, or discuss mental health concerns [3,13,14].

In response to these findings, several researchers are increasingly capturing how and how often therapists incorporate social media data into mental health therapy and treatment. For example, Fisher and Appelbaum [15] report that some mental health clinicians incorporate parts of their patients' Facebook feeds in their care delivery, whereas Hobbs and colleagues [13] found that nearly two-thirds of outpatient psychotherapists report viewing at least one patient's social or electronic media (such as email messages, SMS text messaging, and other messaging apps) as part of psychotherapy. Hobbs et al [13] also report that the psychotherapists who access their patients' electronic or social media data indicated that it improved their ability to provide effective treatment. These findings underscore the utility of social media and digital data in mental health therapy. These examples reflect the growing body of research on the mental health care therapist experience using social media data in mental health therapy; however, research on the patient perspective remains limited. Accordingly, research exploring both patient and therapist perspectives on the use of social media data in mental health therapy is warranted.

Objective

This qualitative study aimed to provide new knowledge on mental health patient and therapist perspectives regarding the use of social media data in mental health therapy. The aims of this study were to (1) explore patients' and therapists' current use(s) of social media data in mental health therapy and (2) identify the advantages and disadvantages of sharing social media data in mental health therapy.

Methods

Recruitment

Individuals in mental health therapy (referred to as patients) were recruited through a clinical research registry at a large academic institution (a private university and medical center in the northeastern region of the United States) from March to May 2018. Thirteen applicants expressed interest from the registry, of which 9 were available for an in-person interview. Patient inclusion criteria included that they attended mental health therapy for anxiety or depression and were aged 18 years and older. Patients completed informed consent and received a US $20 gift card as compensation for their study participation. Mental health therapists (referred to as therapists) were recruited at the same large academic institution through behavioral health research consortiums and word-of-mouth sampling (approximately 50 received an email invitation) from October to December 2018. Twelve therapists expressed interest and 8 were available for an in-person interview. Therapist inclusion criteria included that they provided behavioral health care and worked with clients aged 18 years and older with depression or anxiety disorder(s). The therapists completed informed consent and received a US $100 gift card as compensation for study participation. All participants were willing to complete an in-person interview at a large academic hospital, which was held in a secure private room and lasted for 50-70 minutes.

Qualitative Interview

Two semistructured interview guides were constructed based on a review of the published literature on social media data in therapy and co-design techniques [15,16]. The interview guide was structured as follows: introduction, discussion of interview expectations, current use of data in therapy, experience with social media data (eg, Facebook, Instagram, Twitter) in therapy, and advantages and disadvantages of incorporating social media data in therapy. Interview questions were tailored to separately address the patient and therapist experiences. All interviews were conducted by 2 team members. One individual (a cisgender female researcher with >10 years of experience in qualitative interviewing) acted as the lead facilitator, while the other interviewer (a cisgender female service designer with 5 years of experience) took notes and contributed to probing and follow-up questions. Prompts were included, where appropriate, to elicit participant elaboration about each topic. The full interview guide is available from the authors upon request.

Survey

Patients completed a brief in-person survey at the end of the interview. The survey assessed sociodemographic characteristics such as gender, race/ethnicity, highest level of education, and the Social Media Use Questionnaire (SMUQ) [17]. The SMUQ assesses problematic use of social media and comprises 9 items (eg, "I feel anxious when I am not able to check my social network account"), with response options on a 5-point Likert scale from "Never" to "Always." All items were averaged into a scale for which higher scores corresponded with excessive social media use.

At 24 hours after the interview, therapists received a "thank you" email and a link to a web-based survey, programmed in Qualtrics software. The survey was intended to elicit participants' beliefs pertinent to using data about patients' social media use in therapy sessions, given little existing research on social media in mental health therapy. Survey questions were drawn from established procedures for elicitation studies [18]. These open-ended questions elicited beliefs about using social media data in therapy sessions in the next month. Types of beliefs elicited included behavioral beliefs (ie, potential benefits or drawbacks of using social media data in therapy sessions), injunctive normative beliefs (ie, individuals or groups of people who may or may not approve of using social media data in therapy sessions), descriptive normative beliefs (ie, individuals or groups of people who may or may not use social media data in therapy sessions), and control beliefs (ie, circumstances that may help or hinder the use of social media data in therapy sessions). The survey also assessed the sociodemographic characteristics of the participants, including gender, race, ethnicity, highest level of education, and work environment characteristics (eg, type of practice, caseload).

Ethics Approval

This study protocol was reviewed and approved by the University of Pennsylvania Institutional Review Board (protocol 831246).

Analysis

Qualitative Interviews

Therapist interviews were audio recorded and transcribed verbatim. Patient interviews were not recorded; a research assistant observed the interviews and took detailed notes and collected quotes. After multiple readings, the transcripts and interview notes were then coded by authors LS and RS. The coders created a codebook using the interview guide as themes, whereas subthemes emerged from the transcripts. The transcripts and interview notes were then analyzed using a deductive framework analysis [19] and coded according to the 6 stages of framework analysis: familiarization, identifying a thematic framework, indexing, charting, mapping, and interpretation. Identified subthemes were deduced from the coded passages and placed into separate coded charts. The authors then independently reviewed the charts for consistency and agreement. The coders met regularly to resolve disagreements for any theme or subtheme by consensus following discussion. The codebook is available from the authors upon request.

Survey Findings

All quantitative analyses (ie, descriptive statistics) were conducted in a Microsoft application (Microsoft Excel, version 16.58).

Results

Study Sample

The sociodemographic characteristics of the patients and therapists are provided in Table 1. Patients were mostly female (6/9, 66%) and non-Hispanic White (6/9, 66%). Patients were aged 22-60 years; half reported having completed college while the other half completed graduate school. According to the SMUQ, all patients reported using social media platforms (eg, Facebook, Instagram) more than 3 times a day and exhibited nonproblematic social media use (mean score 22, range 9-32). All therapists were females and non-Hispanic White, with a mean age of 37 years. Most reported that they work in a hospital/medical setting and practice cognitive behavioral therapy (CBT). However, therapists were at various stages in their careers, ranging from graduate student to psychologist or associate professor (Table 1).

Table 1. Sociodemographic characteristics of the patients and therapists.

Sociodemographic variables	Patients (n=9)		Therapists (n=8)	
	Frequency/mean	Range	Frequency/mean	Range
Age (years)	34	22-60	37	29-43
Female	6	N/A[a]	8	N/A
Ethnicity/race				
White	6	N/A	8	N/A
Black	1	N/A	0	N/A
Asian	1	N/A	0	N/A
Non-Hispanic, Latinx	1	N/A	8	N/A
Highest level of education				
College graduate	5	N/A	0	N/A
More than college graduate (eg, master's, doctoral degree)	4	N/A	8	N/A
Social media use questionnaire	22	9-32	N/A	N/A
Cognitive behavioral therapy orientation	N/A	N/A	8	N/A
Clinical experience (years)	N/A	N/A	5.3	2-9
Clinical setting				
Community outpatient	N/A	N/A	1	N/A
Hospital	N/A	N/A	4	N/A
Private practice	N/A	N/A	3	N/A
Caseload (clients)	N/A	N/A	N/A	5-40
Prep time prior to session (minutes)	N/A	N/A	N/A	3-30

[a]N/A: not applicable.

Thematic Areas

The interviews captured patients' and therapists' social media use in mental health therapy. The interviews identified the advantages and disadvantages of sharing social media data in mental health therapy and highlighted the contextual and logistical considerations to incorporating these new data. The interviews were structured on the following themes: (1) current data use in therapy, (2) experience with social media in therapy, (3) advantages of social media in therapy, and (4) disadvantages of social media in therapy (Table 2).

Table 2. Emerging interview subthemes with illustrative quote(s).

Themes, subthemes	Illustrative quote(s)	
	Therapist	Patient
Current data use in therapy		
Applications	*…I think it is a great idea because if you get the notification from the app and it is kind of fun, it's visually enticing, and people are more likely to do it. And then some of them have the built-in mindfulness tracking.*	*…I use the eMoods app to track my triggers and love the biometrics information. It tracks a lot things like eating vegetables, how much bread I eat, or if I went outside. I show my therapist it. eMoods doesn't feel like it's judgmental because I am creating the field(s).*
Surveys	*…Implement surveys during treatment to see change over time in symptoms or quality for life.*	*…Surveys are infrequent, I attend a self-determined care model.*
Experience with social media in therapy		
Patient-initiated	*…The patient brought up Snapchat and showed conversation around the video…It was sharing a sexting video with someone they were interested in and were having anxious thoughts about having done that and the individual saved/downloaded the video.*	*…I am comfortable sharing social media because it's already out there, and everyone puts their business out on social media.* *…I am willing to share [with my therapist] if I can explain and provide context.*
Provider-initiated	*…There's a lot of videos on YouTube to use for exposures. So…if I have somebody that comes in with like a serious vomit fear or emetophobia or fear of like, could be anything like scary movies or clowns. We just go on YouTube and look stuff up.*	*…I trust my instincts [on what to share].*
Advantages of incorporating social media in therapy		
Convenience	*…I would also love for there to be more technology where the person doesn't have to enter anything themselves. That's why the Fitbit is nice or this sleep app where you basically put under your pillow (and) it tracks your sleep… because requiring them to do any work when they're already depressed and anxious [and it can be] a huge burden.*	*…It would be nice to look back [at previous posts].*
Objective	*…I think sleep data could be really useful… Like I've had a lot of patients who tell me that they sleep for like 3 hours and I'm like that's not possible every single night. I think this could be useful.*	*…It would help flag my memory… keep me on track.*
Builds rapport	*…I don't think it derails anything… mostly it's very relevant and helpful. They're willing to share, and I think it builds rapport. Often times it helps me to like really understand what they're talking about.*	*…We talk about [text messages] as a "how are you?"*
User-friendly	*…It's nice that it graphs it for you…I would use it in the beginning of session when I'm asking how they are feeling right now and how they feel how the last month has been to sort of see if their self-report in the moment lines up.*	*…If therapy is holistic having this data might not be bad.*
Disadvantages of incorporating social media in therapy		
Nonreflective	*…I think that people sometimes present very differently in therapy than they might on social media. If I'm thinking of somebody who is super depressed or maybe [have] chronic mental health issues, I think that that person probably doesn't post a lot on social media about their depression.*	*…I trust my instincts on what to share.* *…I want to disclose what I need to [in each session]… maybe it's my eating disorder [that] has been more on my mind, not my substance abuse.*
Ethically ambiguous	*…For them to know that I have access to everything they post could make them feel pretty watched. They don't have any privacy because anything they do online the therapist will see it and will pull it up on their portal and judge them for it.*	*…My phone is so personal and feels like an invasion of privacy. It feels too "big brother."* *…I am not confident in accuracy of an algorithm.*
Nongeneralizable	*…I think there are different social media personalities. There are some people who are very explicitly like "I don't talk about politics on Facebook." And other people who use it as their diary where they are posting every thought. And then people who don't post very much because they're more private.*	*…I [would like to] annotate them (ie, text message or social media post) because things are left out.* *…I don't [want to be] judged by my digital data.*

Current Data Use in Therapy

When asked about current data collection practices, both patients and therapists noted using mental health and well-being apps in their sessions. Apps were mainly used during their practice to augment mental health services. For example, a patient reported using an app designed for patients with bipolar disorders I and II, depression, posttraumatic stress disorder, and anxiety disorder to keep track of symptoms and triggers (Table 2). However, several therapists expressed concern with selecting the "right" app, whereas others noted security and confidentiality issues for their patients. One therapist said, "[My patient] didn't want to put the app in her phone because her friends look at her phone and she thought that they would ask about it" (Participant #1, female, aged 42 years). In addition to using apps, therapists reported that they routinely collect patient-reported outcome data. As a common practice of CBT, validated surveys such as Beck's Depression Inventory, Patient Health Questionnaire 9-item, or Generalized Anxiety Disorder 7-item scales are collected at every therapy visit to track progress. Patients in this sample did not report completing validated surveys at their therapy sessions.

Experience Discussing Social Media in Therapy

The patients interviewed reported sharing and discussing social media data with their therapist. Most patients reported sharing or discussing Instagram posts (n=10), Facebook posts or statuses (n=8), tweets (ie, Twitter posts) (n=3), and YouTube videos (n=3). One patient noted, "I show my therapist my Facebook, it's evidence of my life; yes, I have friends, I exist" (Participant #3, genderqueer, aged 24 years).

Approximately half of the patients reported that they have directly shown or summarized social media posts to their therapist in the last month. Therapists noted similar interactions; one elaborated as follows, "So for [instance] a patient that is dating and wants to show me some of the people that she's talking about… she'll show me on social media. Or they want to show me what they post, a story that's going, something that they are involved with, like a family member or friends" (Participant #2, female, aged 35 years). Once social media data are introduced in therapy sessions, both patients and therapists detailed how social media platforms such as Facebook or YouTube can also be used as homework or in exposure therapy. For example, 1 patient explained that their therapist would have them write on a friend's Facebook wall to address their social anxiety and phobia.

Advantages of Incorporating Social Media in Therapy

Both patients and therapists highlighted how social media data can be convenient, objective, user-friendly, and allow them to build rapport when used in therapy (Table 2). Both patients and therapists focused on how data are shared, may it be automatic or manual and how the algorithm could select posts as an advantage. Patients indicated a general acceptance of an algorithm selecting their social media posts but emphasized a desire to annotate or provide context to the post. One patient said, "I need to be there with [my therapist] to review" (Participant #5, male, aged 34 years). Therapists highlighted a desire for objective metrics derived from social media posts. They commonly referenced social media's metadata such as time of post and language used. For example, therapists noted that patients recounting of events may be influenced by recall bias when individuals have a partial account of prior events [20], whereas seeing metadata provides an objective source of information. Both therapists and patients highlighted their comfort with and interest in digital data such as steps walked via a smartphone's built-in pedometer or screen time metrics.

Social media data are seen to aid discussions and accelerate a patient's account of events. Both participants and therapists highlighted how they would like to see trends over time. A therapist said, "I'm asking how they are feeling right now and how they feel how the last month has been to sort of see if their self-report in the moment lines up" (Participant #2, female, aged 35 years). Therapists also noted instances when their patient would show pictures of friends or family from a social media platform to add a "*face to the name*," implying that the use of such data builds rapport, which in turn could enhance the patient-provider communication.

Disadvantages of Incorporating Social Media in Therapy

Our results underscored perceived disadvantages of using social media data in therapy. Specifically, it can be nonreflective, nongeneralizable, and its use could be ethically ambiguous. Patients reported a sense of fear and uneasiness "*always*" sharing social media data with their therapist. They noted concerns about being "*watched*" and saw it as "*a little creepy*." Few therapists expressed concern that it could also elicit a Hawthorne effect, altering one's behavior due to the awareness of being observed [21]. As 1 therapist noted, "And again, me seeing all of [the] posts patients put up, even if they are agreeing to that…Patients may change [the way] they engage with that social media platform" (Participant #6, female, aged 43 years).

Therapists expressed concern that sharing social media data is nonreflective and would not provide accurate depictions of the patient's true thoughts and emotions. They highlighted how social media posts are often public-facing accounts of people or events and may not be genuinely authentic. One patient said that "*things are left"* out of posts, whereas therapists highlighted how their patients may have a social media personality.

Several participants raised important questions regarding the security of the data collected, with 1 patient saying, "I want to protect [my] autonomy" (Participant #7, female, aged 29 years). Similarly, a patient highlighted that that they trust their intuition on what social media posts to share with their therapist (Participant #7, female, aged 29 years). Half of the therapists expressed concern regarding the social media platforms' security policies. Several therapists also indicated that consistently including social media data in their therapy sessions could negatively impact their workflow.

Furthermore, utility of social media data in therapy may hinge on a patient's age or comfort with technology. Both therapists and patients agreed that including these data would be the most beneficial for younger or more technologically inclined patients. One patient participant noted that social media in their therapy session may not be "beneficial for me [and my treatment goals]

but for a younger generation because they post so much" (Participant #6, female, aged 60 years).

Additional Therapist Beliefs About Social Media Use in Therapy

After the interviews concluded, therapists were asked a series of open-ended survey questions to assess their normative and control beliefs relevant to using patients' social media data in therapy sessions. See Table 3 for the elicitation questions asked and the illustrative quotes. Therapists were asked to indicate what types of individuals or groups of people would be more or less likely to use social media in therapy (descriptive normative beliefs). Responses included being a younger therapist or patient and a digital native and someone who grew up with technology who might be more technologically inclined. Therapists indicated that their colleagues who would be open to this type of data exchange must be made aware of the social media platform themselves and be oriented toward CBT or other measurement-based care orientations. Therapists reported that individuals or groups who would be less likely to use social media in therapy included those who are older, have limited experience with social media platforms, and do not use measurement-based care. When asked to indicate what circumstances would make it difficult or easy to use social media in therapy (control beliefs), responses underscored the importance of convenience of use for both the therapist and the patient and ease of understanding in the context of therapy.

Table 3. Examples of open-ended responses and the corresponding theme categorization for each type of belief elicitation question.

Belief category, elicitation questions, illustrative open-ended responses	Subtheme(s)
Normative beliefs (injunctive)	
Generally, what types of individuals or groups would approve or think you should use data about clients' social media use in therapy sessions in the next month? Please list general groups or personas; do not include specific names.	
Younger clinicians, clinicians who work with young adults, clinicians who use a measurement-based care framework (ie, track their clients' progress using measures)	Convenience Objective
Younger therapists, data-driven/number-oriented people	Objective
Generally, what types of individuals or groups would disapprove or think you should not use data about clients' social media use in therapy sessions in the next month? Please list general groups or personas; do not include specific names.	
Potentially psychodynamic practitioners, individuals with strong privacy concerns	Ethically ambiguous
Individuals who do not use or have social media, clients who may be mistrusting or not have a strong therapeutic rapport with their therapist	Builds rapport Ethically ambiguous
Normative beliefs (descriptive)	
Generally, what types of individuals or groups are most likely to use data about clients' social media use in therapy sessions in the next month? Please list general groups or personas; do not include specific names.	
Younger, more number-oriented practitioners	Convenience Objective
Clinicians who are familiar with and comfortable using social media, clinicians who treat young adults, clinicians who incorporate technology into their treatments (eg, give measures on a computer or iPad, email, or text their clients)	Provider-initiated
Generally, what types of individuals or groups are least likely use data about clients' social media use in therapy sessions in the next month? Please list general groups or personas; do not include specific names.	
Older clinicians: clinicians in an environment in which it is inconvenient to do so	Convenience
Clinicians who do not use social media themselves and may have limited knowledge about how to use it (likely older clinicians), clinicians from orientations that do not emphasize measurement	N/A[a]
Control beliefs	
Please list any factors or circumstances that would make it easy or enable you to use data about clients' social media use in therapy sessions in the next month.	
Easy-to-use interface that logs all the information on the patient unobtrusively	User-friendly Convenience
Electronic platform, automated data collection and reminders	Convenience
Easily downloadable app(s), clear directions on how to use it in sessions with clients	Apps Convenience
Please list any factors or circumstances that would make it difficult or prevent you to use data about clients' social media use in therapy sessions in the next month.	
Lack of perceived need or benefit for a client, privacy concerns, difficult user interface for either myself or my client	Nonreflective
If I had to go through several steps to access the data	Convenience
A client's hesitancy or anxiety	Ethically ambiguous

[a]N/A: not applicable.

Discussion

Principal Findings

This qualitative study provides new knowledge on patient and therapist perspectives regarding the use of social media in mental health therapy. The interviews captured patients' and therapists' current use(s) of social media in mental health therapy and that both patients and therapists initiated its use and discussion in prior sessions. Of note, the individuals interviewed expressed comfort reviewing and discussing social media data. They identified several advantages and disadvantages of sharing social media data in mental health therapy (Table 2), underscoring the utility of and the potential concerns associated with integrating elements of our digital lives into mental health therapy in a world that increasingly relies on digital technologies.

To our knowledge, this is the first account of patients' experiences and perspectives on this novel data source. Patients highlighted how social media posts provide objective evidence

of their social lives or networks. They discussed previous experiences of self-curating posts and questioned if they want an algorithm selecting social media posts to share. Throughout the interviews, a clear understanding of how the data are generated and shared was of great importance. Previous research has captured individuals' willingness to share digital data for health research [3,22]. However, researchers have yet to devise a way to provide comprehensive feedback or snapshots to better inform patients, let alone their health care provider.

Despite already using social media data in therapy, both patients and therapists questioned the utility of certain social media platforms. They noted that social media posts may not fully reflect their true thoughts or feelings. This speaks to a growing trend among Instagram users with 2 types of accounts: "finsta" accounts, on which users post less polished photos of themselves and "rinsta" accounts, which include more authentic posts [23]. "Finsta" and "rinsta" accounts highlight how social media posts may not always accurately portray individuals' experiences. However, there might be clinical relevance if a patient has a "rinsta" account.

The desire for objectivity was noted throughout the interviews. Although the interview questions specifically asked about social media data, therapists discussed the incorporation of objective digital data such as the number of steps taken per day or hours slept per night and highlighted them as variables that can contribute to one's mental health and well-being. Although our interviews did not specifically capture the use of digital data such as steps and screen time, Di Matteo and colleagues [24] reported general acceptance sharing this type of data through their interviews with new patients referred to a tertiary care mood and anxiety disorder clinic. Since Di Matteo et al's work [24] and our interviews were both hypothetical scenarios, further research is needed to assess patient and therapist comfort with sharing social media data in the context of therapy sessions.

Lastly, when discussing how to systematically incorporate social media data into sessions, both patients and therapists stressed the importance of person-centered design. As demonstrated by Yoo and colleagues [25], who conducted co-design workshops to build a social media tool for therapists, the data must map onto the end users' expectations, both in terms of analyses and organization, to avoid being "another layer of noise" [25]. Therapists in this sample questioned how social media data would fit in their workflow (Tables 2 and 3). Despite these concerns, overall, our interviews highlight how sharing social media insights with the patient and therapist could also be clinically relevant and informative. Further work is needed to explore and test how to systematically collect social media and present it to patients and their therapists in a user-friendly format for use in mental health therapy.

Implications and Recommendations

To our knowledge, this is the first study to capture the patient perspective on their experience sharing and discussing social media data in mental health therapy. The patients interviewed provided critical insights that have yet to be characterized. Our findings uniquely underscore the importance of patient autonomy on what and when to share social media data. As the study sample included mostly female, White, college-educated patients and female, White, and CBT therapists, future research should be conducted with more diverse samples in terms of gender, race/ethnicity, therapeutic orientation, and educational attainment. Additional and special attention is also needed to explore how social media is used and discussed in other cultural contexts.

Our interviews were conducted in 2018 and 2019 prior to the COVID-19 pandemic and may not reflect current norms and beliefs of social media data sharing in the therapeutic encounter. During the COVID-19 pandemic and recovery phrases, there has been a dramatic increase in individuals with depressive and anxiety symptoms seeking mental health care [26]. Further research is warranted to capture how social media data are shared in virtual sessions via videoconference or telephone call. We also recommend additional research on the use of digital data such as smartphone metadata as viable data sources in therapy [27]. The combination of social media and digital data could enhance tailored treatment plans and impact therapeutic alliance, a cooperative working relationship between client and therapist, often seen as an essential aspect of successful therapy [28]. Since previous research found that therapeutic alliance is maintained and even enhanced with the introduction of digital mental health interventions [29,30], further research in this area is needed.

Lastly, there are clear educational, practical, and policy implications. As detailed in the American Psychological Association Guidelines for the Optimal Use of Social Media in Professional Psychological Practice (October 2021) [31] and indicated by our findings, mental health therapists should be encouraged to undergo specific educational training on how to safely and ethically use social media data in therapy. The training should encompass core elements of ethics, informed consent, comfort with technology, social media trends per population segment (eg, age, gender, race, ethnicity, sexual identity, language, culture), and how to integrate these new data sources into their workflow. Clinics and practices would need to reserve additional time, personnel, and technology infrastructure to support this training. Noel and colleagues [11] recommend a technology specialist, a new type of health care worker who identifies and reviews electronic resources that may support a client's specific recovery goals. Data infrastructure and protections are critically important and require special attention. Furthermore, if social media data were incorporated into therapy and demonstrated to improve patient outcomes and reduce costs to the clinic and patient, national/state policies and insurance companies could modify current plans and coverage.

Limitations

This study has several limitations. First, the study sample was small and largely homogeneous with respect to sociodemographic data. All participants were recruited from a convenience sample in 1 large metropolitan region in northeastern United States. It is possible that the results from this study do not apply to other population segments or geographic regions. As we advertised for this study online, it is possible that study participants were drawn to the study because of their prior experience with social media in mental health therapy. We were unable to audio record patient

participant interviews. As such, our transcripts were not as robust for patient participants as they were for therapist participants. Our interviews were conducted with patients and therapists separately; future research could interview patient and therapist dyads for further insights. With respect to analyses, there are limitations to thematic coding as a methodological approach, such as inferences made from a small study sample size and coding at the phrase level, which may not fully capture the participants' intended meaning. Furthermore, we were unable to use qualitative data analysis software such as NVivo. Interviews were conducted prior to the COVID-19 pandemic. Accordingly, findings may not fully reflect the current state of mental health delivery in the United States, as most mental health therapy is now delivered virtually via videoconference [32]. Despite these limitations, findings from this pilot study can inform social media use practices and norms in mental health therapy.

Conclusions

In this study, patient and therapist interviews provide important insights on the current utilization of social media in mental health therapy, including the advantages and disadvantages of social media use in such contexts. Our findings highlight that social media data used in therapy provide convenient, objective information that is user-friendly and can promote rapport between the patient and the therapist. However, the use of social media data in mental health therapy is also perceived as nonreflective, ethically ambiguous, and potentially nongeneralizable. Future research is needed to explore and test how to systematically collect social media and present it to patients and their therapists in a user-friendly format for use in mental health therapy.

Acknowledgments

We would like to acknowledge and thank Andrea Ngan and Jessica Diaz who assisted with the interviews and interview transcriptions.

Conflicts of Interest

None declared.

References

1. Social media use in 2021. Pew Research Center. 2021 Apr 07. URL: https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/ [accessed 2021-04-13]
2. Social media fact sheet. Pew Research Center. 2021 Apr 07. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/ [accessed 2021-04-13]
3. Asch DA, Rader DJ, Merchant RM. Mining the social mediome. Trends Mol Med 2015 Sep;21(9):528-529 [FREE Full text] [doi: 10.1016/j.molmed.2015.06.004] [Medline: 26341614]
4. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nat Biotechnol 2015 May;33(5):462-463. [doi: 10.1038/nbt.3223] [Medline: 25965751]
5. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. N Engl J Med 2016 Sep 29;375(13):1216-1219. [doi: 10.1056/nejmp1606181]
6. Merchant RM, Asch DA, Crutchley P, Ungar LH, Guntuku SC, Eichstaedt JC, et al. Evaluating the predictability of medical conditions from social media posts. PLoS One 2019;14(6):e0215476 [FREE Full text] [doi: 10.1371/journal.pone.0215476] [Medline: 31206534]
7. Mohr DC, Zhang M, Schueller SM. Personal sensing: understanding mental health using ubiquitous sensors and machine learning. Annu Rev Clin Psychol 2017 May 08;13:23-47 [FREE Full text] [doi: 10.1146/annurev-clinpsy-032816-044949] [Medline: 28375728]
8. De Choudhury M, Gamon M, Counts S. Predicting depression via social media. Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media. 2013. URL: https://tinyurl.com/2p98c3ke [accessed 2022-06-15]
9. De Choudhury M, Kiciman E, Dredze M. Discovering shifts to suicidal ideation from mental health content in social media. 2016 Presented at: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems; May; San Jose, California, USA p. 2098-2119. [doi: 10.1145/2858036.2858207]
10. Guntuku SC, Schneider R, Pelullo A, Young J, Wong V, Ungar L, et al. Studying expressions of loneliness in individuals using twitter: an observational study. BMJ Open 2019 Nov 04;9(11):e030355 [FREE Full text] [doi: 10.1136/bmjopen-2019-030355] [Medline: 31685502]
11. Noel VA, Carpenter-Song E, Acquilano SC, Torous J, Drake RE. The technology specialist: a 21st century support role in clinical care. NPJ Digit Med 2019;2:61 [FREE Full text] [doi: 10.1038/s41746-019-0137-6] [Medline: 31388565]
12. Eichstaedt JC, Smith RJ, Merchant RM, Ungar LH, Crutchley P, Preoţiuc-Pietro D, et al. Facebook language predicts depression in medical records. Proc Natl Acad Sci U S A 2018 Oct 30;115(44):11203-11208 [FREE Full text] [doi: 10.1073/pnas.1802331115] [Medline: 30322910]

13. Hobbs KW, Monette PJ, Owoyemi P, Beard C, Rauch SL, Ressler KJ, et al. Incorporating information from electronic and social media into psychiatric and psychotherapeutic patient care: survey among clinicians. J Med Internet Res 2019 Jul 12;21(7):e13218 [FREE Full text] [doi: 10.2196/13218] [Medline: 31301127]
14. Carson NJ, Gansner M, Khang J. Assessment of digital media use in the adolescent psychiatric evaluation. Child Adolesc Psychiatr Clin N Am 2018 Apr;27(2):133-143. [doi: 10.1016/j.chc.2017.11.003] [Medline: 29502741]
15. Fisher CE, Appelbaum PS. Beyond googling: the ethics of using patients' electronic footprints in psychiatric practice. Harv Rev Psychiatry 2017;25(4):170-179. [doi: 10.1097/hrp.0000000000000145]
16. Sundram F, Hawken SJ, Stasiak K, Lucassen MF, Fleming T, Shepherd M, et al. Tips and traps: lessons from codesigning a clinician e-Monitoring tool for computerized cognitive behavioral therapy. JMIR Ment Health 2017 Jan 11;4(1):e3 [FREE Full text] [doi: 10.2196/mental.5878] [Medline: 28077345]
17. Xanidis N, Brignell CM. The association between the use of social network sites, sleep quality and cognitive function during the day. Computers in Human Behavior 2016 Feb;55:121-126. [doi: 10.1016/j.chb.2015.09.004]
18. Fishbein M, Ajzen I. Predicting and changing behavior: The reasoned action approach. APA PsychNet. URL: https://psycnet.apa.org/record/2009-17267-000 [accessed 2021-04-18]
19. Srivastava A, Thomson SB. Framework analysis: a qualitative methodology for applied policy research. JOAAG. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2760705 [accessed 2022-06-15]
20. Patten SB. Recall bias and major depression lifetime prevalence. Soc Psychiatry Psychiatr Epidemiol 2003 Jun;38(6):290-296. [doi: 10.1007/s00127-003-0649-9] [Medline: 12799778]
21. Sedgwick P, Greenwood N. Understanding the hawthorne effect. BMJ 2015:h4672. [doi: 10.1136/bmj.h4672]
22. Seltzer E, Goldshear J, Guntuku SC, Grande D, Asch DA, Klinger EV, et al. Patients' willingness to share digital health and non-health data for research: a cross-sectional study. BMC Med Inform Decis Mak 2019 Aug 08;19(1):157 [FREE Full text] [doi: 10.1186/s12911-019-0886-9] [Medline: 31395102]
23. Bach MC. Attachments: the colliding worlds of telepsychiatry, social media, and COVID-19. J Am Acad Child Adolesc Psychiatry 2021 Apr;60(4):426-428. [doi: 10.1016/j.jaac.2020.10.019] [Medline: 33188855]
24. Di Matteo D, Fine A, Fotinos K, Rose J, Katzman M. Patient willingness to consent to mobile phone data collection for mental health apps: structured questionnaire. JMIR Ment Health 2018 Aug 29;5(3):e56 [FREE Full text] [doi: 10.2196/mental.9539] [Medline: 30158102]
25. Yoo DW, Birnbaum ML, Van Meter AR, Ali AF, Arenare E, Abowd GD, et al. Designing a clinician-facing tool for using insights from patients' social media activity: iterative co-design Approach. JMIR Ment Health 2020 Aug 12;7(8):e16969 [FREE Full text] [doi: 10.2196/16969] [Medline: 32784180]
26. Vahratian A, Blumberg SJ, Terlizzi EP, Schiller JS. Symptoms of anxiety or depressive disorder and use of mental health care among adults during the COVID-19 pandemic - United States, August 2020-February 2021. MMWR Morb Mortal Wkly Rep 2021 Apr 02;70(13):490-494 [FREE Full text] [doi: 10.15585/mmwr.mm7013e2] [Medline: 33793459]
27. Moshe I, Terhorst Y, Opoku Asare K, Sander LB, Ferreira D, Baumeister H, et al. Predicting symptoms of depression and anxiety using smartphone and wearable data. Front Psychiatry 2021;12:625247 [FREE Full text] [doi: 10.3389/fpsyt.2021.625247] [Medline: 33584388]
28. APA Dictionary of Psychology. URL: https://dictionary.apa.org/therapeutic-alliance [accessed 2022-03-29]
29. Tremain H, McEnery C, Fletcher K, Murray G. The therapeutic alliance in digital mental health interventions for serious mental illnesses: narrative review. JMIR Ment Health 2020 Aug 07;7(8):e17204 [FREE Full text] [doi: 10.2196/17204] [Medline: 32763881]
30. 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]
31. APA guidelines for the optimal use of social media in professional psychological practice. APA. URL: https://www.apa.org/about/policy/guidelines-optimal-use-social-media.pdf [accessed 2022-05-14]
32. Using telehealth to expand access to essential health services during the COVID-19 pandemic. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/hcp/telehealth.html [accessed 2021-07-21]

Abbreviations

CBT: cognitive behavioral therapy
SMUQ: social media use questionnaire

Edited by A Mavragani; submitted 14.09.21; peer-reviewed by A Sachinopoulou, M Kleib, H Tanaka; comments to author 14.02.22; revised version received 13.05.22; accepted 27.05.22; published 07.07.22.

Please cite as:
Southwick L, Suh R, Kranzler E, Bradley M, Merchant RM
Perspectives of Patients and Therapists on Social Media and Digital Data Use in Mental Health Therapy: Thematic Analysis
JMIR Form Res 2022;6(7):e32103
URL: https://formative.jmir.org/2022/7/e32103
doi:10.2196/32103
PMID:35797103

Original Paper

A German Smartphone-Based Self-management Tool for Psoriasis: Community-Driven Development and Evaluation of Quality-of-Life Effects

Lea C Brandl[1,2], MSc; Claudia Liebram[3]; Wendelin Schramm[4], Prof Dr med; Monika Pobiruchin[2,4], Dr sc hum

[1]Institute of Telematics, University of Lübeck, Lübeck, Germany

[2]Consumer Health Informatics SIG, German Association for Medical Informatics, Biometry & Epidemiology (GMDS e.V.), Cologne, Germany

[3]Psoriasis-Netz, Berlin, Germany

[4]GECKO Institute for Medicine, Informatics & Economics, Heilbronn University, Heilbronn, Germany

Corresponding Author:
Lea C Brandl, MSc
Institute of Telematics
University of Lübeck
Ratzeburger Allee 160
Lübeck, 23562
Germany
Phone: 49 451 3101 6401
Email: brandl@itm.uni-luebeck.de

Abstract

Background: Psoriasis is a chronic disease characterized by inflammation, increased scaling, itching, and other symptoms. Psoriasis is not contagious, but patients have often felt shunned. Therefore, in addition to psoriasis symptoms, stress, anxiety, and depression can also affect quality of life (QoL). Surveys show that only a quarter of patients are satisfied with the success of their therapy. However, in addition to medical therapy, self-management can also make it easier to deal with chronic diseases like psoriasis.

Objective: The aim of this project was to develop a smartphone-based self-management tool (SMT) specifically for patients with psoriasis using a community-driven process. The impact of the SMT on QoL as well as its acceptance and usability were evaluated.

Methods: In collaboration with an internet-based self-help community, 2 user surveys were conducted to determine the requirements for a smartphone-based SMT. The surveys consisted of semistructured questionnaires asking for desired features in an SMT for psoriasis. A pilot study was conducted to evaluate QoL, acceptance, and usability. Community users were recruited to use the app for 21 days and complete the Dermatology Life Quality Index (DLQI) questionnaire at the beginning (T_0) and end (T_1). Afterward, participants were asked to complete another questionnaire on usability and ease of use.

Results: SMT requirements were collected from 97 members of an internet-based community. The SMT was built as a progressive web app that communicates with a server back end and an Angular web app for content management. The app was used by 15 participants who also provided qualitative feedback, and 10 participants answered all questionnaires. The average DLQI score was 7.1 (SD 6.2) at T_0 and 6.9 (SD 6.6) at T_1. The minimal required sample size of 27 was not reached.

Conclusions: The high degree of community participation in the development process and the responses during the requirement engineering process indicated that there is a general need for an independently developed SMT for patients with psoriasis. However, the feedback received after app use shows that the SMT does not meet the needs of the community. It can be concluded that a more customizable app is needed. The focus and needs of the users were very heterogeneous. Similar developments and research could benefit from the findings of this project.

(JMIR Form Res 2022;6(7):e32593) doi:10.2196/32593

KEYWORDS

psoriasis; self-management; mobile apps; quality of life; mobile phones; smartphones

Introduction

Overview

Psoriasis is a chronic skin disease that manifests itself with inflammation, increased scaling, itching, and other symptoms [1]. Flares can be aggravated by stress, medical drugs, or infectious diseases [2-4]. Psoriasis affects approximately 2% to 3% of the German population [1,5]. In addition to psoriatic symptoms, stress, anxiety, and depression can affect quality of life (QoL) [6-8]. Compared to healthy individuals, anxiety disorders were identified more frequently in patients with psoriasis [4,9]. It is not possible to cure psoriasis, but symptoms could be mitigated with medical drugs, special therapies, or complementary methods [10,11]. Nevertheless, approximately 30% of patients do not consult a physician [12,13].

Interviews reveal that a quarter of patients are dissatisfied with the success of their treatment [14,15]. Problems regarding drug therapies are present in up to 40% of patients [15,16]. To increase patient satisfaction and adherence, evidence-based decision guidance for psoriasis therapy is available as an S3 guideline [17]. However, in addition to medical therapy, self-management can facilitate the management of chronic diseases such as psoriasis [18-20]. In comparison to other chronic diseases, there are less evidence-based self-management tools (SMTs) for patients with psoriasis [20].

Related Work

Self-help and self-management can be effective tools for patients to better cope with their (chronic) disease [18-20]. Traditionally, self-help involves in-person group meetings or counseling. In recent years, a shift toward mobile apps, information websites, and web-based communities can be observed. Self-help can also include aspects of citizen science approaches. For example, self-help groups and other organizations collect and organize information on people with a disease [21]. Internet-based communities are often used by patient experts who share information and generate hypotheses. This is done based on shared experiences. Modern technologies can facilitate new approaches to citizen science projects [22].

Citizen science also includes the involvement of patients in the development process of SMTs. Some examples are given in the following paragraphs. Safdari et al [23] conducted a study on the requirements for a self-management app for patients with psoriasis. In a requirements analysis, 100 patients provided information about their requirements for educational information and lifestyle management, among other features. This included information on the disease on the one hand and factors such as physical activity, nutrition, or stress management on the other hand.

Trettin et al [24,25] have developed an app for Danish patients with psoriasis treated with biologics. In the development process, patients were interviewed, and various workshops and a prototype test were carried out. In the app, vital signs, and the Dermatology Life Quality Index (DLQI) [26] can be recorded in preparation for video or telephone consultations. Patients, doctors, and nurses have reported that consultations are more structured, and patients feel safe [24,25].

Aims of the Study

Currently, there are few SMTs in German app stores and no German-speaking apps (co-) developed by self-help organizations (ie, the patients themselves). Therefore, the 2 major objectives of this study are to (1) present an SMT especially developed by and for patients with psoriasis and (2) evaluate the app's impact on QoL as well as its acceptance and usability.

Methods

Recruitment and Requirement Engineering Phase

In Germany, there are 2 main organizations for psoriasis self-help. The larger one is the Deutscher Psoriasis Bund e.V. (German Psoriasis Association) and the smaller one is the Psoriasis Selbsthilfe Arbeitsgemeinschaft e.V. (Psoriasis Self-Help Association, PSOAG) [27,28]. The latter offers self-help on its website. The internet-based community consists of approximately 28,000 users (as of November 2019) and concerns itself with topics such as therapies or nutrition. There are various expert forums and groups that can be joined. Knowledge articles are also published [29]. The project was initiated with PSOAG members. The community collected requirements of a (potential) smartphone-based SMT during 2 online surveys (published via the web-based forum).

The first survey started on February 2, 2019, and was closed on March 15, 2019. It consisted of a semistructured questionnaire that asked about the desired features of an SMT for psoriasis. A translated version of this questionnaire can be found in Multimedia Appendix 1. After completion of the first survey, the responses were read and clustered to capture different categories of functionalities. In the next step, categories were checked for technical feasibility and focus on SMTs.

The second survey that incorporated the community feedback from the first round was open from June 25, 2019, to August 31, 2019. During this phase, 11 mock-ups of the SMT were presented and discussed with the community members. The mock-ups were built with the Balsamiq Mockup software (version 3.5.17; Balsamiq Studios, LLC) [30].

Changes needed in the design of the app and the expansion of certain functionalities were again checked for technical feasibility and incorporated into the final app requirements.

After the requirements were determined, the SMT development started. The system was implemented by one of the authors using software components previously developed [31] as blueprints. The app development and the pilot study were carried out as part of a master's thesis at Heilbronn University.

SMT Evaluation

QoL Assessment

Numerous instruments of QoL assessment exist [32,33]. Fitzpatrick et al [34] developed a list of criteria to select the most appropriate measurement instrument. According to these criteria, we decided to use the DLQI by Finlay et al [26] for this study. The DLQI consists of 10 questions that cover the dimensions of symptoms and daily activity, leisure, work or

school, personal relationships, and therapy [26]. The average response time is 2 minutes [35]. The score of the DLQI ranges from 0 to 30. The larger the value, the worse the QoL [26]. The minimal clinically important difference is 4 score points [36]. A license for the use of the German Translation DLQI was applied for and was provided by Cardiff University [35].

Use and Usability

Considering usability and further use, an additional questionnaire was introduced. It included questions about handling, use, further use, and the severity of psoriasis, and were self-reported and/or measured by the Psoriasis Area and Severity Index (PASI) [37]. This additional questionnaire can be found in Multimedia Appendix 2.

Recruitment for QoL Assessment

Participants for the QoL assessment were recruited through various channels:

1. A call for participation was given via the PSOAG website and forum. The call provided the opportunity to contact the study team directly or to access a study information website.
2. Information flyers were sent to dermatology clinics (13 rehabilitation clinics and 11 university clinics), the Professional Association of German Dermatologists, and the Psoriasis Association.
3. Social media channels were used to actively promote the study (eg, regular tweets, posts on Instagram, and the creation of a Facebook page about the study).

For inclusion in the study, the participants had to have psoriasis and be at least 18 years old.

Study Design

As the study measured the change in QoL, it was necessary to record it at a minimum of 2 time points (T_0 and T_1). The DLQI measures QoL in relation to the last 7 days. A break of at least 7 days is recommended, and frequent questioning is discouraged; otherwise, participants may remember their previous response. Based on these requirements, the intervention duration was set at 21 days.

After 21 days had elapsed, the second DLQI and the extra questionnaire on usability and further use of the SMT were completed. Then, the participants dropped out of the study.

Statistical Analysis

Based on the German-specific data provided by Lesner et al [7] and the minimal clinical effect of 4 points, a sample size of 27 participants is calculated to reach α=.05 and β=.2 [38]. Data were analyzed with the statistical software MATLAB (version 2019b) [39].

Ethical Approval

No ethical approval was obtained for the conducted study. The local ethics committee of Heilbronn University is only an advisory board (see §1(1) [40]), and it does not provide formal ethical approval. This was a community-led project involving interested persons with psoriasis who enrolled voluntarily and by general invitation in the development and evaluation of an app. These were not patients treated by the authors. Treatment changes or any interaction with the medical care delivery team were not included in the objectives of this study.

As the data collected were worthy of protection, a high scientific standard was applied here. All participants were informed about the study details and received written information in accordance with the Declaration of Helsinki [41]. This information was available on the study information website and could be downloaded for offline reading (see Multimedia Appendix 3). All participants had one-to-one contact with the principal investigator and questions from them were answered by email or telephone, as requested.

All participants signed an informed consent form (see Multimedia Appendix 4), which could be revoked at any time. In this case, all documents that could still be assigned were destroyed. To ensure a high level of data security, the questionnaires were recorded under a pseudonym generated by the participants themselves.

Results

Requirement Engineering and SMT

More than 90 community members participated in the survey (N=97). They determined the scope of the SMT's features, created the medical content, and curated it.

During the first user survey, the requirements for the SMT were collected. Nearly 90% (87/97) of the participants indicated willingness to test the SMT. The suggestions of the community about the desired features were evaluated for their feasibility and purpose of use. Four categories emerged during the analysis of the community responses: (1) communication, (2) drug management, (3) tracking of complementary methods, and (4) rate/score doctors. Selected examples of the feedback from the SMT users are presented in Table 1. Furthermore, mock-ups (ie, drafts of the user interface) were created. These were presented to the community and the community members could provide their comments.

The final SMT framework consists of three components, as shown in Figure 1: (1) an Angular web application for the management of the SMT content used by the editorial team, (2) a server back end built with Java Spring Boot and Spring Security, and (3) the SMT for the study participants as an Angular application. The SMT app was implemented as a progressive web app (PWA) to be available for all types of smartphones.

Both web components were built with the Angular Framework (Google; version 7.2.15) [42]. The server back end was built with Java Spring Boot (The Spring Team; version 2.1.5) [43] and Java (Oracle Corporation; version 11) [44].

One feature focus of the SMT is to suggest interaction-free complementary measures to the user, with which physical complaints, such as itching, skin blisters, and dry skin, can be alleviated. Patients can document their psoriasis type, medical drugs, and complaints each day. Based on these data, the SMT suggests complementary methods that are free of interactions. The complementary measures were suggested by the community itself and were reviewed by the forum's editorial team. In total,

55 complementary measures were found, described, and incorporated into the knowledge base of the SMT (see component 1 in Figure 1). The complete list is available in Multimedia Appendix 5. The knowledge was embedded in a PostgreSQL database (PostgreSQL Global Development Group; version 10.9) [45], as shown in Figure 1.

Another aspect of the app was the documentation of psoriasis and the possibility to view the severity of symptoms as they progressed and in relation to the complementary methods that were considered. Moreover, individual body parts could be documented in more detail with the help of photos and text. Figure 2 shows representative screenshots.

A comprehensive data protection concept was developed. All the app data were stored exclusively on the user's own smartphone.

Table 1. Results of the first requirement analysis involving the web-based community[a].

Feedback from the community	Category	Realization in the SMT[b]
"Input of laboratory values and their graphical representation in order to be able to observe their development (eg, CRP[c], leukocytes, lymphocytes)"	2	—[d]
"Alarm clock to remind of injection days, medication intake, doctor's appointments"	2	—
"One should be able to print out what has been written"	3	✓[e]
"Being up to date on care for scalp psoriasis - nutrition tips, new findings"	3	+[f]
"Diet plan, natural remedies"	3	+
"Exchange with other patients (ie, link to the forum)"	1	—
"Notifications about new findings regarding the therapy used (side effects, new variants)"	2	—
"All [measures] that improve my quality of life, take away pain, and build me up"	3	+
"It could remind me to take medication"	2	—
"Study situation and results on complementary therapy methods would be good"	3	✓
"Some kind of daily checklist to check off to-dos related to (psoriasis] would be great. That way you can see correlations, if necessary (eg, if you forgot a supplement or only applied cream once a day instead of twice)."	3	+
"Could you name and comment on current dermatologist (recommend, yes/no and why)"	4	—
"That you can print out the diary would be good to present to the doctor who treated you!"	3	✓
"A list of my current medications as well as a list of medications I have tried but no longer use (and why!)"	2, 3	+
"A chat or a link to a forum where users can exchange information directly. Because I see the main problem with the app as being that something different helps everyone. It's so hard to generalize."	1	+
"Calendar in which one can enter appointments for the doctor, taking one's medication or other things."	1, 2	—
"Logging what you have eaten and then automated evaluation of whether patterns are recognizable, how nutrition affects you. The same with stress/well-being."	3	✓
"Maybe refer to competent doctors, dermatologists and rheumatologists in the respective federal states of the people concerned."	4	—
"To find other patients with psoriasis nearby and find experts or good doctors from the patient's point of view."	1, 4	—
"A digital (medication plan) for all medications, not only for psoriasis treatment; in my case, for example, this was accompanied by a CHD[g] disease."	2	—

[a]All statements were translated from German. Additions or replacements for better understanding are denoted in square brackets.

[b]SMT: self-management tool.

[c]CRP: c-reactive protein.

[d]The function is not implemented.

[e]The function is fully implemented.

[f]The function is partially implemented.

[g]CHD: congenital heart defects.

Figure 1. System architecture. The 2 Angular components are shown on the left. These are provided on a web server that communicates with the back-end server. A Spring Security module protects the back end and the connected database from security attacks. The back-end component is shown on the right. DB: database; SMT: self-management tool; SQL: structured query language.

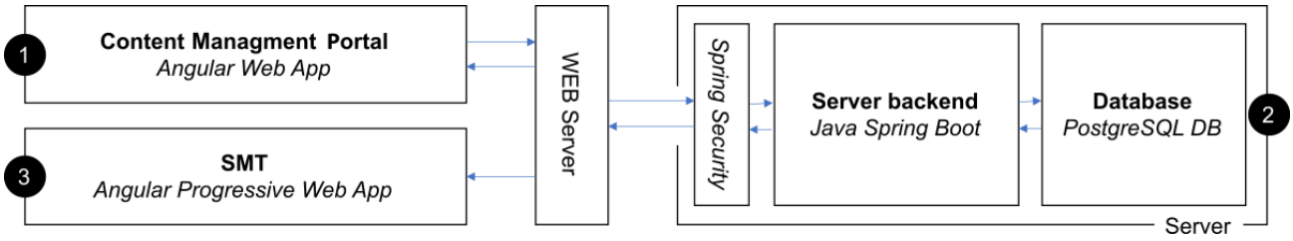
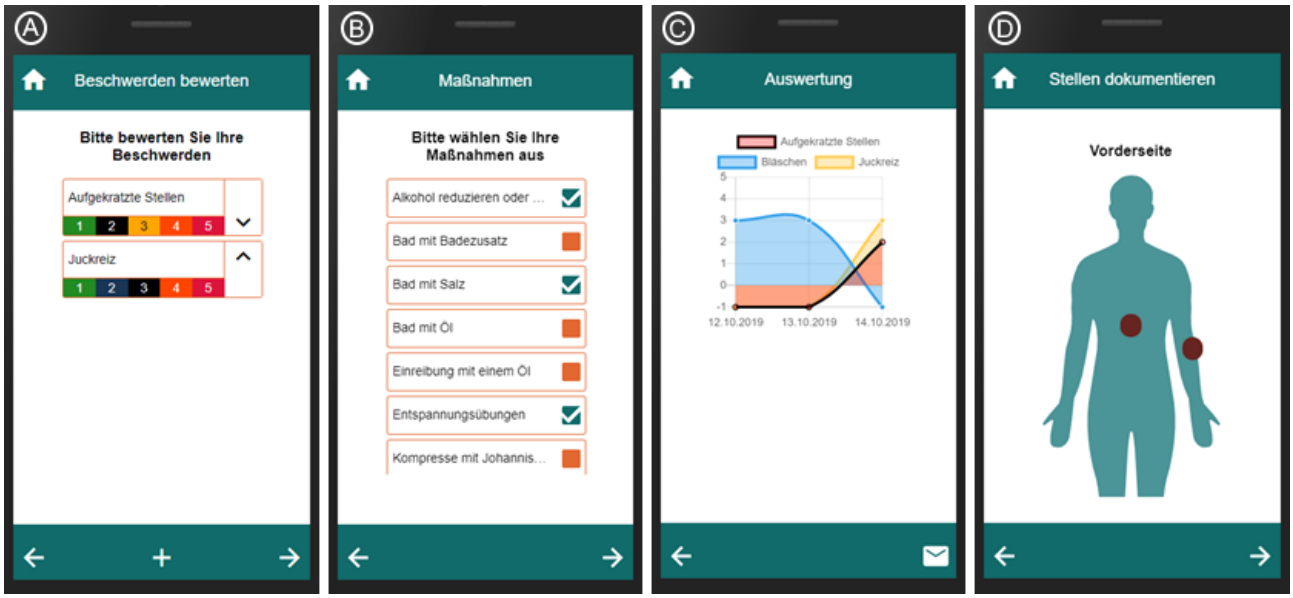


Figure 2. Screenshots of the smartphone app. Left to right: (A) Rating and sorting physical complaints. Complaints can be added via the "plus" sign in the footer. (B) Choosing complementary methods, including reducing alcohol, taking a bath with bath salts, and relaxing. (C) The course of complaints. By clicking on one of the dots in the chart, the complementary methods used during this day are shown. (D) Documentation of the front of the body. A click on the body adds a region of interest. Photos and text could be documented for each region of interest.



SMT Evaluation

QoL Analysis

During the call for participation, 29 persons showed interested in participation, with 21 of them providing consent to the processing of their data. Of these, 18 submitted the initial questionnaires. All persons who returned the first questionnaire are counted as participants. Participants who revoked consent (3 participants) or did not return the second questionnaire (5 participants) are counted as dropouts. Only 10 participants returned both questionnaires that were subsequently analyzed.

In total, 15 valid preintervention and 10 postintervention QoL questionnaires, and questionnaires about acceptance and usability were collected and transcribed into a CSV format. Additionally, 6 participants did not return any questionnaires but gave consent.

The mean DLQI for the preintervention questionnaire was 7.1 (SD 6.2) score points and 6.9 (SD 6.6) score points for the postintervention questionnaire. A lower score after the intervention corresponds to an increased QoL. However, the reduction of the DLQI by 0.2 score points does not correspond to a clinically relevant effect.

Use and Usability

The survey on difficulties in using the app shows that users did face challenges (see Table 2).

The results showed that 4 participants had no difficulties in using the app, and 3 participants found it quite or very difficult to use. Moreover, 3 other participants had slight difficulty in using the app. This was also reflected in the feedback from the participants. When returning the second questionnaire or withdrawing from the study, there was additional qualitative feedback on the SMT given by various participants, which is documented in Textbox 1.

The participants were also asked about their continued use of the app. None of the participants would continue to use the app daily. Only 2 participants continued to use the app regularly, 1 participant used it as needed, and 6 participants did not continue to use the app at all. In 1 questionnaire, this question remained unanswered.

Table 2. Study results[a].

Participant	$DLQI_{T0}$[b]	$DLQI_{T1}$[c]	Usage period (days)[d]	Difficulty[e]	Further use[f]	Severity level[g]	PASI[h] score
1	1	1	3	A little	Not at all	Mild	Not replied
2	5	3	12	Fairly	Not replied	Heavy	Not replied
3	1	0	4	A lot	Not at all	Heavy	Not replied
4	9	19	6	A little	Not at all	Mild	0
5	7	2	21	Not at all	Regularly	Mild	Not replied
6	2	2	4	Fairly	Not at all	Moderate	Not replied
7	1	4	8	Not at all	Not at all	Moderate	<3
8	20	17	5	A little	As needed	Moderate	Not replied
9	15	9	5	Not at all	Regularly	Heavy	Not replied
10	10	12	3	Not at all	Not at all	Moderate	Not replied
11	15	__[i]	—	—	—	—	—
12	15	—	—	—	—	—	—
13	4	—	—	—	—	—	—
14	10	—	—	—	—	—	—
15	10	—	—	—	—	—	—

[a]Average values for participants 1 to 10: $DLQI_{T0}$, 7.1, $DLQI_{T1}$, 6.9, and usage period (days), 7.1. Average $DLQI_{T0}$ value for participants 1 to 15 was 8.3.

[b]$DLQI_{T0}$: Dermatology Life Quality Index before using the app.

[c]$DLQI_{T1}$: Dermatology Life Quality Index after using the app.

[d]Indicates the number of days the app was used.

[e]Represents the difficulties experienced when using the app.

[f]Represents further use of the app.

[g]Shows the subjective perception of the severity of psoriasis.

[h]PASI: Psoriasis Area and Severity Index.

[i]Indicates participants who only completed the first questionnaire but did not withdraw their agreement.

Textbox 1. Qualitative feedback of participants. All statements were translated from German. Additions for better understanding are shown in square brackets. Detailed information on date, time, and medication are omitted because of privacy reasons.

Person A

- "I have psoriatic arthritis and unfortunately the app does not relate to that."
- "And unfortunately, I also have to say that I find the handling very complicated."
- "Basically, I find the idea good, but so [the app] is unfortunately not further usable for me."

Person B

- "It worked great."
- "The "problem" is the photos. I live alone and find it difficult to photograph the back. But that is my only issue at the moment."

Person C

- "I have tried a few apps, but they were all not applicable. Maybe theirs is a little better and more supportive."

Person D

- "Also, I found it inconvenient and difficult to use at times because there was no explanation whatsoever."

Person E

- "The "app" is not a native iOS or Android app, but a web application. I prefer here, especially for my sensitive health data, an app running locally on the iPhone/iPad, where the data is local on the device or encrypted in my own or the iCloud (optional)."
- "Using the app in parallel from multiple devices (smartphone, tablet, smartwatch) would be of great benefit."
- "Such an app would also have to be individually "customizable" for me: Creation of own "therapies". For the individual therapies, storage of more details (eg, for light therapy, the duration of the respective irradiation or the set Joule dose), for ointments, the name, PZN, etc."
- "Usability would also have to be significantly optimized. For regular documentation (skin condition, condition, medication), the [documentation] must happen as quickly and easily as possible."

Person F

- "The app is very interesting and helpful for people who don't have this background knowledge."
- "The documentation via photos, I think is very good, I could have used that a lot from ***."

Person G

- "I only used the app for a very short time, as I perceived filling it out as annoying."
- "However, I think small changes to the app could fix this for the most part."

Person H

- "Since I don't do therapy other than *** and *** and don't yet know what things help me, it would have also become difficult for me to use it meaningfully."
- "I'm just looking more the other way around for a template where I can document what I've done and eaten (how much sleep, how much sun, etc) to figure out what factors are negatively impacting me."

Statistical Analysis

The targeted sample size (N=27) was not achieved. Therefore, the testing of the hypotheses was disregarded.

Discussion

Requirements Engineering and SMT Development

A special feature of the development process was the high level of participation. Community members were asked about additional features and could comment on the designs and mock-ups. Unfortunately, it was not possible to implement all the desired functions (see Table 1). Reminders could not be implemented due to the technical restriction of a PWA. Notifications are possible with this design but cannot be individually configured. Ratings of doctors and exchanges with other patients were not implemented, as these functions would overlap with the internet-based forum of the community. Some suggestions of the stakeholders have been partially implemented because the documentation should be in a structured form instead of a free form to avoid incorrect entries and simplify usage.

When examining for possible confounders, the usability aspects stood out. This was reinforced by the feedback from the individual participants. There were difficulties in dealing with

the app. This was partly due to the handling and partly because patients have a greater need to document their illness than that assumed. In particular, the behavior for which therapy is currently being provided and secondary diseases should be documented more precisely and as quickly as possible. The heterogeneity of the users' feedback shows that the app should be highly customizable. The suggestion to implement the app as a standalone one rather than as a PWA certainly warrants further IT security requirements and offers possibilities for including more features.

The feedback from the community shows a certain ambivalence regarding the privacy aspect. On the one hand, previous experience and feedback on the app show that a high level of data protection is desired. Therefore, all data remain locally stored on the user's device. On the other hand, the fact that the data cannot be accessed on several devices was perceived as a negative feature. Apparently, there are services that enjoy a higher level of trust regarding data protection than others. Further developments to the app must ensure that users can decide whether they want to use the SMT data on 1 or more devices.

Despite the low number of study participants, the project shows the possibility of using citizen science in biomedical projects and research. The community-supported software development process successfully led to a functional SMT. In particular, the internet-based community provided the list of complementary measures, which were used as textual content in the SMT. Therefore, the SMT users could benefit from the collected patient knowledge about complementary measures without the effort of searching information on the internet.

Limitations

There are 2 major limitations of this project. The first is the low overall sample size and the second is a small observed effect.

For showing a statistically significant improvement of a minimum of 4 score points, 27 participants would have been necessary for the study. This number was not reached despite previous experience about the active internet-based community and high involvement in the requirement engineering process. Only 21 participants could be recruited. Of these, only 10 participants completed all questionnaires. We identified potential reasons for this: (1) failure to reach the target group via social media channels, (2) an admission process that was perceived as too complicated, and (3) the Christmas holidays. One indication of this is an increase in the dropout rate in the month of December. The additional leaflets sent in December could not mitigate the low involvement.

The observed effect of 0.2 score points reported in the results is negligible. The fact that the desired effect with a difference of 4 DLQI points could not be measured can be attributed to different causes. Besides the low effectiveness of the intervention, the intervention period (21 days) may play a role. The reason is that regular use over a long period of time may be necessary to effect changes. Therefore, a measured effect could be considered a placebo effect. In addition, the DLQI was recorded without considering the occurrence of relapses, possible rehabilitation stays, or an existing concomitant disease such as psoriatic arthritis. Comparing the average DLQI score determined in this study with the results of the study by Augustin et al [46] (mean 7.5, SD 6.4 points) reveals that the values do not differ much. Moreover, the variances in the respective samples are similar. This difference of less than 1 score point with the score in this study is noteworthy because Augustin et al included not only patients of dermatology outpatient hospital clinics but also patients of dermatologists in ambulant care. In this study, the participants were recruited via the community, which is why it is unlikely that the sample consists exclusively of dermatology outpatient hospital clinic patients.

Future Directions

Based on the feedback provided by the community and app users, the existing smartphone app should be completely revised. The authors strongly recommend developing such an app using a community-driven process again.

The improvements to the subsequent versions of the SMT can be roughly divided into four areas: (1) therapies including conventional medical therapies, hospital and rehabilitation stays, and complementary measures; (2) psoriasis relapses and triggers that can be analyzed; (3) behavioral patterns or environmental data in everyday life; and (4) QoL, as well as psychological and physical factors measured textually or visually using questionnaires.

To implement SMTs sustainably in care, it should be better integrated into the existing structures of care processes and medical practices. Direct data transmission to practitioners and further development of SMTs as approved medical products are conceivable.

Conclusions

The high participation of the internet-based community in the development process and the response to the first survey (N=97) shows a general need for independently developed SMTs for people with psoriasis. However, the collected feedback shows that the solution presented in the paper does not meet the needs of the community. The authors conclude that a more customizable app is needed.

Acknowledgments

The authors would like to thank all the members of the psoriasis-netz.de community who provided feedback; all the participants of the study for their time, effort, and feedback on the app; the team from Cardiff University for their support in granting the DLQI license; and the staff in Heilbronn University (especially the groups in E112 and E117) for their continuous support throughout the SMT development.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured German questionnaire and English translation used in the requirement engineering process.
[DOCX File , 16 KB - formative_v6i7e32593_app1.docx]

Multimedia Appendix 2
Original German questionnaire and English translation for usability and acceptance of the self-management tool.
[DOCX File , 26 KB - formative_v6i7e32593_app2.docx]

Multimedia Appendix 3
Study information website with information on the project that can be downloaded as a PDF file.
[PDF File (Adobe PDF File), 1489 KB - formative_v6i7e32593_app3.pdf]

Multimedia Appendix 4
Informed consent of participants.
[PDF File (Adobe PDF File), 1208 KB - formative_v6i7e32593_app4.pdf]

Multimedia Appendix 5
List of complementary measures suggested in the app, considering interactions.
[DOCX File , 17 KB - formative_v6i7e32593_app5.docx]

References

1. Schäfer T. Epidemiology of psoriasis. Review and the German perspective. Dermatology 2006;212(4):327-337. [doi: 10.1159/000092283] [Medline: 16707882]
2. Lee EB, Wu KK, Lee MP, Bhutani T, Wu JJ. Psoriasis risk factors and triggers. Cutis 2018 Nov;102(5S):18-20. [Medline: 30566552]
3. Rousset L, Halioua B. Stress and psoriasis. Int J Dermatol 2018 Oct;57(10):1165-1172. [doi: 10.1111/ijd.14032] [Medline: 29729012]
4. Russo PAJ, Ilchef R, Cooper AJ. Psychiatric morbidity in psoriasis: a review. Australas J Dermatol 2004 Aug;45(3):155-9;-quiz 160-1. [doi: 10.1111/j.1440-0960.2004.00078.x] [Medline: 15250891]
5. Augustin M, Glaeske G, Hagenström K. Psoriasisreport: Ergebnisse von Routinedaten-Analysen der Techniker Krankenkasse aus den Jahren 2017–2019. 2019. URL: https://www.tk.de/resource/blob/2105142/9fc664ddbfa42f94c3796a2211469f7f/hautreport-psoriasis-lang-data.pdf [accessed 2022-06-24]
6. Jankowiak B, Kowalewska B, Krajewska-Kułak E, Khvorik DF. Stigmatization and quality of life in patients with psoriasis. Dermatol Ther (Heidelb) 2020 Apr;10(2):285-296 [FREE Full text] [doi: 10.1007/s13555-020-00363-1] [Medline: 32146709]
7. Lesner K, Reich A, Szepietowski JC, Dalgard FJ, Gieler U, Tomas-Aragones L, et al. Determinants of psychosocial health in psoriatic patients: a multi-national study. Acta Derm Venereol 2017 Nov;97(10):1182-1188 [FREE Full text] [doi: 10.2340/00015555-2760] [Medline: 28795763]
8. Greb JE, Goldminz AM, Elder JT, Lebwohl MG, Gladman DD, Wu JJ, et al. Psoriasis. Nat Rev Dis Primers 2016 Nov;2:16082. [doi: 10.1038/nrdp.2016.82] [Medline: 27883001]
9. Dalgard FJ, Gieler U, Tomas-Aragones L, Lien L, Poot F, Jemec GBE, et al. The psychological burden of skin diseases: a cross-sectional multicenter study among dermatological out-patients in 13 European countries. J Invest Dermatol 2015 Apr;135(4):984-991 [FREE Full text] [doi: 10.1038/jid.2014.530] [Medline: 25521458]
10. Rendon A, Schäkel K. Psoriasis pathogenesis and treatment. Int J Mol Sci 2019 Mar;20(6):1475 [FREE Full text] [doi: 10.3390/ijms20061475] [Medline: 30909615]
11. Gamret AC, Price A, Fertig RM, Lev-Tov H, Nichols AJ. Complementary and alternative medicine therapies for psoriasis: a systematic review. JAMA Dermatol 2018 Nov;154(11):1330-1337. [doi: 10.1001/jamadermatol.2018.2972] [Medline: 30193251]
12. Patienten oft beim falschen Arzt. AerzteZeitung. URL: https://www.aerztezeitung.de/Medizin/Patienten-oft-beim-falschen-Arzt-242017.html [accessed 2021-07-14]
13. Online-Umfrage: Schuppenflechte-Patienten gehen oft nicht zum Arzt. NetDoktor GmbH. URL: https://www.netdoktor.at/news/psoriasis-umfrage-patienten-8514075 [accessed 2021-07-14]
14. Stern RS, Nijsten T, Feldman SR, Margolis DJ, Rolstad T. Psoriasis is common, carries a substantial burden even when not extensive, and is associated with widespread treatment dissatisfaction. J Investig Dermatol Symp Proc 2004 Mar;9(2):136-139 [FREE Full text] [doi: 10.1046/j.1087-0024.2003.09102.x] [Medline: 15083780]

15. Nast A, Boehncke WH, Mrowietz U, Ockenfels HM, Philipp S, Reich K, Deutsche Dermatologische Gesellschaft, Berufsverband Deutscher Dermatologen. German S3-guidelines on the treatment of psoriasis vulgaris (short version). Arch Dermatol Res 2012 Mar;304(2):87-113. [doi: 10.1007/s00403-012-1214-8] [Medline: 22350179]
16. Richards HL, Fortune DG, O'Sullivan TM, Main CJ, Griffiths CE. Patients with psoriasis and their compliance with medication. J Am Acad Dermatol 1999 Oct;41(4):581-583. [Medline: 10495380]
17. Nast A, Altenburg A, Augustin M, Boehncke W, Härle P, Klaus J, et al. Deutsche S3-Leitlinie zur Therapie der Psoriasis vulgaris, adaptiert von EuroGuiDerm - Teil 1: Therapieziele und Therapieempfehlungen. J Dtsch Dermatol Ges 2021 Jun;19(6):934-951. [doi: 10.1111/ddg.14508_g] [Medline: 34139080]
18. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov;288(19):2469-2475. [Medline: 12435261]
19. Newman S, Steed L, Mulligan K. Self-management interventions for chronic illness. The Lancet 2004 Oct;364(9444):1523-1537. [doi: 10.1016/S0140-6736(04)17277-2] [Medline: 15500899]
20. Larsen MH, Hagen KB, Krogstad A, Aas E, Wahl AK. Limited evidence of the effects of patient education and self-management interventions in psoriasis patients: a systematic review. Patient Educ Couns 2014 Feb;94(2):158-169. [doi: 10.1016/j.pec.2013.10.005] [Medline: 24184041]
21. Wink M, Funke J. Wissenschaft für alle: Citizen Science. Heidelberg, Germany: Heidelberg University Publishing; 2017.
22. Wiggins A, Wilbanks J. The rise of citizen science in health and biomedical research. Am J Bioeth 2019 Aug;19(8):3-14. [doi: 10.1080/15265161.2019.1619859] [Medline: 31339831]
23. Safdari R, Firoz A, Masoorian H. Identifying training and informational components to develop a psoriasis self-management application. Med J Islam Repub Iran 2017 Oct;31:67 [FREE Full text] [doi: 10.14196/mjiri.31.67] [Medline: 29445696]
24. Trettin B, Danbjørg DB, Andersen F, Feldman S, Agerskov H. An mHealth app to support patients with psoriasis in relation to follow-up consultations: qualitative study. JMIR Dermatol 2021 Jun 8;4(1):e28882 [FREE Full text] [doi: 10.2196/28882]
25. Trettin B, Danbjørg DB, Andersen F, Feldman S, Agerskov H. Development of an mHealth app for patients with psoriasis undergoing biological treatment: participatory design study. JMIR Dermatol 2021 May 10;4(1):e26673 [FREE Full text] [doi: 10.2196/26673]
26. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. Clin Exp Dermatol 1994 May;19(3):210-216. [doi: 10.1111/j.1365-2230.1994.tb01167.x] [Medline: 8033378]
27. Selbsthilfe vor Ort. Deutscher Psoriasis Bund eV. URL: https://www.psoriasis-bund.de/selbsthilfe-vor-ort/ [accessed 2021-07-08]
28. Herzlich willkommen!. Psoriasis Selbsthilfe Arbeitsgemeinschaft eV (PSOAG). URL: https://www.psoriasis-selbsthilfe.org/ [accessed 2021-07-08]
29. Gemeinsam mit Schuppenflechte. Psoriasis-Netz. URL: https://www.psoriasis-netz.de/ [accessed 2021-07-08]
30. Mockups 3 for Desktop | Balsamiq. Balsamiq Studios, LLC. URL: https://balsamiq.com/wireframes/mockups3fordesktop/ [accessed 2021-07-08]
31. Hollenbenders Y, Brandl L, Kappler R, Mezger M, Prokosch M, Zsebedits D, et al. Supporting Cancer Patients by recommending Complementary Methods based on Individual Complaints. Germany: German Medical Science GMS Publishing House; 2017.
32. Hyland ME. A brief guide to the selection of quality of life instrument. Health Qual Life Outcomes 2003 Jul;1:24 [FREE Full text] [doi: 10.1186/1477-7525-1-24] [Medline: 12848895]
33. Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. BMJ 2002 Jun;324(7351):1417 [FREE Full text] [doi: 10.1136/bmj.324.7351.1417] [Medline: 12065262]
34. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess 1998;2(14):i-iv,-i1-74 [FREE Full text] [Medline: 9812244]
35. Dermatology Life Quality Index. Cardiff University. URL: https://www.cardiff.ac.uk/medicine/resources/quality-of-life-questionnaires/dermatology-life-quality-index [accessed 2021-07-15]
36. Basra MKA, Salek MS, Camilleri L, Sturkey R, Finlay AY. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. Dermatology 2015 Jan;230(1):27-33 [FREE Full text] [doi: 10.1159/000365390] [Medline: 25613671]
37. Fredriksson T, Pettersson U. Severe psoriasis--oral therapy with a new retinoid. Dermatologica 1978;157(4):238-244. [doi: 10.1159/000250839] [Medline: 357213]
38. Bortz J, Schuster C. Statistik für Human und Sozialwissenschaftler. Berlin, Heidelberg: Springer; 2010.
39. MathWorks stellt Release 2019b von MATLAB und Simulink vor. MathWorks. URL: https://de.mathworks.com/company/newsroom/mathworks-announces-release-r2019b-of-matlab-and-simulink.html [accessed 2021-07-08]
40. Brecht U. Satzung der Ethikkommission der Hochschule Heilbronn. URL: https://tinyurl.com/2hdh4txw [accessed 2022-05-30]
41. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
42. The modern web developer's platform. Angular. URL: https://angular.io/ [accessed 2021-07-08]

43. Webb P. Spring Boot 2.1.5 released. Spring. 2019 May 15. URL: https://spring.io/blog/2019/05/15/spring-boot-2-1-5-released [accessed 2021-07-08]
44. Java SE Development Kit 11- - Downloads. Oracle Deutschland. URL: https://www.oracle.com/de/java/technologies/javase-jdk11-downloads.html [accessed 2021-07-08]
45. PostgreSQL 10 Released. PostgreSQL Global Development Group. URL: https://www.postgresql.org/about/news/postgresql-10-released-1786/ [accessed 2021-07-08]
46. Augustin M, Sommer R, Kirsten N, Danckworth A, Radtke MA, Reich K, et al. Topology of psoriasis in routine care: results from high-resolution analysis of 2009 patients. Br J Dermatol 2019 Aug;181(2):358-365. [doi: 10.1111/bjd.17403] [Medline: 30430557]

Abbreviations

DLQI: Dermatology Life Quality Index
PASI: Psoriasis Area and Severity Index
PSOAG: Psoriasis Selbsthilfe Arbeitsgemeinschaft eV (Psoriasis Self-Help Association)
PWA: progressive web app
QoL: quality of Life
SMT: self-management tool

Edited by T Leung; submitted 03.08.21; peer-reviewed by S Gulliver, T van Hal; comments to author 11.10.21; revised version received 08.11.21; accepted 01.06.22; published 07.07.22.

Please cite as:
Brandl LC, Liebram C, Schramm W, Pobiruchin M
A German Smartphone-Based Self-management Tool for Psoriasis: Community-Driven Development and Evaluation of Quality-of-Life Effects
JMIR Form Res 2022;6(7):e32593
URL: https://formative.jmir.org/2022/7/e32593
doi:10.2196/32593
PMID:35797109

Original Paper

Exploring Users' Experiences With a Quick-Response Chatbot Within a Popular Smoking Cessation Smartphone App: Semistructured Interview Study

Alice Alphonse[1], MSc; Kezia Stewart[1], MSc; Jamie Brown[1], PhD; Olga Perski[1], PhD

Department of Behavioural Science and Health, University College London, London, United Kingdom

Corresponding Author:
Olga Perski, PhD
Department of Behavioural Science and Health
University College London
1-19 Torrington Place
London, WC1E 7HB
United Kingdom
Phone: 44 0207679 ext 1258
Email: olga.perski@ucl.ac.uk

Abstract

Background: Engagement with smartphone apps for smoking cessation tends to be low. Chatbots (ie, software that enables conversations with users) offer a promising means of increasing engagement.

Objective: We aimed to explore smokers' experiences with a quick-response chatbot (*Quit Coach*) implemented within a popular smoking cessation app and identify factors that influence users' engagement with Quit Coach.

Methods: In-depth, one-to-one, semistructured qualitative interviews were conducted with adult, past-year smokers who had voluntarily used Quit Coach in a recent smoking cessation attempt (5/14, 36%) and current smokers who agreed to download and use Quit Coach for a minimum of 2 weeks to support a new cessation attempt (9/14, 64%). Verbal reports were audio recorded, transcribed verbatim, and analyzed within a constructivist theoretical framework using inductive thematic analysis.

Results: A total of 3 high-order themes were generated to capture users' experiences and engagement with Quit Coach: anthropomorphism of and accountability to Quit Coach (ie, users ascribing human-like characteristics and thoughts to the chatbot, which helped foster a sense of accountability to it), Quit Coach's interaction style and format (eg, positive and motivational tone of voice and quick and easy-to-complete check-ins), and users' perceived need for support (ie, chatbot engagement was motivated by seeking distraction from cravings or support to maintain motivation to stay quit).

Conclusions: Anthropomorphism of a quick-response chatbot implemented within a popular smoking cessation app appeared to be enabled by its interaction style and format and users' perceived need for support, which may have given rise to feelings of accountability and increased engagement.

(JMIR Form Res 2022;6(7):e36869) doi:10.2196/36869

KEYWORDS

chatbot; conversational agent; engagement; smartphone app; smoking cessation; accountability; mobile phone

Introduction

Diseases caused by cigarette smoking are a leading cause of preventable death, killing approximately 8 million people each year globally [1]. Smoking-attributable diseases place significant financial burden on health care systems, with costs estimated to be approximately 5.7% of the total annual global health care expenditure [2]. Therefore, improved behavioral or pharmacological smoking cessation support is a priority for individuals, public health bodies, and governments. However, in-person smoking cessation services are challenged by scalability and face substantial funding cuts across many countries [3], especially after many had to offer only remote services owing to the COVID-19 pandemic [4]. With growing internet access and smartphone ownership, digital interventions (including smartphone apps) provide a low-cost means of scaling up the delivery and optimizing the reach of evidence-based smoking cessation support [5]. However, the available smoking cessation apps tend to generate low average levels of user engagement [6,7]—although estimates vary between apps

[8]—which may reduce the likelihood of success in quitting smoking. Therefore, identifying modifiable factors (eg, content and design elements) that positively influence engagement with smoking cessation apps is important. A promising means of improving digital engagement is through provision of concurrent health care professional support [9,10]. However, this is limited by the cost and availability of health care professionals. Therefore, an underexplored and relatively low-cost means of improving engagement is to replicate health care professional support via chatbots (ie, software that enables 2-way conversations with app users). This study aimed to explore smokers' experiences with a quick-response chatbot (*Quit Coach*) implemented within a popular smoking cessation app and identify factors that influence users' engagement with Quit Coach, using a qualitative approach.

Engagement with digital interventions can be defined as (1) the extent (eg, amount, frequency, duration, and depth) of use and (2) a subjective experience characterized by attention, interest, and affect [9]. A distinction between *microengagement* (ie, engaging with the technology itself) and *macroengagement* (ie, engaging in the behavior change process, such as abstaining from smoking) has also been proposed [11]. A common pattern observed across several studies is that engagement is positively associated with intervention effectiveness [12-14]. Therefore, many researchers and intervention designers have focused their efforts on identifying factors that increase engagement with digital interventions in general and with smoking cessation apps in particular. These studies have identified factors such as app content or behavior change techniques (eg, goal setting, reminders, self-monitoring, social support, and health care professional support), design elements (eg, tailoring of content, using a nonjudgmental message tone, and gamification), and cognitive considerations (eg, minimizing cognitive load and providing user guidance) as being important for increased engagement [10,15-17]. However, despite such studies, early disengagement from smoking cessation apps remains common.

Chatbots (also referred to as *conversational agents*) are computer programs that have tailored conversations with users via text or audio-visual messaging [18]. Some chatbots are specifically designed to appear as social actors (ie, *relational agents*), with the intention that users form social-emotional relationships with the bot [19,20]. Current chatbot implementations include structured (or decision tree–based) and unstructured bots. The former type enables the user to select relevant options from a list of predefined responses, with the bot responding with prewritten messages following conditional if-then rules (also referred to as *quick reply responses* or *quick-response* bots [21]). The latter type typically relies on natural language processing, with the user inputting open or unstructured messages that are processed and responded to by the bot. According to the Model of Supportive Accountability, human support (eg, from a health care professional or coach) is expected to promote engagement with digital interventions by fostering a sense of accountability to a trustworthy, benevolent, and competent coach [22]. Although underexplored, it is plausible that chatbots may fulfill the role of such human support by offering *human-like* support [19].

Chatbots are a relatively new addition in the health care domain, with recent systematic reviews identifying only a handful of studies of chatbots for improving mental health [23] and increasing physical activity and healthy diets [24]. Within the substance use and smoking cessation domains, a few early single-arm and 2-arm randomized studies have yielded promising results [25-28]. For example, a chatbot incorporating the principles of motivational interviewing—designed specifically to support smokers who are unmotivated to stop—tested positively in an early user-testing study [25]. An adapted version of the cognitive behavioral therapy–informed chatbot, *Woebot*, for people who use addictive substances was found to be acceptable to deliver, engaging, and associated with improvements in mental health and substance use outcomes in a single-arm study [28]. We found that the addition of a supportive, quick-response chatbot to a popular smoking cessation app more than doubled the user engagement and improved short-term quit success in a large, 2-arm, experimental study [27]. However, the available single-arm and 2-arm quantitative studies have not focused on the potential mechanisms underpinning this increased engagement (eg, owing to limited data collection). Qualitative studies of users' experiences with the relational agent, *Replika*, have found that such companion chatbots can mimic human interaction, with users perceiving their relationships with the designated bot as rewarding [20,29]. However, qualitative investigations of users' experiences with chatbots designed specifically to support smoking cessation are lacking. Therefore, this qualitative study aimed to address the following research questions:

1. What are smokers' experiences with a quick-response chatbot (*Quit Coach*) implemented within a popular smartphone app?
2. What are the factors that influence users' engagement with Quit Coach?

Methods

Study Design

The Consolidated Criteria for Reporting Qualitative Research checklist was used in the design and reporting of this study [30]. Semistructured, one-to-one interviews were conducted.

Theoretical Framework

A constructivist theoretical framework was used to inform data collection and analysis [31]. This theoretical approach was selected because constructivism recognizes the active role of the researcher in the generation and interpretation of qualitative data.

Participants

For pragmatic purposes, participants were recruited across 2 periods: June 2020 to August 2020 (led by KS and OP) and April 2021 to August 2021 (led by AA and OP). Owing to the COVID-19 pandemic, it was challenging to recruit as planned during the summer of 2020. Therefore, we continued the recruitment in 2021. The project team decided that it would be useful to recruit participants from 2 different subgroups (ie, past-year smokers who had voluntarily used Quit Coach in a recent smoking cessation attempt and current smokers who

agreed to download and use Quit Coach for a minimum of 2 weeks to support a new cessation attempt) as a form of triangulation [32]. We reasoned that such triangulation of results when varying the eligibility criteria (rather than the methods) would either help to validate the results (eg, if smokers who did not self-select to download and use the Smoke Free app had similar experiences with Quit Coach as those who had voluntarily used the app) or highlight different experiences owing to smoking status or treatment-seeking behavior.

Participants recruited in 2020 were eligible to participate if they (1) were aged ≥18 years, (2) were fluent English speakers based in the United Kingdom, (3) were past-year smokers and had used the *pro* (ie, paid) version of the Smoke Free app (ie, the app version that included Quit Coach) for at least two weeks, and (4) had interacted with Quit Coach at least once during the 2-week period.

Participants recruited in 2021 were eligible to participate if they (1) were aged ≥18 years; (2) were fluent or highly competent English speakers, with no restrictions on geography; (3) were current cigarette smokers; (4) were willing to make a quit attempt within 1 week from initial contact with the researchers and use Quit Coach for at least two weeks; and (5) owned a smartphone.

All the participants used the *pro* version of the Smoke Free app for a period of at least 2 weeks before participating in the semistructured interviews. We expected this time window to be sufficient for enabling detailed conversation about participants' chatbot experiences.

Sampling

Participants recruited in 2020 were approached through advertisements (unpaid) shared on social media platforms (ie, Facebook and Twitter) and through a mailing list of Smoke Free app users. The recruitment materials stated that Smoke Free users were invited to participate in a web-based interview about their experiences with the app (good or bad), with particular focus on the Quit Coach feature. Participants were incentivized to win 1 of 5 gift vouchers worth £20 (approximately US $24).

Participants recruited in 2021 were approached through advertisements (unpaid) shared on social media platforms (ie, LinkedIn, Facebook, and Instagram), directly through the researchers' networks (ie, WhatsApp, email, SMS text messages, and flyers), and through professional web-based recruitment platforms (ie, Prolific and Call for Participants). The recruitment materials stated that smokers interested in making a quit attempt with the use of a smartphone app were invited to participate in a web-based interview about their experiences with the app (good or bad), with particular focus on its chatbot feature. Participants received a gift voucher worth £10 (approximately US $12) after completing the interview.

Participants were recruited in batches of 4 to 5 participants each until theoretical saturation was judged to have occurred (ie, a point in the data collection process when no new information alters the identified themes) [33]. Preliminary data analysis was conducted by KS and subsequently by AA after each batch of 4 to 5 participants, to determine whether additional participants were needed.

Measures

Eligibility and Sample Characteristics

Data were collected to determine eligibility and characterize the sample based on (1) age; (2) gender (female, male, or in another way); (3) country of residence; (4) whether they were fluent or highly competent English speakers (yes or no); (5) time to first cigarette (<5, 6-30, 31-60, or >60 minutes or not applicable); (6) cigarettes smoked per day (<10, 11-20, 21-30, ≥31, or not applicable); and (7) motivation to stop, measured with the validated Motivation to Stop Scale [34].

Participants recruited in 2020 were asked the questions mentioned previously and to provide additional information on the following: (1) job type (manual or nonmanual); (2) smoking status ("I smoke cigarettes [including hand-rolled] every day"; "I smoke cigarettes [including hand-rolled], but not every day"; "I don't smoke cigarettes at all, but I do smoke tobacco of some kind [e.g. pipe, cigar or shisha]"; "I have stopped smoking completely in the last year"; "I stopped smoking completely more than a year ago"; or "I have never been a smoker [i.e. smoked for a year or more]"); and (3) self-reported use of Quit Coach (none at all, a little, moderately, a lot, or extremely).

Participants recruited in 2021 were asked the questions mentioned previously and to provide additional information on the following: (1) the number of past-year quit attempts and (2) whether they had ever used any app-based support to help stop smoking (and if so, the name of the app).

Interview Topic Guide

The topic guide was informed by the Model of Supportive Accountability [22] to address specific theoretical concepts (eg, information quality, reliability, and accountability) and split into 3 sections: an introductory section to allow participants to warm up, covering general experiences with the app; a second section exploring users' experiences with Quit Coach; and a final section exploring situations in which participants engaged with Quit Coach (Multimedia Appendix 1). Prompts were used to encourage participants to elaborate on their impressions and experiences. The topic guide was pilot-tested by KS on 2 graduate colleagues from the MSc program in Behavior Change and adapted following their feedback. The topic guide was further adapted following the interviews conducted in 2020 to facilitate elaboration by adding specific probes in addition to a new question ("To what extent would you say you formed a relationship of sorts with the chatbot? How was this?"). As participants mentioned their relationship with the chatbot, we considered it useful (and consistent with the flexible, semistructured style of interviewing) to specifically prompt subsequent participants about this. The adapted topic guide was piloted by AA on a graduate colleague from the MSc program in Behavior Change and a current smoker from AA's network and updated according to their feedback. Interviews remained flexible, facilitated by the semistructured style of questioning.

Procedure

Upon expressing interest, participants were asked to read the participant information sheet, provide informed consent, and complete the eligibility questionnaire via Qualtrics (Qualtrics

International Inc). If eligible, participants recruited in 2020 were contacted by the researchers to arrange the interview. If eligible, participants recruited in 2021 were contacted to confirm study acceptance and provided with information on how to download the *pro* version of the Smoke Free app (using a free access code). Participants were asked to select a quit date and nominate a day that was 2 weeks from their quit date to complete the interview.

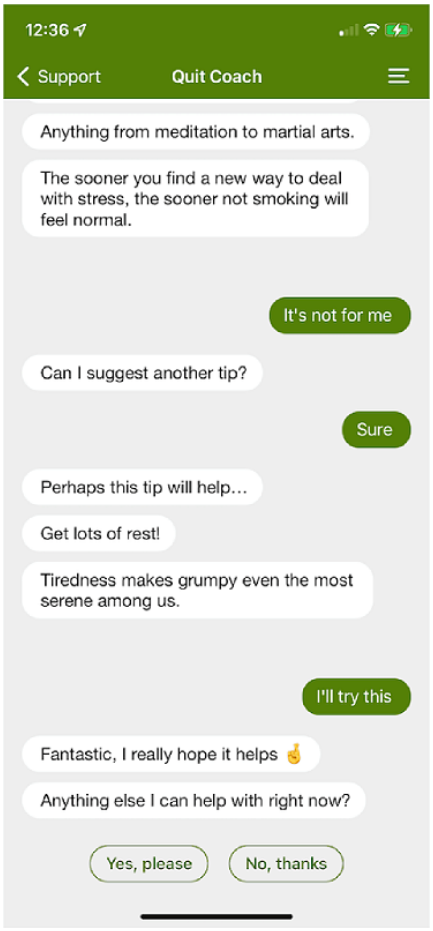
Owing to the COVID-19 pandemic, in-person interviews were not possible. Therefore, interviews were conducted via the web by KS or AA (graduate students enrolled in an MSc program in Behavior Change) via Microsoft Teams. Besides the participant and the researcher, no one else was present during the interviews. KS had limited experience in conducting qualitative interviews before this study; AA had extensive experience from working for a consultancy firm. Before conducting the interviews, OP (PhD in Health Psychology, extensive experience in conducting qualitative interviews through previous academic work) provided training to KS and AA. Interviews were audio recorded and lasted between 30 and 45 minutes. Following completion, the participants were thanked, verbally debriefed, and presented with an incentive.

The Smoke Free App and Quit Coach

Smoke Free [35] is an evidence-informed app with a large user base (approximately 4000 downloads per day). The app contains behavior change techniques that are expected from theory and evidence from other settings to aid smoking cessation [17,36]. Refer to the study by Jackson et al [37] for a summary of the behavior change techniques included in the Smoke Free app, coded against a 44-item taxonomy of techniques used in individual behavioral support for smoking cessation [36].

The *pro* (ie, paid) version of Smoke Free contains a text-based, quick-response chatbot called *Quit Coach* (Figure 1). During the first 2 weeks of a user's quit attempt, Quit Coach initiates twice-daily check-ins with users regarding their cessation attempt through a push notification. Check-in frequency is programmed to reduce after 1 month and cease entirely after 90 days (when users are anticipated to have quit smoking). Users engage with Quit Coach via text messages, selecting from prewritten responses. There are a few exceptions to this format, such as when users complete certain exercises that require free-text input (eg, typing the mantra, "not another puff, no matter what"), to occasionally type what is influencing their craving, or for providing feedback on whether they found a piece of advice useful (refer to Multimedia Appendix 1 for additional screenshots of such interactions). A bespoke Node .js natural language processing framework (adapted from freely available, state-of-the-art source code by Smoke Free's developers) is used to map free-text inputs onto their likely intent—this constitutes the only machine learning element of the chatbot. Users can also initiate check-ins themselves by opening Quit Coach to record a craving or ask for assistance via a *get help now toolkit*, which provides different options for directing the conversation with Quit Coach. The conversational options include craving management, relapse, difficult situations, and withdrawal. Quit Coach's communications (typically in text form, but also through emojis and Graphics Interchange Formats [GIFs]) contain information about the health consequences of smoking, quitting tips, and motivational messages, simulating a text message conversation.

Figure 1. Example screenshot of the Smoke Free chatbot (Quit Coach).



Data Analysis

Interviews conducted in 2020 and 2021 were combined to form a single data set. Analysis was performed through an inductive thematic approach following the methodology by Braun and Clarke [38,39]: (1) familiarizing with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing the themes, (5) defining and naming the themes, and (6) producing the report. We focused on latent (rather than semantic) meanings [39]. Although the Model of Supportive Accountability was used to inform the interview topic guide, terminology from this model was used for coding only if deemed relevant, with alternative codes considered throughout the process.

Interviews were coded using Microsoft Word in batches of 4 to 5 by AA to facilitate an iterative and reflexive approach. Each transcript was read multiple times to familiarize with the data. Then, initial codes were generated. Coded transcripts were reread following initial code generation and discussed with OP, adding or refining codes as appropriate. Then, the coded extracts were examined and used to generate preliminary themes. Next, a second, independent coder (another student from the same MSc program in Behavior Change) helped to assess coding reliability. The second coder coded 2 interviews, 1 each from the 2020 and 2021 samples, which were selected using a random number generator. The second coder was instructed to inductively code each interview. The resulting codes were compared with those of the first coder conceptually, rather than for perfect word matching. Discrepancies were discussed and reconciled. Then, themes were reviewed, refined, named, and agreed upon through discussion among AA, OP, and JB. During coding, the possibility for differences between the 2020 and 2021 samples was considered. Theoretical saturation was judged to have been reached after 12 interviews.

External Validation

A subsample of 14% (2/14) of randomly selected participants were contacted and agreed to read the results and comment on the congruence of the themes and narrative generated by the researchers with their own experiences. Both participants agreed with the researchers' interpretations.

Reflexivity

The interviewers (women, White ethnicity, nonsmokers, and unfamiliar with most participants before the interview) felt that a good rapport was built with all the participants. Some were more immediately verbose, whereas others took a little time to open up, but did so with prompting and encouragement. For the first few interviews conducted, the interviewers closely followed the topic guide; however, as salient conversation topics emerged, a more discursive style of questioning was adopted to explore salient topics in great depth. Before commencing the interviews, participants were told about the goals of the study and that the interviewers were not directly involved in the development of

the Smoke Free app; however, participants were unaware of the interviewers' smoking status or theoretical assumptions regarding user engagement or smoking cessation.

Ethics Approval

Ethics approval for this study was obtained from University College London's Research Ethics Committee (CEHP/2020/579). Participants provided written informed consent before participating in the study.

Results

Participant Characteristics

A total of 40 participants completed the screening survey and were eligible to participate in the study. Of the 40 participants, 26 (65%) participants did not complete an interview, as they later decided that it was not the right time to quit, became uncontactable, or failed to attend the interview. Table 1 shows a summary of the demographic and smoking characteristics of the 35% (14/40) included participants.

Table 1. Participants' demographic and smoking characteristics (n=14).

Participant ID	Recruitment year	Age (years)[a]	Country	Gender	Cigarettes per day at baseline	Smoking status at the time of interview	Quit Coach use	Number of past-year quit attempts	Use of app-based support to stop smoking (name of the app)
P1	2020	30	United Kingdom	Female	N/A[b]	I have stopped smoking completely in the last year	Moderate	N/A	N/A
P2	2020	20	United Kingdom	Male	N/A	I have stopped smoking completely in the last year	A lot	N/A	N/A
P3	2020	33	United Kingdom	Female	N/A	I have stopped smoking completely in the last year	Extreme	N/A	N/A
P4	2020	39	United Kingdom	Female	N/A	I have stopped smoking completely in the last year	A lot	N/A	N/A
P5	2020	44	United Kingdom	Female	N/A	I have stopped smoking completely in the last year	A lot	N/A	N/A
P6	2021	25-34	United Kingdom	Male	7	Quit[c]	N/A	1	No (N/A)
P7	2021	18-24	France	Female	3	Cut down[c]	N/A	3	Yes (Qwit)
P8	2021	18-24	France	Female	10	Cut down[c]	N/A	1	No (N/A)
P9	2021	25-34	France	Male	12	Cut down[c]	N/A	3	Yes (Smoke Free)
P10	2021	18-24	United Kingdom	Female	5	Cut down[c]	N/A	0	No (N/A)
P11	2021	18-24	Mauritius	Female	10	Cut down[c]	N/A	1	No (N/A)
P12	2021	25-34	India	Male	6	Cut down[c]	N/A	2	No (N/A)
P13	2021	18-24	United Kingdom	Male	10	Quit[c]	N/A	3	No (N/A)
P14	2021	18-24	United Kingdom	Male	8	Quit[c]	N/A	1	No (N/A)

[a]For participants recruited in 2021, age was measured as a range.

[b]N/A: not applicable.

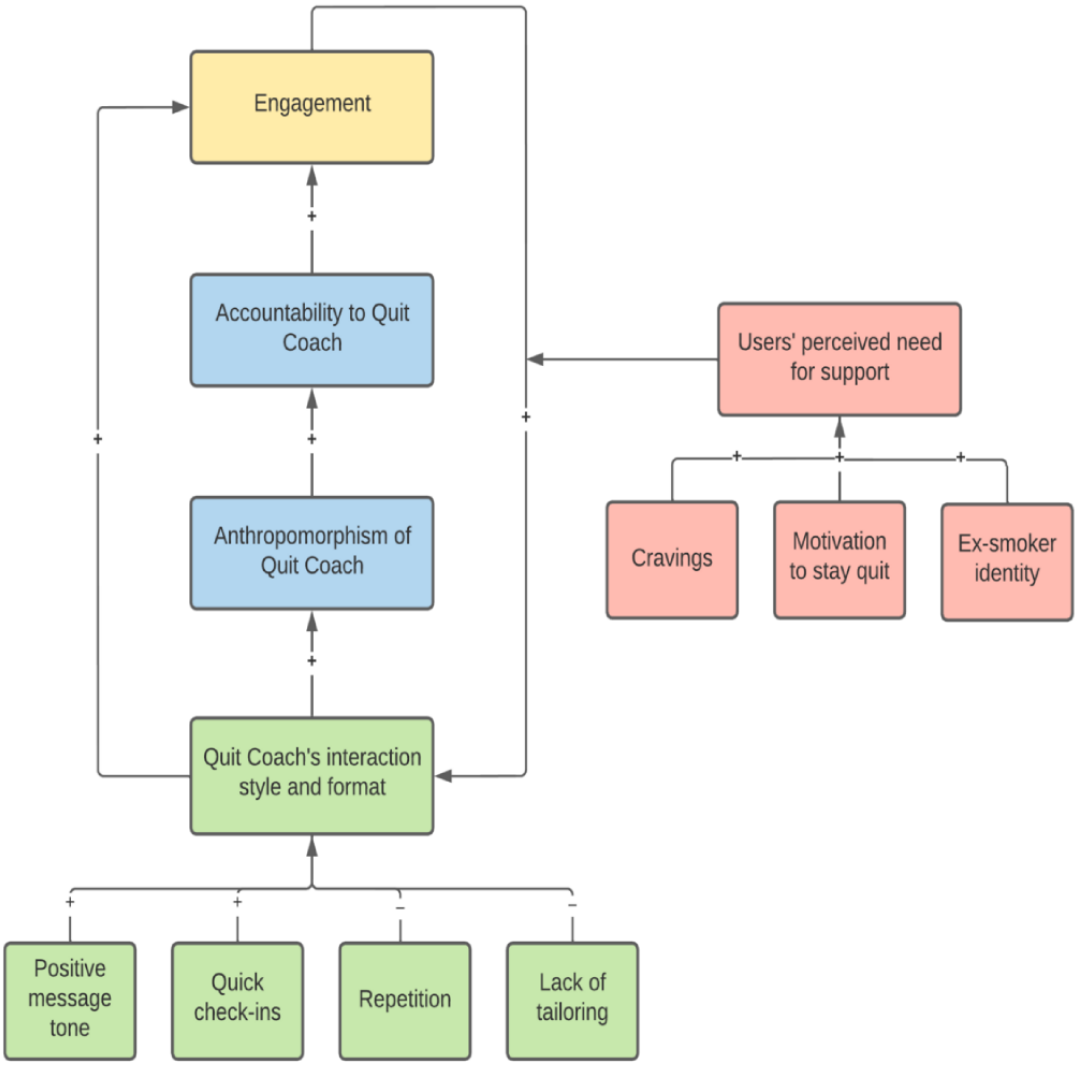
[c]Ascertained qualitatively during the interview.

Themes

Overview

A total of three high-order themes were developed to capture the participants' experiences with Quit Coach and the potential mechanisms underpinning user engagement: (1) anthropomorphism of and accountability to Quit Coach, (2) Quit Coach's interaction style and format, and (3) users' perceived need for support. Refer to Multimedia Appendix 1 for additional quotations. Figure 2 presents a thematic map of the themes. Here, anthropomorphism of Quit Coach, which is influenced by its interaction style and format, and users' perceived need for support leads to feelings of accountability and increased engagement. Continued engagement with Quit Coach reinforces this accountability through a bidirectional relationship. In addition to this indirect link, Quit Coach's interaction style and format directly influences users' engagement.

Figure 2. Thematic map. Arrows indicate the direction of relationships and the + and – signs indicate their valence. Blue boxes relate to theme 1, green boxes relate to theme 2, and red boxes relate to theme 3.



Anthropomorphism of and Accountability to Quit Coach

Many users ascribed human-like characteristics, thoughts, and behavior to Quit Coach—despite the awareness that it was fully automated—following an interaction experience (eg, colloquial language, GIFs, and emojis) that closely mimicked those with peers and family members via SMS text messages or WhatsApp. Users' anthropomorphism of Quit Coach was evident from almost all users frequently referring to it as an embodied entity (ie, him, them, or someone) rather than an object (eg, it, Quit Coach, or chatbot):

> *It felt like you were really connected to someone.* [P5; 2020; quit smoking]
>
> *Throughout this conversation I have referred to "someone" rather than "something" quite a few times, which I haven't done on purpose.* [P6; 2021; quit smoking]

Users compared Quit Coach's motivational and monitoring support with that of close friends or family members. Consequently, users described feeling that Quit Coach *cared* about them, which helped foster a relational bond and promoted

a feeling of accountability to Quit Coach to succeed in cessation. Feeling that Quit Coach was fulfilling a supportive, social role justified the daily monitoring (ie, push notifications), which was experienced by some participants as very frequent, boring, or annoying. This social presence promoted continued app engagement, as users were motivated to succeed and willing to engage and cooperate with Quit Coach:

> *It felt like a friend or family member seeing how I was...It cared about whether I was smoking or not...You feel like you've got someone who cares enough [that you] stop. A reason not to...* [P10; 2021; cut down smoking]

Many users even described Quit Coach's support as superior to human friends or family members, owing to its immediate availability, single purpose (ie, smoking cessation support), and nonjudgmental tone of voice (particularly when reporting a lapse) and the relationship being *1-way* (ie, users do not have to reciprocate support or worry about being boring or bothering Quit Coach). Such responses contradict the users' comparison of Quit Coach's support with that provided by friends or family members, indicating that they may instead have experienced the support as similar to that provided by a therapist or coach (ie, a 1-way rather than 2-way relationship). However, users' perceptions of Quit Coach's support as superior to that of human friends or family members appeared to serve as an advantage rather than to negate the human-like experience. A minority of users anthropomorphized Quit Coach to a lesser extent, owing to previous experience with using chatbots and great awareness of their synthetic nature. Nonetheless, they still embraced the support offered by Quit Coach:

> *[Friends/family] don't wanna talk about it every evening...To have the QC for that purpose [is helpful]...You're not always gonna be lucky enough to have someone who's just gonna be there to just listen to what you want to talk about on any given evening.* [P6; 2021; quit smoking]

> *It's cool that he wasn't judgemental...He says "lots of people do smoke again when they start to stop but it doesn't mean that it's their loss and that they need to start all over again, it's just the first step"...The point is that long term you are trying not to smoke.* [P8; 2021; cut down smoking]

Many users perceived having access to Quit Coach's thoughts and feelings, reporting that they worried that Quit Coach would feel angry, upset, or disappointed if they did not check in or reported a lapse and wanting to *please* it by checking in regularly. Consequently, many users felt positively accountable to Quit Coach for frequent engagement and abstinence. For example, many users reported feeling proud or particularly motivated to engage when they could report not smoking. In addition, most users reported feelings of worry or shame about Quit Coach's anticipated reaction if they had lapsed. For a few users, this caused an unintended consequence; their feelings were so significant that they reported completely avoiding checking in or lying in their reports. However, for most users, this anticipated worry or shame appeared to be beneficial, representing a source of motivation to stay quit. A few participants who failed in their quit attempt over time reported that they started ignoring check-ins, possibly owing to negative avoidance as the prompts would remind them of their *failure*:

> *I wanted to find an opportunity to make it happy.* [P3; 2020; quit smoking]

> *[I didn't want] to tell that [I smoked] to the robot every morning and every afternoon...I felt that I was accountable to it.* [P11; 2021; cut down smoking]

Some self-contradiction was observed in how accountable users reported feeling to Quit Coach. At some time points during the interviews, users stated feeling accountable to Quit Coach, but at different time points, they stated feeling accountable to themselves, friends, or family members. This contradiction typically arose after being specifically prompted about accountability. Users may have retrospectively changed how accountable they felt to Quit Coach, because the interviewer brought to the fore that Quit Coach was not real, despite an anthropomorphic experience:

> *The accountability thing was definitely my relationship with [people] rather than my relationship with the app.* [P6; 2021; quit smoking]

Quit Coach's Interaction Style and Format

Quit Coach's interaction style and format appeared to both directly and indirectly (ie, by giving rise to anthropomorphic experience of Quit Coach) influence users' engagement. Quit Coach's motivational and positive tone of voice encouraged many users to stay on track by reminding them of and praising them for their progress. GIFs and emojis were used alongside written text messages to create a positive mood and inject humor, thus enhancing the motivational and positive tone beyond that created through written text alone. Many users noted that the GIFs and emojis promoted a human-like perception of Quit Coach:

> *There was a certain level of wanting to go back and get those little GIFs or whatever...I was always glad to go and have a check-in.* [P6; 2021; quit smoking]

Engagement was generally driven by Quit Coach prompting check-ins rather than by the users themselves. The prompts were perceived as useful, as users may otherwise have forgotten to check in. Most users reported that the daily check-in time requirement was acceptable; it did not take up much of their day. During working hours, some users felt that check-ins were sufficiently short to be manageable, whereas some users were very busy, preferring to complete check-ins before or after working hours. This was supported by the ability to set preferred check-in times. Where users reported check-ins being *long*, it typically referred to a subjective experience of *long*, which was linked to boredom and lack of interest, often driven by repetition, forced choice, or limited opportunities for free-text inputs. Many users reported that forced choice made engagement feel less burdensome (ie, easy and less time-consuming), which contributed to check-ins being "just the right amount of time" and was particularly welcome when users were craving cigarettes. In contrast, many users reported that forced-choice interactions quickly became boring as they could not express what they wanted more precisely (as they would with a human).

This reduced the interest in engagement with Quit Coach, with many users indicating a preference for typing free-text questions and responses:

> *[If you're] distracted by wanting a cigarette [it's] just easier if you've got options in front of you to just pick one.* [P10; 2021; cut down smoking]

> *It was a bit...Samey. Sometimes I would just kind of click through it all and not have to react as much...You feel like you're just going through the motions a little bit rather than actually thinking about it.* [P6; 2021; quit smoking]

Most users mentioned that they would have liked tailoring of the chatbot interactions on 2 levels. First, although Quit Coach broadly aligned with most users' cessation motivations, it did not discuss the specific personal motives that users had inputted elsewhere in the app. Second, many users wanted Quit Coach to remember more about what worked for them and modify the messages and advice accordingly. Although users did not report the lack of desired tailoring to be particularly detrimental to engagement, most users indicated that enhanced tailoring could have a positive impact. The ability to set check-in times according to personal preferences or anticipated times of need promoted engagement. Several users agreed that if Quit Coach could prompt them to engage before or during triggering situations, it would be very useful:

> *It wasn't really a tailored fit for myself. It talked about infertility and [that's] not something that bothers me.* [P13; 2021; quit smoking]

> *A reminder of what I've done so far [that worked] would be helpful.* [P9; 2021; cut down smoking]

> *[It] would've been great if...he could send me a notification at [or before a] time [of anticipating being triggered] saying "you're gonna be OK" or "you're gonna make it!"* [P7; 2021; cut down smoking]

Users' Perceived Need for Support

For all users, the perceived need for support from Quit Coach appeared to be related to the frequency and intensity of cravings. Engaging with Quit Coach provided a useful behavioral alternative to smoking or attentional distraction from cravings. This was particularly helpful when check-ins aligned with strong cravings. In moments where users were not *smoking or thinking about smoking*, the need for support and, consequently, engagement interest appeared to be low. For some users, Quit Coach triggered cravings by reminding them of smoking:

> *It was [a] great thing to keep my hands [busy] [and] just give me time to let the craving pass.* [P2; 2020; quit smoking]

> *When it was going well and when I didn't smoke I just didn't even use the app because I was OK...I was like I'm doing well so what can the app give me right now?* [P8; 2021; cut down smoking]

As cravings reduced, users' perceived need for support decreased, leading to reduced engagement interest. However, several users recruited in 2020 (all of whom had successfully quit smoking) reported engaging with Quit Coach even after their cravings reduced or disappeared. For these users, the need for support shifted from primarily needing a distraction from cravings to maintaining motivation to stay quit and reinforce ex-smoker identity. Self-selection bias may explain this prolonged engagement; users recruited in 2020 were already using Quit Coach, had successfully quit smoking, and self-identified as *heavy* Quit Coach users:

> *[Usage] kind of dwindles down to kind of extreme need, its more about the check-ins in the morning...The chatbot doesn't have much use now it's like ten weeks or something since I stopped.* [P2; 2020; quit smoking]

For users recruited in 2021 who were unsuccessful in cessation, the motivation to engage with Quit Coach decreased after a period of *failing*. These users felt discouraged and despondent and did not want to be reminded of their *failure*. Interestingly, this was sometimes accompanied by a reversal of the anthropomorphic experience (eg, referring to Quit Coach as a *robot*):

> *[Not succeeding in quitting] made me resent – not resent, that's a big word – but made me not want to tell that to the robot every morning and every afternoon.* [P11; 2021; cut down smoking]

Discussion

Principal Findings

Using a qualitative approach, this study aimed to explore smokers' experiences and engagement with a quick-response chatbot implemented within a popular smoking cessation app. Users' experiences with the chatbot were largely positive. Anthropomorphism of the chatbot (ie, ascribing human-like characteristics, thoughts, and behavior to the chatbot) was enabled by its specific interaction style and format (eg, positive message tone and quick and easy-to-complete check-ins) and users' perceived need for support, which appeared to give rise to feelings of accountability to the chatbot and increased engagement. Our results build on and extend previous qualitative findings pertaining to users' experiences with companion chatbots [20,29] to a chatbot specifically designed to support smoking cessation.

A previous experimental study has shown that social responses to computers—that is, a direct consequence of anthropomorphism—are common and relatively *easy* to generate (eg, by providing the computer with human-like attributes, such as a language output) [40]. However, according to the *uncanny valley hypothesis*, chatbot designers have a fine line to tread [41]. Strong feelings of affinity are generated by great human-like qualities in a robot or computer program up to a point. Once it becomes very similar to or indistinguishable from a real human, people's reactions can reverse because they may find them *creepy* or *eerie* [41]. Although the quick-response Quit Coach in this study was relatively simple (eg, it did not allow many free-text inputs from users) and made it clear to users that it is an automated bot, the presence of social cues (eg, its positive message tone and communication format similar to

text messaging with GIFs and emojis) appeared to be sufficient for generating social responses from many users without entering the territory of the uncanny valley. This is positive, as it implies that simple (and relatively low-cost) chatbots generate feelings of affinity and that more complex chatbots that can more closely mimic human interactions may not be necessary.

However, users mentioned that they would have liked it if Quit Coach tailored its questions and responses to their unique situations and momentary needs. Therefore, future studies would benefit from exploring—for example, through user-centered design activities and experimental studies—additional design elements that can enhance Quit Coach's similarity to humans, as this may further promote user engagement. In addition, some design elements appeared to detract from users' anthropomorphism of and engagement with Quit Coach, such as its repetitive questions and responses and forced-choice interactions. For users who struggled to stay quit, the repetitiveness and inflexibility of Quit Coach appeared particularly salient, sometimes leading to a reversal of the anthropomorphic experience. Previous studies indicate that users' perceived need for support and the target behavior itself (eg, progress toward smoking cessation) are important for continued engagement [9,10,15]. Similarly, high perceived need for support may be important for users to *suspend disbelief* and anthropomorphize conversational agents within the health care domain. The *Three-Factor Theory of Anthropomorphism* predicts that people are more likely to anthropomorphize nonhuman agents or objects (1) when anthropocentric knowledge is readily accessible and applicable (ie, when knowledge about how humans interact, think, and feel is judged as relevant for the interaction), (2) when motivated to be effective social agents (ie, motivation to master one's environment by increasing its predictability and controllability), and (3) when lacking a sense of social connection to other humans (ie, feeling lonely or isolated) [42]. Future studies would benefit from building on and empirically testing such a theory of anthropomorphism within the human-computer interaction domain, with a view to improving the design of future conversational and relational agents for health and well-being.

Our findings also lend partial support to the Model of Supportive Accountability [22] in that users reported feeling accountable to checking in and updating the nonjudgmental and supportive Quit Coach. Finding a balance between nonjudgmental tone of voice and human-like social cues to generate feelings of affinity and accountability (as discussed previously) may be important for future behavior change chatbots. However, trust and competence (which are additional cornerstones of the Model of Supportive Accountability) were largely missing from users' accounts. It is plausible that competence (eg, legitimacy of the information provided) and trust (eg, data security and confidentiality) were already assumed by users who had either voluntarily downloaded the Smoke Free app from a digital marketplace—likely selecting an app they trusted among the myriad of available apps [15]—or were asked to download it based on recommendation from university researchers. Alternatively, trust and competence may not be necessary conditions for supportive accountability to arise within human-chatbot relationships; this should be further explored in future studies.

Strengths and Limitations

This study was strengthened by recruiting both experienced and novice app users, having a second coder to help in validating the coding, using external validation to ensure that the researchers' interpretations aligned with participants' narratives, and achieving theoretical saturation. However, this study also had several limitations. First, self-selection bias may limit the applicability of the findings to other populations and settings. For example, the 36% (5/14) of participants recruited in 2020 had quit smoking successfully and were considered as *heavy* Quit Coach users, and most participants were young (ie, aged 18-44 years). Second, we did not record any additional support used by participants during their quit attempts (eg, pharmacological support), which may have influenced their perceived need for support. Third, for pragmatic purposes, we did not record participants' actual engagement with Quit Coach, but instead relied on self-reports. Going forward, triangulation of qualitative and quantitative findings would be an important addition to the research literature. Fourth, the study was conducted during the COVID-19 pandemic—a time of significant change in people's life and work conditions, including their smoking behavior [43]—which may also limit the applicability of our findings to other periods and contexts.

Implications for Research and Practice

Findings from this study have both theoretical and practical implications. First, our results indicate that the Model of Supportive Accountability [22] may usefully be extended from human to *human-like* support within digital interventions. However, future studies should further explore the specific conditions under which chatbot interactions lead to feelings of accountability and whether accountability is more easily generated within human-to-human (rather than human-to-bot) interactions. Second, our findings suggest that chatbots for smoking cessation may benefit from including more variation in conversations to prevent boredom and incorporating different levels of tailoring (including context-sensitive tailoring).

Conclusions

Anthropomorphism of a quick-response chatbot implemented within a popular smoking cessation app appeared to be enabled by its interaction style and format (eg, positive message tone and quick check-ins) and users' perceived need for support, which may have given rise to feelings of accountability and increased engagement.

Acknowledgments

OP receives salary support from Cancer Research UK (C1417/A22962). AA, KS, JB, and OP are members of Shaping Public Health Policies to Reduce Inequalities and Harm—a UK Prevention Research Partnership Consortium (MR/S037519/1). The authors gratefully acknowledge the funding.

Data Availability

Participants did not provide consent for their anonymized transcripts to be shared openly. Therefore, the qualitative data underpinning the analyses are available to bona fide researchers upon reasonable request.

Conflicts of Interest

JB has received unrestricted funding to study smoking cessation outside the submitted study from Johnson & Johnson and Pfizer, who manufacture smoking cessation medications. JB and OP are unpaid members of the scientific advisory board for the Smoke Free app.

Multimedia Appendix 1
Interview topic guide, additional screenshots of Quit Coach, and additional participant quotations.
[DOCX File , 484 KB - formative_v6i7e36869_app1.docx]

References

1. WHO report on the global tobacco epidemic 2021: addressing new and emerging products Internet. World Health Organization. URL: https://www.who.int/teams/health-promotion/tobacco-control/global-tobacco-report-2021 [accessed 2022-01-28]
2. Goodchild M, Nargis N, Tursan d'Espaignet E. Global economic cost of smoking-attributable diseases. Tob Control 2018 Jan;27(1):58-64 [FREE Full text] [doi: 10.1136/tobaccocontrol-2016-053305] [Medline: 28138063]
3. Iacobucci G. Stop smoking services: BMJ analysis shows how councils are stubbing them out. BMJ 2018 Aug 24;362:k3649. [doi: 10.1136/bmj.k3649] [Medline: 30143515]
4. Cox S, Ward E, Ross L, Notley C. How a sample of English stop smoking services and vape shops adapted during the early COVID-19 pandemic: a mixed-methods cross-sectional survey. Harm Reduct J 2021 Aug 31;18(1):95 [FREE Full text] [doi: 10.1186/s12954-021-00541-0] [Medline: 34465346]
5. Barnett A, Ding H, Hay KE, Yang IA, Bowman RV, Fong KM, et al. The effectiveness of smartphone applications to aid smoking cessation: a meta-analysis. Clinical eHealth 2020;3:69-81. [doi: 10.1016/j.ceh.2020.09.001]
6. Herbec A, Shahab L, Brown J, Ubhi H, Beard E, Matei A, et al. Does addition of craving management tools in a stop smoking app improve quit rates among adult smokers? Results from BupaQuit pragmatic pilot randomised controlled trial. PsyArXiv. URL: https://psyarxiv.com/d5kjq/ [accessed 2022-01-28]
7. Ubhi HK, Michie S, Kotz D, Wong WC, West R. A mobile app to aid smoking cessation: preliminary evaluation of SmokeFree28. J Med Internet Res 2015 Jan 16;17(1):e17 [FREE Full text] [doi: 10.2196/jmir.3479] [Medline: 25596170]
8. Bricker JB, Watson NL, Mull KE, Sullivan BM, Heffner JL. Efficacy of smartphone applications for smoking cessation: a randomized clinical trial. JAMA Intern Med 2020 Nov 01;180(11):1472-1480 [FREE Full text] [doi: 10.1001/jamainternmed.2020.4055] [Medline: 32955554]
9. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
10. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the uptake of and engagement with health and well-being smartphone apps: systematic review. J Med Internet Res 2020 May 29;22(5):e17572 [FREE Full text] [doi: 10.2196/17572] [Medline: 32348255]
11. 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]
12. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 05;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
13. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008 Nov 04;10(5):e36 [FREE Full text] [doi: 10.2196/jmir.1002] [Medline: 18984557]
14. Zeng EY, Heffner JL, Copeland WK, Mull KE, Bricker JB. Get with the program: adherence to a smartphone app for smoking cessation. Addict Behav 2016 Dec;63:120-124 [FREE Full text] [doi: 10.1016/j.addbeh.2016.07.007] [Medline: 27454354]

15. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Feb 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]
16. Rajani NB, Weth D, Mastellos N, Filippidis FT. Use of gamification strategies and tactics in mobile applications for smoking cessation: a review of the UK mobile app market. BMJ Open 2019 Jun 17;9(6):e027883 [FREE Full text] [doi: 10.1136/bmjopen-2018-027883] [Medline: 31213452]
17. Ubhi HK, Michie S, Kotz D, van Schayck OC, 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 23;6(3):410-417 [FREE Full text] [doi: 10.1007/s13142-015-0352-x] [Medline: 27528530]
18. Rapp A, Curti L, Boldi A. The human side of human-chatbot interaction: a systematic literature review of ten years of research on text-based chatbots. Int J Human Comput Stud 2021 Jul;151:102630. [doi: 10.1016/j.ijhcs.2021.102630]
19. Bickmore TW, Mitchell SE, Jack BW, Paasche-Orlow MK, Pfeifer LM, Odonnell J. Response to a relational agent by hospital patients with depressive symptoms. Interact Comput 2010 Jul 01;22(4):289-298 [FREE Full text] [doi: 10.1016/j.intcom.2009.12.001] [Medline: 20628581]
20. Skjuve M, Følstad A, Fostervold KI, Brandtzaeg PB. My Chatbot Companion - a study of human-chatbot relationships. Int J Human Comput Stud 2021 May;149:102601. [doi: 10.1016/j.ijhcs.2021.102601]
21. Rich response messages. Google Cloud. URL: https://cloud.google.com/dialogflow/es/docs/intents-rich-messages [accessed 2022-01-28]
22. 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]
23. 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]
24. Oh YJ, Zhang J, Fang M, Fukuoka Y. A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. Int J Behav Nutr Phys Act 2021 Dec 11;18(1):160 [FREE Full text] [doi: 10.1186/s12966-021-01224-6] [Medline: 34895247]
25. Almusharraf F, Rose J, Selby P. Engaging unmotivated smokers to move toward quitting: design of motivational interviewing-based chatbot through iterative interactions. J Med Internet Res 2020 Nov 03;22(11):e20251 [FREE Full text] [doi: 10.2196/20251] [Medline: 33141095]
26. Masaki K, Tateno H, Kameyama N, Morino E, Watanabe R, Sekine K, et al. Impact of a novel smartphone app (CureApp smoking cessation) on nicotine dependence: prospective single-arm interventional pilot study. JMIR Mhealth Uhealth 2019 Feb 19;7(2):e12694 [FREE Full text] [doi: 10.2196/12694] [Medline: 30777848]
27. Perski O, Crane D, Beard E, Brown J. Does the addition of a supportive chatbot promote user engagement with a smoking cessation app? An experimental study. Digital Health 2019 Sep 30;5:205520761988067 [FREE Full text] [doi: 10.1177/2055207619880676]
28. Prochaska JJ, Vogel EA, Chieng A, Kendra M, Baiocchi M, Pajarito S, et al. A therapeutic relational agent for reducing problematic substance use (Woebot): development and usability study. J Med Internet Res 2021 Mar 23;23(3):e24850 [FREE Full text] [doi: 10.2196/24850] [Medline: 33755028]
29. 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]
30. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
31. Schwandt T. Constructivist, interpretivist approaches to human inquiry. In: Denzin NK, Lincoln YS, editors. Handbook of Qualitative Research. London: SAGE Publications; 1994:118-137.
32. Jick TD. Mixing qualitative and quantitative methods: triangulation in action. Administrative Sci Q 1979 Dec;24(4):602. [doi: 10.2307/2392366]
33. 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]
34. Kotz D, Brown J, West R. Predictive validity of the Motivation To Stop Scale (MTSS): a single-item measure of motivation to stop smoking. Drug Alcohol Depend 2013 Feb 01;128(1-2):15-19 [FREE Full text] [doi: 10.1016/j.drugalcdep.2012.07.012] [Medline: 22943961]
35. Smoke Free homepage. Smoke Free. URL: https://smokefreeapp.com/ [accessed 2022-01-28]
36. Michie S, Hyder N, Walia A, West R. Development of a taxonomy of behaviour change techniques used in individual behavioural support for smoking cessation. Addict Behav 2011 Apr;36(4):315-319. [doi: 10.1016/j.addbeh.2010.11.016] [Medline: 21215528]

37. Jackson SE, Perski O, Crane D, Michie S, West R, Brown J. Effectiveness of an offer of the Smoke Free smartphone application for smoking cessation: protocol for a randomized controlled trial. Addiction 2019 Nov 20;114(11):2078-2086. [doi: 10.1111/add.14652] [Medline: 31083767]
38. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101 [FREE Full text]
39. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exercise Health 2019 Jun 13;11(4):589-597. [doi: 10.1080/2159676x.2019.1628806]
40. Nass C, Steuer J, Tauber E. Computers are social actors. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1994 Presented at: CHI94: ACM Conference on Human Factors in Computer Systems; Apr 24 - 28, 1994; Boston Massachusetts USA. [doi: 10.1145/191666.191703]
41. Złotowski JA, Sumioka H, Nishio S, Glas DF, Bartneck C, Ishiguro H. Persistence of the uncanny valley: the influence of repeated interactions and a robot's attitude on its perception. Front Psychol 2015 Jun 30;6:883 [FREE Full text] [doi: 10.3389/fpsyg.2015.00883] [Medline: 26175702]
42. Epley N, Waytz A, Cacioppo JT. On seeing human: a three-factor theory of anthropomorphism. Psychol Rev 2007 Oct;114(4):864-886. [doi: 10.1037/0033-295X.114.4.864] [Medline: 17907867]
43. Jackson SE, Garnett C, Shahab L, Oldham M, Brown J. Association of the COVID-19 lockdown with smoking, drinking and attempts to quit in England: an analysis of 2019-20 data. Addiction 2021 May;116(5):1233-1244 [FREE Full text] [doi: 10.1111/add.15295] [Medline: 33089562]

Abbreviations

GIF: Graphics Interchange Format

Edited by A Mavragani; submitted 28.01.22; peer-reviewed by J Rose, A Graham, J Heffner; comments to author 14.03.22; revised version received 04.04.22; accepted 05.04.22; published 07.07.22.

Please cite as:
Alphonse A, Stewart K, Brown J, Perski O
Exploring Users' Experiences With a Quick-Response Chatbot Within a Popular Smoking Cessation Smartphone App: Semistructured Interview Study
JMIR Form Res 2022;6(7):e36869
URL: https://formative.jmir.org/2022/7/e36869
doi:10.2196/36869
PMID:35797093

Original Paper

Smartphone Ownership, Smartphone Utilization, and Interest in Using Mental Health Apps to Address Substance Use Disorders: Literature Review and Cross-sectional Survey Study Across Two Sites

Michael Hsu[1], MD; Bianca Martin[1], BA; Saeed Ahmed[2], MD; John Torous[3], MBI, MD; Joji Suzuki[1], MD

[1]Department of Psychiatry, Brigham and Women's Hospital, Boston, MA, United States

[2]West Ridge Center, Rutland Regional Medical Center, Rutland, VT, United States

[3]Digital Psychiatry Division, Department of Psychiatry, Beth Israel Deaconess Medical Center, Boston, MA, United States

Corresponding Author:
Michael Hsu, MD
Department of Psychiatry
Brigham and Women's Hospital
75 Francis Street
Boston, MA, 02115
United States
Phone: 1 617 732 6753
Fax: 1 407 550 7505
Email: mhsu24@emory.edu

Abstract

Background: In recent years, there has been increasing interest in implementing digital technologies to diagnose, monitor, and intervene in substance use disorders. Smartphones are now a vehicle for facilitating telepsychiatry visits, measuring health metrics, and communicating with health care professionals. In light of the COVID-19 pandemic and the movement toward web-based and hybrid clinic visits and meetings, it has become especially salient to assess phone ownership among individuals with substance use disorders and their comfort in navigating phone functionality and using phones for mental health purposes.

Objective: The aims of this study were to summarize the current literature around smartphone ownership, smartphone utilization, and the acceptability of using smartphones for mental health purposes and assess these variables across two disparate substance use treatment sites.

Methods: We performed a focused literature review via a search of two academic databases (PubMed and Google Scholar) for publications since 2007 on the topics of smartphone ownership, smartphone utilization, and the acceptability of using mobile apps for mental health purposes among the substance use population. Additionally, we conducted a cross-sectional survey study that included 51 participants across two sites in New England—an inpatient detoxification unit that predominantly treats patients with alcohol use disorder and an outpatient methadone maintenance treatment clinic.

Results: Prior studies indicated that mobile phone ownership among the substance use population between 2013 and 2019 ranged from 83% to 94%, while smartphone ownership ranged from 57% to 94%. The results from our study across the two sites indicated 96% (49/51) mobile phone ownership and 92% (47/51) smartphone ownership among the substance use population. Although most (43/49, 88%) patients across both sites reported currently using apps on their phone, a minority (19/48, 40%) reported previously using any apps for mental health purposes. More than half of the participants reported feeling at least neutrally comfortable with a mental health app gathering information regarding appointment reminders (32/48, 67%), medication reminders (33/48, 69%), and symptom surveys (26/45, 58%). Most patients were concerned about privacy (34/51, 67%) and felt uncomfortable with an app gathering location (29/47, 62%) and social (27/47, 57%) information for health care purposes.

Conclusions: The majority of respondents reported owning a mobile phone (49/51, 96%) and smartphone (47/51, 92%), consistent with prior studies. Many respondents felt comfortable with mental health apps gathering most forms of personal information and with communicating with their clinician about their mental health. The differential results from the two sites, namely greater concerns about the cost of mental health apps among the methadone maintenance treatment cohort and less experience with downloading apps among the older inpatient detoxification cohort, may indicate that clinicians should tailor technological

interventions based on local demographics and practice sites and that there is likely not a one-size-fits-all digital psychiatry solution.

(JMIR Form Res 2022;6(7):e38684) doi:10.2196/38684

KEYWORDS

smartphone; mobile phone; addiction; substance use; phone ownership; health equity; digital psychiatry; digital phenotyping; phone applications; substance abuse; mHealth; phone utilization; mental health; mindfulness; digital mental health

Introduction

Handheld phones evolved at a lightning pace over the past 2 decades. Once devices that were solely used to text or call, smartphones now connect millions through social media, track health metrics, and have GPS-sensing capabilities. Smartphones also have a burgeoning role in telepsychiatry, which has been widely adopted during the COVID-19 pandemic. Tens of thousands of mental health apps are available through app stores. Digital phenotyping, which involves using passively and continuously collected sensory and user data to track movement, phone utilization, and communication, has potential apps for relapse prediction in schizophrenia [1], relapse prediction in bipolar disorder [2], and depressive episode detection [3].

Smartphone apps have expanded and can be used for the diagnosis, monitoring, and treatment of substance use disorders [4]. For instance, Reset-O is the first US Food and Drug Administration–approved prescription mobile app for the treatment of opioid use disorder with evidence of improving abstinence and treatment retention [5]. Additionally, a smartphone app may be able to detect an opioid overdose by using a short-range active sonar [6]. Further, A-CHESS (Addiction-Comprehensive Health Enhancement Support System) is a mobile app suite that buttresses alcohol recovery through features such as delivering alerts to patients when they approach known triggers, such as a bar or liquor store, via GPS [7].

With smartphones playing an increasingly integral role in the delivery of substance use care, addressing equity and understanding the adoption and utilization of smartphones and the acceptability of related technologies have become ever more pressing. We performed a nonsystematic review of literature around smartphone ownership and utilization among individuals with substance use disorders, and we present original data from our 2-site, cross-sectional survey study assessing smartphone ownership, smartphone utilization, and the acceptability of using mental health apps. We hypothesize that a vast majority of individuals own smartphones and would be open to using smartphone apps to address substance use. The aim of this study is to inform clinicians who plan to develop or implement digital interventions for patients with substance use disorders on how to anticipate and optimize adoption and engagement.

Methods

Literature Review

The purpose of the literature review was to identify articles that assessed smartphone or phone ownership and/or utilization to answer the following questions: what proportion of patients with substance use disorders own mobile phones or smartphones, how did they utilize their phones, and how open were they to mobile health (mHealth) intervention through smartphones?

A nonsystematic search was conducted in January 2022 within PubMed and Google Scholar, using the key terms *smartphone ownership*, *phone ownership*, *access to a mobile phone*, *access to a smartphone*, *substance use disorder*, *addiction*, and *substance abuse*, by two independent researchers. The selected publication types included primary survey or questionnaire studies, systematic or nonsystematic reviews, and secondary research studies based on prior survey or census data published in or after 2007 (the year that the Apple iPhone was introduced). Other inclusion criteria included articles written in the English language and participants with substance use disorders. The exclusion criteria included clinical trials of specific digital interventions, such as a smartphone app, as the purpose of this study was to understand general smartphone ownership and use. Other exclusion criteria included papers written in foreign languages and studies on participants with behavioral addictions, such as internet gaming disorder, internet addiction, or gambling disorder, as opposed to substance use disorders. A total of 8 abstracts were identified by using these criteria. Study purposes, study designs, descriptions of methods, sample sizes, sample demographics, smartphone or phone ownership rates, and smartphone utilization and acceptability data were extracted.

Cross-sectional, 2-Site Survey Study

We developed a survey that closely mirrored the one developed by Torous et al [8] (Textbox 1). The 5- to 10-minute survey was completed via paper and pencil and included sections on demographics, phone ownership, phone utilization, and the acceptability of using phones for mental health care purposes. Surveys were administered between June 2021 and February 2022 at two sites—a level 4 inpatient detoxification unit at a community hospital in Massachusetts (Brigham and Women's Faulkner Hospital [BWFH]) that predominantly treats individuals with alcohol use disorder and an outpatient methadone clinic in Rutland, Vermont (West Ridge Clinic). Participants were offered a US $5 *Dunkin' Donuts* gift card for completing this study. There was a total of 51 participants from both sites—22 from the inpatient detoxification unit and 29 from the outpatient methadone clinic. Descriptive statistics were used to summarize the data; all analyses were performed by using Stata Statistical Software: Release 17 (StataCorp LLC). Demographic and mobile phone use variables for those at the detoxification unit and those at the outpatient methadone clinic were compared by using a chi-square test for categorical variables and a 2-tailed Student *t* test for continuous variables.

Textbox 1. Survey questions. Demographic questions are excluded.

Questions

1. Do you currently own any type of phone?
   - If no, for what reason do you not have a phone?
2. Is your phone a smartphone (eg, iPhone, Android, etc)?
3. What is the name of your smartphone (eg, Samsung Galaxy)?
4. How comfortable are you sending text messages on your phone?
5. What type of payment plan do you use for text messages?
6. Do you download apps onto your phone?
7. Have you ever downloaded an app for your mental health?
8. Do you currently use any apps for your phone?
9. How comfortable or uncomfortable would you feel about a mental health app gathering and/or sending the following data from your smartphone to your clinician in the context of your care?
   - Appointment reminders
   - Medication reminders
   - Symptom surveys (eg, survey questions about your mood or thoughts throughout the day)
   - Your location (phone GPS sensor)
   - Your social information (call and text logs without any phone numbers or context of messages; eg, how many people you called and for how long)
   - Coaching for healthy living (eg, exercise, sleep, and diet)
   - Mindfulness or therapy exercises
   - Communicating with my clinician about my mental health
10. Select up to 3 top concerns you may have about mental health apps or apps for substance use disorders:
    - Privacy
    - Accuracy of recommendations from app
    - Hard to use
    - Sharing information with clinician
    - Cost
    - Time
    - Hard to set up
11. Select up to 3 top benefits you may see in mental health apps or apps for substance use disorders:
    - Privacy
    - Accuracy of recommendations from app
    - Easy to use
    - Sharing information with clinician
    - Cost
    - Time
    - Easy to set up

Ethical Considerations

This study was approved and monitored by the Mass General Brigham Institutional Review Board (approval number 2020P001656) and Rutland Regional Hospital Institutional Review Board (approval number 2020P001656/RRMC24). This study was identified as human subjects research and was classified as exempt by the Mass General Brigham Institutional Review Board, given the minimal risk to subjects and the use of a survey tool with no identifiable information obtained. This study adheres to the ethical guidelines set forth by the Mass

General Brigham Human Research Protection Program [9]. Participants were consented prior to the survey study and were given the choice to opt out of specific questions or this study entirely at any time. No personal health or identifiable information was recorded.

Results

Participant Characteristics

For the inpatient detoxification cohort, the study team approached 35 participants, of whom 25 consented to answering the survey and 22 completed it (response rate: 22/35, 63%). Further, 2 participants were discharged from detoxification prior to completion, and 1 participant dropped out.

The BWFH inpatient detoxification sample skewed toward male (12/22, 55%) and White (20/22, 91%) patients, and the West Ridge Clinic sample skewed toward female (18/29, 62%) and White (23/29, 79%) patients (Table 1). Notably, the West Ridge Clinic had a relatively higher proportion of individuals experiencing homelessness (10/28, 36%) compared to that of the inpatient detoxification clinic (1/22, 5%), though this did not reach statistical significance (P=.10). There were no additional differences in demographics between the two clinic sites except for the fact that the mean age at the methadone maintenance treatment (MMT) clinic was significantly lower (48.71 vs 36.04 years; P<.001).

Table 1. Summary of participant demographic characteristics (N=51).

Variable	All participants (N=51)	BWFH[a] inpatient detoxification clinic (n=22)	West Ridge Clinic (n=29)	*P* value
Age (years), mean (SD)	41.47 (13.0)	48.71 (11.8)	36.04 (11.2)	.001
Sex, n (%)				.24
Male	23 (45)	12 (55)	11 (38)	
Female	28 (55)	10 (45)	18 (62)	
Race/ethnicity, n (%)				.32
Black or African American	1 (2)	0 (0)	1 (4)	
White	43 (84)	20 (91)	23 (79)	
Hispanic or Latinx	3 (6)	2 (9)	1 (4)	
Other	3 (6)	0 (0)	3 (10)	
Alaska Native	1 (2)	0 (0)	1 (4)	
Education (highest level completed), n (%)				.20
Completed high school or General Educational Development	13 (25)	5 (23)	8 (28)	
Some high school	7 (14)	1 (4)	6 (21)	
Completed college or associate degree	6 (12)	3 (14)	3 (10)	
Some college or associate degree	17 (33)	7 (32)	10 (34)	
Graduate school	8 (16)	6 (27)	2 (7)	
Work, n (%)				.18
Full-time employment	12 (23)	4 (18)	8 (28)	
Part-time employment	5 (10)	1 (5)	4 (14)	
Unemployed	18 (35)	7 (32)	11 (38)	
SSD[b] or SSI[c]	9 (18)	4 (18)	5 (17)	
Other	3 (6)	2 (9)	1 (3)	
Retired	4 (8)	4 (18)	0 (0)	
Residence (n=50)[d], n (%)				.10
Own or rent apartment	18 (36)	10 (45)	8 (29)	
Family or friends	10 (20)	6 (27)	4 (14)	
Single room occupancy	8 (16)	4 (18)	4 (14)	
Halfway house	3 (6)	1 (5)	2 (7)	
Homeless	11 (22)	1 (5)	10 (36)	

[a]BWFH: Brigham and Women's Faulkner Hospital.

[b]SSD: Social Security Disability.

[c]SSI: Supplemental Security Income.

[d]One participant did not answer the question regarding place of residence.

Phone Ownership in the Substance Use Population

Prior studies indicated that mobile phone ownership among the substance use population from 2013 to 2019 ranged from 83% to 94%, while smartphone ownership ranged from 57% to 94% [10-17] (Table 2). Individuals of the baby boomer generation (aged >52 years) may be less likely to own a mobile phone with app capability when compared to Generation X (age: range 36 to 51 years; *P*=.001; odds ratio 3.52, 95% CI 1.65-7.52) and millennials (age: range 18 to 35; *P*<.001; odds ratio 4.53, 95% CI 2.19-9.35) [14].

In our study, a large proportion of respondents reported owning a mobile phone (49/51, 96%) and smartphone (47/51, 92%). All patients (22/22, 100%) at the inpatient detoxification clinic reported owning a mobile phone, while 27 of 29 patients (93%) reported owning a mobile phone at the outpatient methadone clinic. Of the 49 respondents who owned mobile phones across both sites, 47 (96%) categorized them as smartphones; 2 participants opted out of this question.

Table 2. Mobile phone and smartphone ownership among individuals with substance use disorders across studies.

Authors, year	Patient population	Sample size, N	Mobile phone ownership, %	Smartphone ownership, %
McClure et al, 2013 [16]	Adult patients who were undergoing substance abuse treatment and were enrolled at 8 drug-free psychosocial or opioid-replacement therapy clinics in Baltimore	266	91	N/A[a]
Dahne and Lejuez, 2015 [12]	Adult patients admitted to a residential substance use treatment center in Washington, District of Columbia	251	86.9	68.5
Tofighi et al, 2015 [17]	Adult patients with opiate dependence in an urban, safety-net office–based buprenorphine program in New York City	71	93	63
Milward et al, 2015 [13]	Patients enrolled in 4 UK community drug treatment services (74% were undergoing treatment for heroin addiction)	398	83	57
Masson et al, 2018 [11]	Adult patients enrolled in methadone maintenance treatment in San Francisco	178	87	N/A
Ashford et al, 2018 [14]	Adult patients in 4 intensive outpatient substance use disorder treatment facilities in Philadelphia	259	93.8	64.1
Curtis et al, 2019 [10]	Adolescents (aged 13-17 years) and emerging adults (aged 18-35 years) engaged in outpatient substance use treatment in the Southwest and Northeast regions of the United States	164	92.2	80.9
Tofighi et al, 2019 [15]	Adult patients enrolled in an inpatient detoxification program at a safety-net tertiary referral center in New York City	206	86	66

[a]N/A: not applicable.

Phone Utilization in the Substance Use Population

Based on our review of the literature, 79% to 96% of individuals with substance use disorders have phones with text messaging capabilities [12,15,16], with one study published in 2015 suggesting that 55% of mobile phone owners use their phones to text daily [13]. Between 61% to 68% of individuals use their mobile phones to access the internet [10,12,15], and 61.3% of adult smartphone owners use mobile apps on their phones [12]. One study found that of individuals who accessed the internet, 80% accessed it primarily through their mobile phones, with no generational differences in terms of using phones to access the internet versus other means [14]. Masson et al [11] concluded that 40% (n=70) of their participants used their mobile devices as a reminder to take medications, while 8% (n=13) of smartphone users utilized a medication reminder app. McClure et al [16] found that 44% of their adult participants reported being called by substance use clinic staff, while 0.4% reported being texted by them.

Patients were also asked about their mobile phone use patterns in our study (Table 3). Nearly all patients across both sites (45/49, 92%) reported feeling extremely, very, or somewhat comfortable with sending text messages. Of the 47 patients who reported their type of text message payment plan, most (42/47, 89%) reported paying a flat fee for unlimited text messages. The remaining 5 patients reported either paying a flat fee for a limited text message plan (2/47, 4%) or using a pay-per-text plan (3/47, 6%). The majority of respondents (43/49, 88%) reported having downloaded apps, although this differed significantly across sites (West Ridge Clinic patients: 26/27, 96%; inpatient detoxification clinic patients: 17/22, 77%; P=.04). Although most patients (43/49, 88%) across both sites reported currently using apps on their phone, only a minority (19/48, 40%) reported having previously used any app for their mental health; this did not differ significantly across sites (P=.21).

Table 3. Mobile phone use patterns (N=51).

Variable	All participants (N=51), n (%)	BWFH[a] inpatient detoxification clinic (n=22), n (%)	West Ridge Clinic (n=29), n (%)	*P* value
Mobile phone ownership	49 (96)	22 (100)	27 (93)	.21
Smartphone (n=47)[b]	47 (100)	20 (100)	27 (100)	
Sending text messages (n=49)[c]				
Extremely, very, or somewhat comfortable	45 (92)	19 (86)	26 (96)	.21
Text message payment plan (n=47)[d]				.30
Flat fee for unlimited text messages	42 (89)	19 (95)	23 (85)	
Flat fee for limited text messages	2 (4)	1 (5)	1 (4)	
Pay-per-text plan	3 (7)	0 (0)	3 (11)	
Downloads apps onto phone (n=49)[e], n (%)	43 (88)	17 (77)	26 (96)	.04
Has downloaded app for mental health	19 (40)	10 (48)	9 (33)	.32
Currently uses any apps on phone	43 (88)	18 (82)	25 (93)	.25
Very, somewhat, or neutrally comfortable with mental health app gathering information (n=48)[f]				
Appointment reminders	32 (67)	11 (52)	21 (78)	.06
Medication reminders	33 (69)	12 (57)	21 (78)	.13
Symptom surveys	26 (58)	11 (52)	15 (63)	.49
Location	18 (38)	7 (33)	11 (42)	.53
Social information	20 (43)	8 (38)	12 (46)	.58
Coaching for healthy living	27 (56)	11 (52)	16 (59)	.63
Mindfulness or therapy exercises	31 (65)	12 (57)	19 (70)	.34
Communicating with clinician about mental health	30 (63)	11 (52)	19 (70)	.20
Perceived concerns about mental health apps				
Privacy	34 (67)	14 (64)	20 (69)	.69
Accuracy of recommendations	12 (24)	4 (18)	8 (28)	.43
Hard to use	9 (18)	5 (23)	4 (14)	.41
Sharing information with clinician	14 (28)	7 (32)	7 (24)	.54
Cost	15 (29)	2 (9)	13 (45)	.006
Time	16 (31)	5 (23)	11 (38)	.25
Hard to set up	11 (22)	7 (32)	4 (14)	.12
Perceived benefits of mental health apps				
Privacy	11 (22)	4 (18)	7 (24)	.61
Accuracy of recommendations	14 (28)	3 (14)	11 (38)	.05
Easy to use	22 (43)	10 (46)	12 (41)	.77
Sharing information with clinician	19 (37)	8 (36)	11 (38)	.91
Cost	8 (16)	2 (9)	6 (21)	.26
Time	20 (39)	8 (36)	12 (41)	.72
Easy to set up	18 (35)	4 (18)	14 (48)	.03

[a]BWFH: Brigham and Women's Faulkner Hospital.

[b]Two participants who reported owning a mobile phone did not provide information about whether it was a smartphone.

[c]Two participants did not answer questions regarding their comfort with sending text messages.

[d]Four participants did not answer questions regarding their current text message payment plan.

[e]Two participants did not answer questions regarding downloading apps onto their phones.

[f]Three participants did not answer questions regarding their comfort with mental health apps gathering personal information.

Acceptability of Using Smartphones for Mental Health Purposes

Based on our review of the literature, 70.4% of adult participants in one study stated that they would use a relapse prevention app [14], though a different study by Curtis et al [10] suggested that 36.9% of millennials and 45.3% of the Generation Z population thought that a mobile phone app would be helpful toward recovery. In another study, 46% of participants found it unacceptable to use geolocation in a smartphone app for health care purposes [13], but 86% of surveyed adults were willing to be contacted via mobile phone by their clinicians, with telephone calls (53%) being the most preferred method when compared to text messages (41%) and letters (41%) [13]. The most preferred frequency of contact was 1 to 2 messages weekly (46%) [13]. Further, 72% of adults surveyed by Ashford et al [14] in 2018 reported that it was acceptable to receive text messages for relapse prevention. Curtis et al [10] found that 28.8% of millennials and 45.3% of the Generation Z population thought that texting could be helpful toward recovery.

Over half of all participants at both sites were at least neutrally comfortable with a mental health app gathering information regarding appointment reminders (32/48, 67%), medication reminders (33/48, 69%), and symptom surveys (26/45, 58%). Most participants also found it acceptable to use mental health apps to engage in coaching for healthy living (27/48, 56%), mindfulness or therapy exercises (31/48, 65%), and communication with their clinician about their mental health (30/48, 62%). Notably, most of our sample expressed concerns about privacy (34/51, 67%) and reported being uncomfortable with an app gathering information about location (29/47, 62%) and social information (27/47, 57%) for health care purposes. The top three noted concerns about using mental health apps were privacy, cost, and time; patients at the outpatient methadone clinic were significantly more likely to perceive cost as a top-three concern (13/29, 45% vs 2/22, 9%; *P*=.006). Overall, the top three perceived benefits of using mental health apps or apps for substance use disorders were ease of use, the ability to share information with clinicians, and time. Patients at the outpatient methadone clinic were also more likely to perceive ease of setup as a benefit of using such apps (14/29, 48% vs 4/22, 18%; *P*=.03). The acceptability results are summarized in Table 3 and Figures 1-3.

Figure 1. Patients' comfort with a mental health app gathering information on smartphone by clinic location. BWFH: Brigham and Women's Faulkner Hospital.

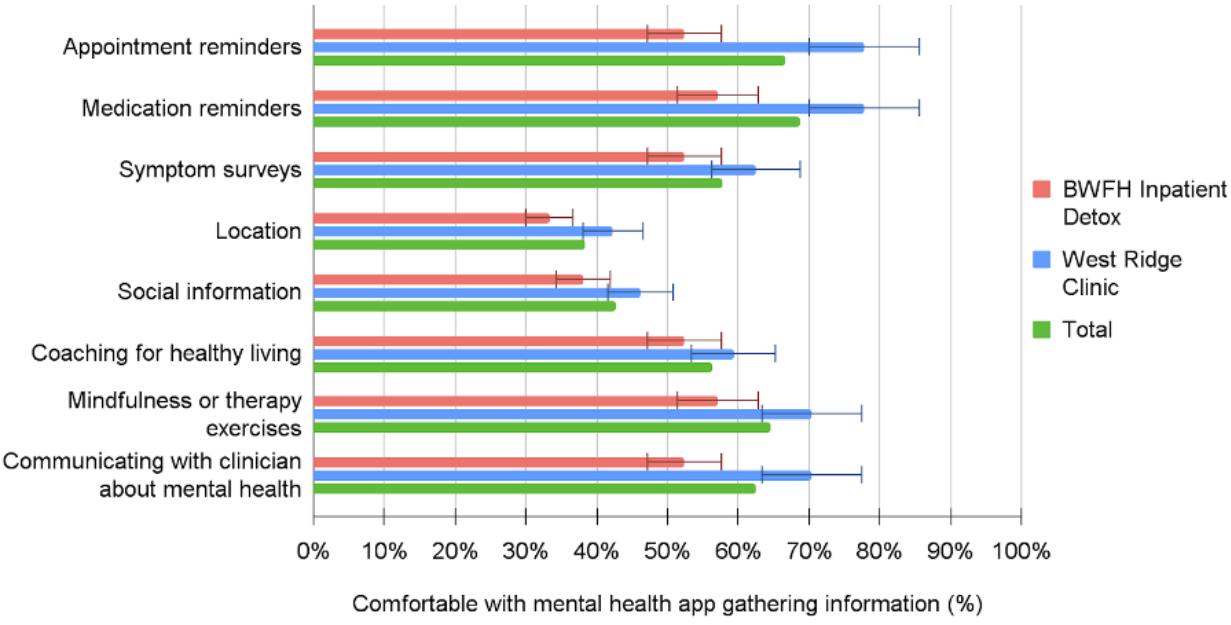


Figure 2. Patients' perceived concerns about mental health apps by clinic location. BWFH: Brigham and Women's Faulkner Hospital.

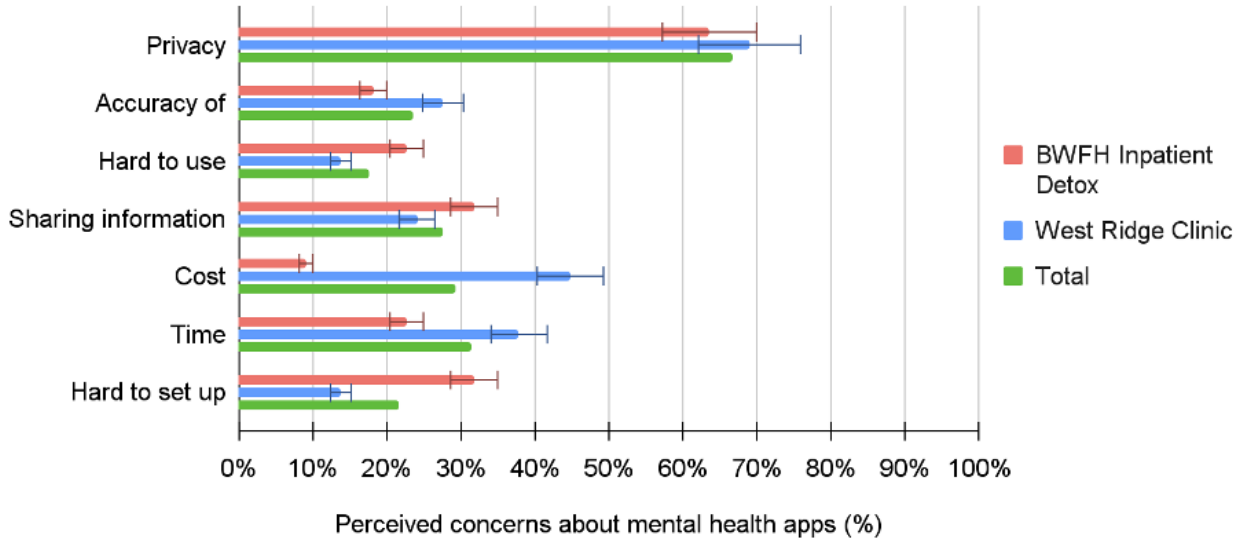
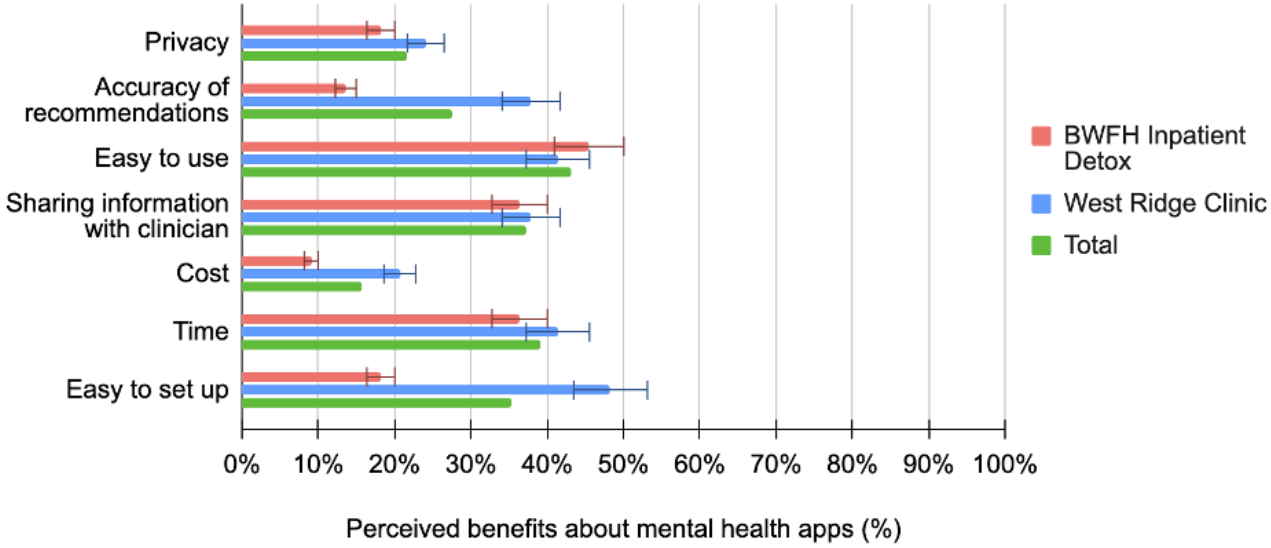


Figure 3. Patients' perceived benefits about mental health apps by clinic location. BWFH: Brigham and Women's Faulkner Hospital.



Discussion

Principal Findings

The overall rate of mobile phone ownership was 96% (49/51), and the overall rate of smartphone ownership was 92% (47/51). The participants recruited at the community inpatient detoxification site were overall older (mean 48.71 vs mean 36.04 years; *P*<.001), and a greater percentage were housed. Participants at the MMT clinic were more likely to report downloading apps when compared to the detoxification sample (26/27, 96% vs 17/22, 77%; *P*=.04), which could be potentially explained by the younger mean age of the methadone cohort (36.04 vs 48.71 years; *P*<.001) [18]. Most patients (43/49, 88%) reported regularly downloading and using apps on their phone, although only 40% (19/48) reported ever downloading an app specifically for mental health or substance use disorder purposes.

A majority of participants across both clinic sites indicated feeling comfortable with mental health apps gathering most forms of personal information, specifically appointment reminders (32/48, 67%), medication reminders (33/48, 69%), symptom surveys (26/45, 58%), coaching for healthy living (27/48, 56%), mindfulness or therapy exercises (31/48, 65%), and communications with their clinician about their mental health (30/48, 62%). Most individuals were uncomfortable with a mental health app tracking location (29/47, 62%) or social information (ie, their call and text logs; 27/47, 57%). The differential views on cost as a barrier to using a mental health app across the two sites (methadone clinic: 13/29, 45%; inpatient detoxification clinic: 2/22, 9%; *P*=.006) might reflect socioeconomic differences across the cohorts. As previously mentioned, the MMT sample had a relatively higher proportion of individuals experiencing homelessness (10/28, 36%) compared to that of the inpatient detoxification sample (1/22, 5%), though the comparative rates of homelessness across the

two samples did not reach statistical significance (P=.10). Most participants cited privacy (34/51, 67%) as a primary concern with using mental health apps and reported that they would not be comfortable with geolocation (29/47, 62%) or social information tracking (27/47, 57%). The participants recruited at the MMT site were significantly more likely to perceive the ease of setting up a mobile app as a perceived benefit when compared to those of the inpatient detoxification unit (14/29, 48% vs 4/22, 18%; P=.03), which may again speak to participants at the MMT clinic being more comfortable with using app functionalities.

Overall, our cross-sectional study suggests that individuals with substance use disorders are generally amenable to using a smartphone app for mental health monitoring or treatment purposes. Interestingly, while smartphone ownership was slightly lower among participants in the MMT site compared to that among the detoxification site participants, which is unsurprising given that the participants at the MMT site were of lower socioeconomic status, our data suggest that the individuals recruited at the MMT site had higher digital literacy, as reflected by their comfort with downloading apps and their perception that ease of use is a benefit of using a mental health app for substance use interventions. In conclusion, clinicians should consider patient demographics, digital literacy, and practice sites when implementing mHealth interventions for substance use disorders in an equitable fashion.

Comparison to Prior Work

To our knowledge, this study represents the first literature review of smartphone ownership, smartphone utilization, and the acceptability of using mHealth among individuals with substance use disorders and the first cross-sectional survey study to address this topic since the beginning of the COVID-19 pandemic. Smartphone and mobile phone ownership rates in our cross-sectional survey study were higher than those reported in all prior studies, likely reflecting the growing adoption of smartphones. Overall rates of downloading apps across both survey sites (43/49, 88%) were also higher than the 61% to 64% of participants who reported downloading mobile apps in a study by Dahne and Lejuez [12], which recruited patients from 2014 to 2015. The proportion of participants who felt uncomfortable with location being tracked (29/47, 62%) was slightly higher than that reported by Milward et al [13]. Comfort with utilizing specific functions of apps for substance use disorders, such as appointment reminders or social functions, and specific perceived benefits of using mobile apps for substance use disorders were not assessed in prior studies of individuals with substance use disorders.

Strengths and Limitations

This study has several limitations. First, we performed a brief, nonsystematic review, and it is likely that relevant papers may not have been included. We attempted to strengthen the robustness of this focused literature review by utilizing two independent reviewers, two separate search engines, and broad key words to capture and screen more abstracts. Future works should incorporate a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)-based systematic review to capture a broader range of studies. Second, the relatively small sample of our cross-sectional survey study precluded our ability to explore the impacts that race, socioeconomic status, age, and other factors have on smartphone ownership, utilization, and acceptability. Third, the predominantly White sample, especially the predominantly White inpatient detoxification cohort, limits the generalizability of this study. However, we recruited from two disparate clinical sites in two very different geographic locations to expand the diversity of recruited participants. Fourth, the degree of selection bias among our outpatient methadone clinic cohort is difficult to assess without ascertaining a survey response rate and consequently may impact the reliability of our results. However, we were able to obtain a survey response rate among individuals at our inpatient detoxification site. Those who turned down the survey at the inpatient detoxification site were asleep, were medically unwell, or were preoccupied at the time of survey distribution.

Future Directions

Future work in this area should include larger patient populations across various sites, which might include non–methadone outpatient substance use clinics. Further, in vivo, randomized controlled studies of promising mental health apps for substance use disorders are needed to establish clinical efficacy. Studies clarifying the effects of socioeconomic status, race, and other factors on digital literacy, smartphone utilization, smartphone ownership, and the acceptability of using apps for substance use interventions among individuals with substance use disorders are needed. Privacy and security concerns around mental health apps will need to be addressed, especially given that individuals with mental health and substance use disorders are particularly vulnerable.

Acknowledgments

We acknowledge Faith Stone from Vermont Blueprint for Health. This study was supported in part by the American Psychiatric Association/Substance Abuse and Mental Health Services Administration Minority Fellowship Program Grant, the Brigham Research Institute Microgrant, and a National Institutes of Health/ National Institute on Drug Abuse grant (K23DA042326).

Data Availability

All data are represented in the tables and figures displayed in this publication. The data sets that were generated and/or analyzed during this study are available from the corresponding author on reasonable request.

Disclaimer

This paper has not been published in and is not in review by another journal, and it has not been presented at any conference.

Authors' Contributions

The final manuscript has been seen and approved by all authors.

Conflicts of Interest

MH is on the scientific advisory board of Healthy Gamer, LLC. JT is a cofounder of Precision Mental Wellness.

References

1. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela JP. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacology 2018 Jul;43(8):1660-1666 [FREE Full text] [doi: 10.1038/s41386-018-0030-z] [Medline: 29511333]
2. Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as an electronic biomarker of illness activity in bipolar disorder. Bipolar Disord 2015 Nov;17(7):715-728. [doi: 10.1111/bdi.12332] [Medline: 26395972]
3. De Angel V, Lewis S, White K, Oetzmann C, Leightley D, Oprea E, et al. Digital health tools for the passive monitoring of depression: a systematic review of methods. NPJ Digit Med 2022 Jan 11;5(1):3 [FREE Full text] [doi: 10.1038/s41746-021-00548-8] [Medline: 35017634]
4. Hsu M, Ahern DK, Suzuki J. Digital phenotyping to enhance substance use treatment during the COVID-19 pandemic. JMIR Ment Health 2020 Oct 26;7(10):e21814 [FREE Full text] [doi: 10.2196/21814] [Medline: 33031044]
5. Maricich YA, Bickel WK, Marsch LA, Gatchalian K, Botbyl J, Luderer HF. Safety and efficacy of a prescription digital therapeutic as an adjunct to buprenorphine for treatment of opioid use disorder. Curr Med Res Opin 2021 Feb;37(2):167-173 [FREE Full text] [doi: 10.1080/03007995.2020.1846022] [Medline: 33140994]
6. Nandakumar R, Gollakota S, Sunshine JE. Opioid overdose detection using smartphones. Sci Transl Med 2019 Jan 09;11(474):eaau8914. [doi: 10.1126/scitranslmed.aau8914] [Medline: 30626717]
7. Gustafson DH, McTavish FM, Chih MY, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. JAMA Psychiatry 2014 May;71(5):566-572 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.4642] [Medline: 24671165]
8. Torous J, Chan SR, Tan SYM, Behrens J, Mathew I, Conrad EJ, et al. Patient smartphone ownership and interest in mobile apps to monitor symptoms of mental health conditions: A survey in four geographically distinct psychiatric clinics. JMIR Ment Health 2014 Dec 23;1(1):e5 [FREE Full text] [doi: 10.2196/mental.4004] [Medline: 26543905]
9. Human research protection program policy and guidance. Mass General Brigham. URL: https://www.massgeneralbrigham.org/researcher-support-and-resources/resources-collaborators-and-sponsors/human-research-protection-program-policy-guidance [accessed 2022-06-06]
10. Curtis BL, Ashford RD, Magnuson KI, Ryan-Pettes SR. Comparison of smartphone ownership, social media use, and willingness to use digital interventions between Generation Z and millennials in the treatment of substance use: Cross-sectional questionnaire study. J Med Internet Res 2019 Apr 17;21(4):e13050 [FREE Full text] [doi: 10.2196/13050] [Medline: 30994464]
11. Masson CL, Chen IQ, Levine JA, Shopshire MS, Sorensen JL. Health-related internet use among opioid treatment patients. Addict Behav Rep 2018 Dec 20;9:100157 [FREE Full text] [doi: 10.1016/j.abrep.2018.100157] [Medline: 31193741]
12. Dahne J, Lejuez CW. Smartphone and mobile application utilization prior to and following treatment among individuals enrolled in residential substance use treatment. J Subst Abuse Treat 2015 Nov;58:95-99 [FREE Full text] [doi: 10.1016/j.jsat.2015.06.017] [Medline: 26231698]
13. Milward J, Day E, Wadsworth E, Strang J, Lynskey M. Mobile phone ownership, usage and readiness to use by patients in drug treatment. Drug Alcohol Depend 2015 Jan 01;146:111-115. [doi: 10.1016/j.drugalcdep.2014.11.001] [Medline: 25468818]
14. Ashford RD, Lynch K, Curtis B. Technology and social media use among patients enrolled in outpatient addiction treatment Programs: Cross-sectional survey study. J Med Internet Res 2018 Mar 06;20(3):e84 [FREE Full text] [doi: 10.2196/jmir.9172] [Medline: 29510968]
15. Tofighi B, Leonard N, Greco P, Hadavand A, Acosta MC, Lee JD. Technology use patterns among patients enrolled in inpatient detoxification treatment. J Addict Med 2019;13(4):279-286 [FREE Full text] [doi: 10.1097/ADM.0000000000000494] [Medline: 30589653]
16. McClure EA, Acquavita SP, Harding E, Stitzer ML. Utilization of communication technology by patients enrolled in substance abuse treatment. Drug Alcohol Depend 2013 Apr 01;129(1-2):145-150 [FREE Full text] [doi: 10.1016/j.drugalcdep.2012.10.003] [Medline: 23107600]

17. Tofighi B, Grossman E, Buirkle E, McNeely J, Gourevitch M, Lee JD. Mobile phone use patterns and preferences in safety net office-based buprenorphine patients. J Addict Med 2015;9(3):217-221 [FREE Full text] [doi: 10.1097/ADM.0000000000000121] [Medline: 25918966]
18. U.S. mobile owners who downloaded apps by age 2018. Statista. URL: https://www.statista.com/statistics/243794/us-adult-cell-phone-owners-who-have-downloaded-apps-by-age-group/ [accessed 2022-04-05]

Abbreviations

A-CHESS: Addiction-Comprehensive Health Enhancement Support System
BWFH: Brigham and Women's Faulkner Hospital
mHealth: mobile health
MMT: methadone maintenance treatment
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by T Leung; submitted 13.04.22; peer-reviewed by K Uludag, X Xiong; comments to author 20.05.22; revised version received 10.06.22; accepted 21.06.22; published 07.07.22.

Please cite as:
Hsu M, Martin B, Ahmed S, Torous J, Suzuki J
Smartphone Ownership, Smartphone Utilization, and Interest in Using Mental Health Apps to Address Substance Use Disorders: Literature Review and Cross-sectional Survey Study Across Two Sites
JMIR Form Res 2022;6(7):e38684
URL: https://formative.jmir.org/2022/7/e38684
doi:10.2196/38684
PMID:35797102

Original Paper

Feasibility of Mobile Health and Social Media–Based Interventions for Young Adults With Early Psychosis and Clinical Risk for Psychosis: Survey Study

Olivia H Franco[1], BS; Monica E Calkins[1], PhD; Salvatore Giorgi[2], MS; Lyle H Ungar[2], PhD; Raquel E Gur[1], MD, PhD; Christian G Kohler[1], MD; Sunny X Tang[1,3], MD

[1]Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, United States
[2]Computer and Information Science, University of Pennsylvania, Philadelphia, PA, United States
[3]Feinstein Institutes for Medical Research, Northwell Health, Glen Oaks, NY, United States

Corresponding Author:
Sunny X Tang, MD
Feinstein Institutes for Medical Research
Northwell Health
75-59 263rd St
Glen Oaks, NY, 11004
United States
Phone: 1 631 786 3737
Email: stang3@northwell.edu

Abstract

Background: Digital technology, the internet, and social media are increasingly investigated as promising means for monitoring symptoms and delivering mental health treatment. These apps and interventions have demonstrated preliminary acceptability and feasibility, but previous reports suggest that access to technology may still be limited among individuals with psychotic disorders relative to the general population.

Objective: We evaluated and compared access to and use of technology and social media in young adults with psychotic disorders (PD), young adults with clinical risk for psychosis (CR), and psychosis-free youths (PF).

Methods: Participants were recruited through a coordinated specialty care clinic dedicated toward early psychosis as well as ongoing studies. We surveyed 21 PD, 23 CR, and 15 PF participants regarding access to technology and use of social media, specifically Facebook and Twitter. Statistical analyses were conducted in R. Categorical variables were compared among groups using Fisher exact test, continuous variables were compared using 1-way ANOVA, and multiple linear regressions were used to evaluate for covariates.

Results: Access to technology and social media were similar among PD, CR, and PF participants. Individuals with PD, but not CR, were less likely to post at a weekly or higher frequency compared to PF individuals. We found that decreased active social media posting was unique to psychotic disorders and did not occur with other psychiatric diagnoses or demographic variables. Additionally, variation in age, sex, and White versus non-White race did not affect posting frequency.

Conclusions: For young people with psychosis spectrum disorders, there appears to be no “technology gap” limiting the implementation of digital and mobile health interventions. Active posting to social media was reduced for individuals with psychosis, which may be related to negative symptoms or impairment in social functioning.

(JMIR Form Res 2022;6(7):e30230) doi:10.2196/30230

KEYWORDS

social media; psychosis; clinical high risk; technology; digital health

Introduction

Digital Health in Psychiatry

The increasing availability of technology has opened a new avenue for making health care more accessible to a broader population and for the development and implementation of digital interventions. In recent years, mobile and web-based health interventions have been used with great promise to assess and coordinate with mental health clients [1]. Social media and mobile health interventions present a unique opportunity to engage young people experiencing the emergence of a psychotic disorder (PD) or at clinical risk (CR) for psychosis, when there are clinically significant symptoms of psychosis, but threshold criteria are not yet met. The importance of early intervention and prevention in psychosis is a key factor in long-term success, and it has been shown that decreasing the duration of untreated psychosis increases long-term treatment outcome [2,3].

Mobile and social media interventions and assessment tools are relatively new in psychiatry, but initial studies show promising engagement, efficacy, and prognostic utility [4,5]. For example, a classification model based on social media language was able to identify early relapse signs for people experiencing first episode psychosis [6]. Relapse events were also predicted in another study by changes in smartphone-enabled social activities [7]. Among interventions, patients with PD benefitted from a combined treatment protocol involving monitoring via mobile health interventions and psychosocial services. Digital social networks have been integrated into mobile health apps to facilitate peer support and improve amotivation symptoms [8]. Other promising mobile health interventions include reminders for medication adherence and appointments, interventions, and case management via SMS text messages and apps on a smartphone. Mobile interventions can also vary from standardized communication to live interactions with clinicians; for example, the patient can have live sessions with a therapist via videoconferencing or SMS text messaging, or they can receive standardized communications like prerecorded videos, and tip sheets based on their input at the time of communication [9].

Access to Technology

Access to technology among the target population is critical for the feasibility of mobile and social media–based interventions. Globally, access to technology has increased dramatically across all age groups, but a "technology gap" may exist for individuals at risk for or diagnosed with psychotic disorders [2]. Among adult patients recruited from an inpatient and outpatient clinic, only 48% have access to the internet, and only 27% used social networking sites on a daily basis [10]. However, younger age may be associated with greater access to technology. In one study of young people in their first episode of psychosis, 100% used social media, and 90% of those reported daily use for an average of about 2 hours daily [5]. Another study completed in Spain showed that patients with first episode psychosis had similar interest in but decreased access to digital technologies including computers and smartphones [11]. If young people with psychosis commonly access digital technology and social media, it would provide an avenue for outreach, and could be a method of monitoring symptoms and side effects and providing treatment.

Objective

To our knowledge, no study has directly compared access to and use of technology and social media across the psychosis spectrum among CR, PD, and psychosis-free (PF) youths. Most studies also do not distinguish between passive access of social media and active posting. In this study, we surveyed CR and PD youths and PF comparison participants about their access to digital technology, the internet, and both access and posting to Facebook and Twitter. We hypothesized that the "technology gap" is decreased or absent among young adults with psychosis and clinical risk for psychosis, suggesting that social media and mobile health interventions would be feasible and implementable in this population.

Methods

Participants

For this study, three groups were evaluated: PD, CR, and PF. A total of 59 participants aged 18-32 years were included. PD participants were recruited from the University of Pennsylvania Psychosis Evaluation and Recovery Center, a coordinated specialty care clinic dedicated toward early psychosis. CR and PF participants were recruited among participants of other ongoing studies at the University of Pennsylvania and Children's Hospital of Philadelphia Lifetime Brain Institute. We were interested in the main effect of psychosis on access to technology and use of social media. Because mental health disorders are common in the community for young adults and adolescents [12], our PF comparison group did not exclude people with nonpsychotic disorders.

All stable outpatients at the Psychosis Evaluation and Recovery Center aged 18-35 years as well as potential participants in Lifetime Brain Institute studies were approached and asked for their consent to participate in this survey as a screening procedure for larger studies. Participants all resided in the greater Philadelphia area, which includes both an urban and surrounding suburban environment. There was no remuneration for completing this survey. Several eligible individuals were not interested in research in general, but no one specifically refused to participate in this survey. This was a brief single timepoint assessment and no one withdrew prior to its completion.

Ethics Approval

All procedures were approved by the Institutional, Review Board at the University of Pennsylvania under protocol 831509, and the Children's Hospital of Philadelphia, protocol number 16-013305.

Study Design and Assessments

The survey was completed verbally with a research coordinator either in person or over the telephone. Individuals who agreed to participate were surveyed via a questionnaire regarding their access to technology and use of social media (Facebook and Twitter). Access to technology was defined generally as the ability to use the technology in a dependable manner, including

personal ownership, shared devices, and public access points. Active posting was defined as commenting, posting status updates, or otherwise contributing original content, versus passive access of social media, which includes scrolling, viewing, or liking. We did not distinguish between accessing the internet via Wi-Fi or a data plan.

The PF and CR groups underwent semistructured interviewing with the Structured Interview for Prodromal Syndromes, which assessed threshold and subthreshold symptoms of psychosis. Determination for CR or PF status was made by consensus case conference. The PD group underwent consensus clinical diagnosis by psychiatrists and psychologists specializing in psychotic disorders and were determined to have one of the following Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) psychotic disorders: schizophrenia (n=13), schizoaffective disorder (n=4), bipolar I disorder with psychotic features (n=1), or unspecified psychotic disorder (n=3). Comorbid nonpsychotic disorders were evaluated based on DSM-5 criteria and were also present in all 3 groups, in proportion to expected population rates.

Statistical Analysis

Statistical analyses were conducted in R (version 3.5.2; R Foundation for Statistical Computing). Categorical variables were compared among groups using Fisher exact test, including the main outcomes of technology access and posting frequency versus group. Fisher exact test was also used to compare social media access and posting rates for other diagnoses as well as for sex and race effects. Age differences among the groups and between active versus nonactive users were compared using 1-way ANOVA. To account for group differences in age, sex, and race, we additionally performed logistic regressions predicting social media access and posting with group as the independent variable of interest, and covarying for the demographic variables. Significance was 2-tailed with =.05.

Results

Table 1 displays details for 21 PD, 23 CR, and 15 PF participants. There was uniformly high access to mobile phones (96%-100%, $P>.99$), smartphones (95%-100%, $P>.99$), computers (85%-95%, $P=.84$), and the internet (95%-100%, $P=.61$). The majority of young adults accessed Facebook, but not Twitter. Social media access rates were similar for all 3 groups (62%-74% at least weekly, $P=.73$). However, there was a significant main group effect for social media posting (5%-43% at least weekly, $P<.009$). CR actively posted at a similar rate compared to PF individuals (CR 43% vs PF 27%, $P=.31$), but PD actively posted at a significantly lower rate than the nonpsychotic groups (5%, $P<.01$). When examining Facebook use alone, as this was the platform with the most users, we found consistent results: access was not affected by psychosis group ($P>.99$), but there was a significant main effect for group on posting ($P=.02$). When covarying for sex, age, and race, psychosis diagnosis remained a significant predictor of decreased active social media posting (standardized $\beta=-1.22$, $P=.03$), but it was not a significant predictor for social media access (standardized $\beta=-0.29$, $P=.40$).

Decreased active social media posting was unique to psychotic disorders and there was no group effect for diagnosis with attention-deficit/hyperactivity disorder ($P=.67$), mood ($P=.54$), or anxiety disorders ($P>.99$) when comparing prevalence of weekly or greater use. Variation in age ($P=.63$), sex ($P=.13$), and White versus non-White race ($P=.77$) did not affect posting frequency.

Table 1. Sample characteristics, access to technology, and social media use.

Variable	PF[a] (n=15)	CR[b] (n=23)	PD[c] (n=21)	*P* value
Age (years), mean (SD)	23.5 (4.2)	22.0 (2.9)	24.2 (3.4)	.10
Sex, n (%)				.01
Female	11 (73)	12 (52)	5 (24)	
Male	4 (27)	11 (48)	16 (76)	
Race, n (%)				.09
White	4 (29)	11 (50)	14 (67)	
African American	7 (50)	9 (41)	6 (29)	
Mixed	1 (7)	2 (9)	0 (0)	
Asian	2 (14)	0 (0)	1 (4)	
Other	0 (0)	0 (0)	0 (0)	
Comorbidities, n (%)				
Attention-deficit/hyperactivity disorder	2 (13)	3 (13)	2 (10)	>.99
Anxiety disorder	2 (13)	15 (65)	5 (24)	.002
Mood disorder	3 (20)	12 (52)	7 (33)	.14
Access to technology, n (%)				
Mobile phone	15 (100)	22 (96)	21 (100)	>.99
Smartphone	15 (100)	22 (96)	20 (95)	>.99
Computer	13 (87)	21 (91)	20 (95)	.84
Internet	15 (100)	23 (100)	20 (95)	.61
Social media use, n (%)				
At least weekly access	10 (67)	17 (74)	13 (62)	.73
Facebook	10 (67)	15 (65)	13 (62)	
Twitter	3 (20)	5 (22)	0 (0)	
At least weekly posting	4 (27)	10 (43)	1 (5)	.01
Facebook	4 (27)	9 (39)	1 (5)	
Twitter	2 (13)	3 (13)	0 (0)	
Facebook (ever)	14 (93)	17 (74)	15 (71)	.27
Twitter (ever)	4 (27)	6 (26)	2 (10)	.30

[a]PF: psychosis-free.

[b]CR: clinical risk for psychosis.

[c]PD: psychotic disorder.

Discussion

Principal Findings

We found that youths with PD and CR who participate in our clinical care and research programs have similar access to technology and use social media to a similar degree compared with PF individuals. This supports the feasibility of mobile health and social media interventions in young CR and PD populations in terms of access to technology. Our results were consistent with those found in previous studies, which showed that the use of digital technology in the treatment of people with psychosis spectrum disorder and other chronic serious mental illnesses is a viable method of delivery of services. In our study, the absolute rate of mobile phone and smartphone ownership and internet use was higher than those reported by Young et al [13], likely because of the younger age of our participants. Most previous studies have compared access to technology between mental health clients and published normative data for the general population but have not directly compared across groups using a consistent standardized methodology. One study directly compared the use of (not access to) technology between patients with first episode psychosis and healthy control subjects; they found significant but small to moderate effects of decreased frequency of use for computers, tablets, smartphones, and smart televisions (but not game consoles) [11]. However, it is unclear whether demographic differences may have accounted for some of the disparity. No studies have included a psychosis clinical risk comparison group.

Another relevant consideration is comfort with technology among clinicians. A recent study by Camacho and Torous [14] surveyed 42 Coordinated Specialty Care clinics and found that health care providers were supportive of implementing technology in their care model for early psychosis; although 69% of surveyed staff were confident in their ability to provide technical assistance for others, 78% indicated that additional digital skills training would be beneficial.

We also found a psychosis-specific decrease in active social media posting. Social cognitive impairments and negative symptoms may contribute to this finding. This interpretation is supported by a study by Rehki et al [15] showing that an increase in severity of negative symptoms is associated with a lower likelihood of social media use. In that study, they suggested that social interactions via social media should be considered in the clinical evaluation of individuals with schizophrenia, as it is a prominent form of communication and fostering relationships. Although individuals with CR may also experience significant negative symptoms, our results suggest that these are not reflected in significant changes in social media posting and access rates. Although youths with PD post significantly less frequently than youths with CR and PF, individuals with PD nevertheless access social media at a similar rate, which provides a potential avenue for intervention. To our knowledge, no other study has distinguished between social media access and posting, so we are the first to report a decrease in social media posting for individuals with psychotic disorders while there was no difference in level of passive use of social media. This effect did not appear with mood disorders, attention-deficit/hyperactivity disorder, and anxiety, and therefore appears to be specific to psychosis.

Limitations of this study include sampling of one geographical area and limited sample size. Socioeconomic status may also influence access to technology but was not consistently measured in this sample. It is possible that patients with access to a specialized psychosis treatment center and those who volunteer for research participation may have higher access to technology than the general population of people with psychosis. Moreover, our survey only included two social media platforms, Facebook and Twitter. Future studies may consider including access and use of other social media platforms that have recently become more popular among youths, as well as questions about their willingness to share their digital access and to receive therapeutic interactions via a digital platform. Objective usage statistics and other metadata may also provide useful signals. Privacy risks and concerns may also be technical as well as subjective barriers to implementation of such interventions. These were not addressed with this study.

Conclusions

Overall, our findings encourage further development of mobile health and social media–based interventions and monitoring for young people with psychosis and youths at clinical risk for psychosis. There appears to be no significant “technology gap” for young people with psychotic disorders relative to young people without psychosis. Lower active engagement may reflect impairments in social cognition and functioning. Notably, access to technology does not mean that digital health interventions will be ultimately efficacious or implementable—however, the availability of technology to young people with psychosis does provide some basis for feasibility. Future studies are needed to directly evaluate the efficacy and usability of digital health and social media–based strategies for intervention and assessment, while considering additional potentially harmful effects related to privacy concerns and increased time spent on social media.

Acknowledgments

SXT received funding from the Brain and Behavioral Research Foundation NARSAD Young Investigator Grant and the ASCP Early Career Research Award. We thank the participants and their families for their collaboration. We also thank Lyndsay Schmidt, Zeeshan Huque, LaTonya McCurry, Victoria Pietruszka, Kosha Ruparel, Thomas Hohing, and Eehwa Ung from the University of Pennsylvania.

Conflicts of Interest

SXT is a consultant for Neurocrine Biosciences and North Shore Therapeutics, received funding from Winterlight Labs, and holds equity in North Shore Therapeutics. The other authors have no disclosures to report.

References

1. Batra S, Baker RA, Wang T, Forma F, DiBiasi F, Peters-Strickland T. Digital health technology for use in patients with serious mental illness: a systematic review of the literature. Med Devices (Auckl) 2017;10:237-251 [FREE Full text] [doi: 10.2147/MDER.S144158] [Medline: 29042823]
2. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile Phone Ownership and Endorsement of "mHealth" Among People With Psychosis: A Meta-analysis of Cross-sectional Studies. Schizophr Bull 2016 Mar;42(2):448-455 [FREE Full text] [doi: 10.1093/schbul/sbv132] [Medline: 26400871]
3. Moe AM, Rubinstein EB, Gallagher CJ, Weiss DM, Stewart A, Breitborde NJ. Improving access to specialized care for first-episode psychosis: an ecological model. Risk Manag Healthc Policy 2018;11:127-138 [FREE Full text] [doi: 10.2147/RMHP.S131833] [Medline: 30214330]

4. Camacho E, Levin L, Torous J. Smartphone Apps to Support Coordinated Specialty Care for Prodromal and Early Course Schizophrenia Disorders: Systematic Review. J Med Internet Res 2019 Nov 12;21(11):e16393 [FREE Full text] [doi: 10.2196/16393] [Medline: 31714250]
5. Birnbaum ML, Rizvi AF, Faber K, Addington J, Correll CU, Gerber C, et al. Digital Trajectories to Care in First-Episode Psychosis. Psychiatr Serv 2018 Dec 01;69(12):1259-1263 [FREE Full text] [doi: 10.1176/appi.ps.201800180] [Medline: 30256181]
6. Birnbaum ML, Ernala SK, Rizvi AF, De Choudhury M, Kane JM. A Collaborative Approach to Identifying Social Media Markers of Schizophrenia by Employing Machine Learning and Clinical Appraisals. J Med Internet Res 2017 Aug 14;19(8):e289 [FREE Full text] [doi: 10.2196/jmir.7956] [Medline: 28807891]
7. Buck B, Scherer E, Brian R, Wang R, Wang W, Campbell A, et al. Relationships between smartphone social behavior and relapse in schizophrenia: A preliminary report. Schizophr Res 2019 Jun;208:167-172 [FREE Full text] [doi: 10.1016/j.schres.2019.03.014] [Medline: 30940400]
8. Schlosser DA, Campellone TR, Truong B, Etter K, Vergani S, Komaiko K, et al. Efficacy of PRIME, a Mobile App Intervention Designed to Improve Motivation in Young People With Schizophrenia. Schizophr Bull 2018 Aug 20;44(5):1010-1020 [FREE Full text] [doi: 10.1093/schbul/sby078] [Medline: 29939367]
9. Ben-Zeev D, Brian RM, Aschbrenner KA, Jonathan G, Steingard S. Video-based mobile health interventions for people with schizophrenia: Bringing the "pocket therapist" to life. Psychiatr Rehabil J 2018 Mar;41(1):39-45 [FREE Full text] [doi: 10.1037/prj0000197] [Medline: 27295133]
10. Miller BJ, Stewart A, Schrimsher J, Peeples D, Buckley PF. How connected are people with schizophrenia? Cell phone, computer, email, and social media use. Psychiatry Res 2015 Feb 28;225(3):458-463. [doi: 10.1016/j.psychres.2014.11.067] [Medline: 25563669]
11. Fernández-Sotos P, Fernández-Caballero A, González P, Aparicio AI, Martínez-Gras I, Torio I, et al. Digital Technology for Internet Access by Patients With Early-Stage Schizophrenia in Spain: Multicenter Research Study. J Med Internet Res 2019 Mar 14;21(4):e11824 [FREE Full text] [doi: 10.2196/11824] [Medline: 30950798]
12. Merikangas KR, He J, Burstein M, Swanson SA, Avenevoli S, Cui L, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication--Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry 2010 Oct;49(10):980-989 [FREE Full text] [doi: 10.1016/j.jaac.2010.05.017] [Medline: 20855043]
13. Young AS, Cohen AN, Niv N, Nowlin-Finch N, Oberman RS, Olmos-Ochoa TT, et al. Mobile Phone and Smartphone Use by People With Serious Mental Illness. Psychiatr Serv 2020 Mar 01;71(3):280-283 [FREE Full text] [doi: 10.1176/appi.ps.201900203] [Medline: 31744429]
14. Camacho E, Torous J. Interest and readiness for digital mental health in coordinate specialty care for early course psychosis: A survey study of 42 programs in 30 states. Early Interv Psychiatry 2021 Oct;15(5):1243-1255. [doi: 10.1111/eip.13073] [Medline: 33260266]
15. Rekhi G, Ang MS, Lee J. Clinical determinants of social media use in individuals with schizophrenia. PLoS One 2019;14(11):e0225370 [FREE Full text] [doi: 10.1371/journal.pone.0225370] [Medline: 31747434]

Abbreviations

CR: clinical risk for psychosis
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
PD: psychosis disorder
PF: psychosis-free

Edited by J Torous; submitted 03.06.21; peer-reviewed by R Brian, J Lee, W Buente; comments to author 27.08.21; revised version received 21.03.22; accepted 15.04.22; published 08.07.22.

Please cite as:
Franco OH, Calkins ME, Giorgi S, Ungar LH, Gur RE, Kohler CG, Tang SX
Feasibility of Mobile Health and Social Media–Based Interventions for Young Adults With Early Psychosis and Clinical Risk for Psychosis: Survey Study
JMIR Form Res 2022;6(7):e30230
URL: https://formative.jmir.org/2022/7/e30230
doi:10.2196/30230
PMID:35802420

Original Paper

Development of a Maternal and Child mHealth Intervention With Aboriginal and Torres Strait Islander Mothers: Co-design Approach

Sarah Jane Perkes[1*], BSc, MPH; Belinda Huntriss[2*], BEd; Noelene Skinner[2*], ATSIHP; Bernise Leece[2], ATSIHP; Rosie Dobson[3], PhD; Joerg Mattes[2], MBBS, PhD; Kerry Hall[4], PhD; Billie Bonevski[1], PhD

[1]Flinders Health and Medical Research Institute, College of Medicine and Public Health, Flinders University, Bedford Park, South Australia, Australia

[2]School of Medicine and Public Health, College of Health, Medicine and Wellbeing, University of Newcastle, Newcastle, Australia

[3]National Institute for Health Innovation, The University of Auckland, Auckland, New Zealand

[4]First Peoples Health Unit, Griffith University, Southport, Australia

[*]these authors contributed equally

Corresponding Author:
Sarah Jane Perkes, BSc, MPH
Flinders Health and Medical Research Institute
College of Medicine and Public Health
Flinders University
Sturt Rd
Bedford Park, South Australia, 5042
Australia
Phone: 61 451994827
Email: sarah.perkes@newcastle.edu.au

Abstract

Background: Despite their growing popularity, there are very few mobile health (mHealth) interventions for Aboriginal and Torres Strait Islander people that are culturally safe and evidence based. A co-design approach is considered a suitable methodology for developing health interventions with Aboriginal and Torres Strait Islander people.

Objective: The aim of this study was to co-design an mHealth intervention to improve health knowledge, health behaviors, and access to health services for women caring for young Aboriginal and Torres Strait Islander children.

Methods: Aboriginal researchers led engagement and recruitment with health services and participants in 3 Aboriginal and Torres Strait Islander communities in New South Wales, Australia. Focus groups and interviews were facilitated by researchers and an app developer to gather information on 3 predetermined themes: design characteristics, content modules, and features and functions. Findings from the co-design led to the development of an intervention prototype. Theories of health behavior change were used to underpin intervention components. Existing publicly available evidence-based information was used to develop content. Governance was provided by an Aboriginal advisory group.

Results: In total, 31 mothers and 11 health professionals participated in 8 co-design focus groups and 12 interviews from June 2019 to September 2019. The 6 design characteristics identified as important were credibility, Aboriginal and Torres Strait Islander designs and cultural safety, family centeredness, supportive, simple to use, and confidential. The content includes 6 modules for women's health: Smoke-free families, Safe drinking, Feeling good, Women's business, Eating, and Exercising. The content also includes 6 modules for children's health: Breathing well; Sleeping; Milestones; Feeding and eating; Vaccinations and medicines; and Ears, eyes, and teeth. In addition, 6 technology features and functions were identified: content feed, social connection, reminders, rewards, communication with health professionals, and use of videos.

Conclusions: An mHealth intervention that included app, Facebook page, and SMS text messaging modalities was developed based on the co-design findings. The intervention incorporates health behavior change theory, evidence-based information, and the preferences of Aboriginal and Torres Strait Islander women and health professionals. A pilot study is now needed to assess the acceptability and feasibility of the intervention.

(JMIR Form Res 2022;6(7):e33541) doi:10.2196/33541

KEYWORDS

mHealth; co-design; Aboriginal and Torres Strait Islander; mother; baby; young children; mobile phone

Introduction

Background

The health and well-being of Aboriginal and Torres Strait Islander people have been significantly affected by dispossession, interruption of culture, and intergenerational trauma since the colonization of Australia [1]. The ongoing impact has resulted in an unequal opportunity for good health. The life expectancy of Aboriginal and Torres Strait Islander women is 8 years less than that of non-Indigenous Australian women [2]. In the 2018-2019 National Aboriginal and Torres Strait Islander Health Survey, the majority of women aged ≥15 years were not meeting guidelines for physical activity, vegetable intake, or fruit intake; 36% reported that they smoked tobacco daily; and 35% reported that they experienced high or very high levels of psychological distress [3]. Infant mortality continues to be unacceptability high for Aboriginal and Torres Strait babies at 2.1 times the rate of non-Indigenous infants (6.3 and 3.1 per 1000 live births, respectively) [4]. Mothers and babies getting the best possible care and support for a good start to life is 1 of 12 health priorities of the National Aboriginal and Torres Strait Islander Health Plan 2013-2023 [1].

Aboriginal and Torres Strait Islander people make up 3.3% (798,400/24,193,939) of the Australian population [5] and include many distinct groups with their own language and culture. In total, 44% of Aboriginal and Torres Strait Islanders live in regional areas, 37% in cities, and 18% in remote or very remote areas [5]. Those living in regional and remote areas have less access to primary health care and overall poorer health [6]. Nationally, Aboriginal and Torres Strait Islander people have less access to the internet at home (75.3% compared with 85.8% of all Australians); there are significant differences based on location: 82.8% in cities, 73.2% in regional areas, 61.3% in remote areas, and 49.9% in very remote areas [7]. More than 1 in 3 (35%) Aboriginal and Torres Strait Islander people are mobile-only users compared with a national rate of 1 in 5 (19.9%); these figures are linked to socioeconomic factors [8]. Using only a mobile is likely to incur more costs for data, less capability, and less access to more sophisticated digital health information and tools [8]. It is of importance that mobile health (mHealth) interventions are developed with a goal to increase digital inclusion.

mHealth is the use of mobile technology to improve health. Functions include SMS text messaging, multimedia messaging service, voice, internet access, and software apps, which range in complexity. mHealth is used for a variety of purposes, including health education, health behavior change, sensors and point-of-care diagnostics, registries and vital-event tracking, and data collection [9]. mHealth is being used increasingly for health promotion because of its reach, with >7 billion mobile phone subscriptions globally [10]; the Be He@lthy, Be Mobile initiative by the World Health Organization has reached >3.5 million people [11]. There are limited recent national figures on smartphone ownership among Aboriginal and Torres Strait Islander people, although available data indicate that ownership is high; a survey with 400 Aboriginal and Torres Strait Islander people in 2014 reported that 70% of Aboriginal and Torres Strait Islander people owned a smartphone and 69% used Facebook compared with 66% and 40% respectively for non-Indigenous Australians [12]. The top reason for using a mobile phone in this group was to send SMS text messages [12].

Studies focused on Aboriginal and Torres Strait Islander people using SMS text messaging to improve health show high acceptability of the modality [13-15]. SMS text messaging has the advantage of being accessible on all mobile phones and not requiring access to a data service. There are few technical barriers to SMS text messaging and high acceptability of the modality among new mothers [16,17]. In a metareview (23 systematic reviews, 371 studies, and 79,665 participants) on the impact of mHealth on a range of outcomes, including clinical outcomes, adherence to treatment and care, health behavior change, disease management, and attendance rates, SMS text messaging was the most frequently examined function and reported to be the most successful overall [18]. SMS text messaging seems to be particularly effective at increasing smoking cessation rates (in adult smokers from mostly high-income countries) [19]. The evidence for SMS text messaging helping to improve nutrition and physical activity is not as strong; however, SMS text messaging used in conjunction with other mHealth functionality has shown significant positive effects for healthy eating [18].

Health apps continue to be popular, although the evidence suggests that apps have limited effectiveness on changing health behaviors [18,20-22]. Some studies have found that apps can be effective at changing behavior among some clinical groups [18], although overall there is limited evidence to date. Of the few trials focused on Indigenous populations, app use has been reported to be low [23,24]. A recent pilot randomized controlled trial of a smoking cessation app with 49 Aboriginal people in Australia reported low to moderate level of app use, and at 6-month follow-up, only 1 participant was abstinent [24]. The authors concluded that although there was broad acceptability for the app, mHealth interventions should be designed with functions that are commonly used, including social media platforms [24]. A co-designed mHealth app developed in New Zealand with Māori and Pacific Islander people was tested in a cluster randomized controlled trial in 2019 (n=1451) [23]. Adherence to health-related–behavior guidelines increased at 12 weeks in both groups, with no difference between the groups. Engagement with the app overall was low, although those who did engage with the app as it was designed saw greater benefit. The co-design approach was reported to have drawn a very positive response from the community, as was reflected in the high participation and follow-up rates [23].

Social media is a form of mHealth, with potential to support health. The Aboriginal and Torres Strait Islander health sector was an early adopter of social media networks to promote health [25,26]. Social media campaigns on COVID-19 by Aboriginal and Torres Strait Islander health organizations is a recent example [27]. A recent Cochrane review on behavioral interventions delivered through social media for health behavior change, health outcomes, and health equity (88 studies; n=871,378) reported varied effects; overall, social media was found to improve physical activity, weight loss, and general

well-being, and small to no effects were found for other outcomes [28]. No studies focusing on Aboriginal and Torres Strait Islander people were included in the review.

Objectives

In response to the limited mHealth interventions available for Aboriginal and Torres Strait Islander women and children, we aimed to co-design a prototype focused on the needs and ideas of Aboriginal and Torres Strait Islander mothers. Co-design is a partnership approach where end users are actively involved from conception to dissemination [29]. Using co-design methodologies is one of the guiding principles of the Aboriginal Health and Medical Research Council of New South Wales (NSW) Ethical Guidelines for conducting health research with Aboriginal people [30]. In this paper, we describe the co-design processes and findings, as well as provide a description of the mHealth prototype.

Methods

Study Design

In total, 8 focus groups and 12 interviews were conducted from June 2019 to September 2019. Surveys were used to collect demographics at the start of focus groups and interviews. An Aboriginal advisory group that included Aboriginal team members who were also members of the participating communities met quarterly to oversee design, implementation, analysis, and reporting. An expert mHealth research group was consulted for opinion on research and intervention design.

Ethics Approval

Human research ethics approval was received from the Aboriginal Health and Medical Research Council (1485/19) and the University of Newcastle (H-2019-00760).

Co-design Framework

A co-design framework for an mHealth intervention with Māori and Pacific communities in New Zealand [29] based on work by Bratteteig et al [31] was used to guide the methods used in this study. Co-design is a coherent methodology with a range of tools and techniques used to favor the preferences of end users [31]. The co-design methods used included focus group and interview discussions, card sorting, storyboarding, design activities, survey, guidance from expert groups, and an iterative design phase with the research team.

Setting

Focus groups and interviews were held at 3 regional NSW locations: Newcastle, Coffs Harbour, and Inverell. In total, 5 Aboriginal organizations (including 3 Aboriginal health services, an Aboriginal preschool, and an Aboriginal corporation) and 3 NSW Health sites participated. Venues for focus groups and interviews were decided in consultation with participants.

Participants

Women aged ≥16 years who were either mothers or primary carers of an Aboriginal or Torres Strait Islander child aged 0 to 5 years or were pregnant (≥30 weeks gestation), owned or regularly used a smartphone, and had accessed a participating service (Aboriginal health service or NSW Health service) were eligible to participate. Health professionals at participating services who worked with women or children were eligible.

Procedures

Convenience sampling was used to recruit participants. Aboriginal researchers (BH, NS, and BL) who worked within the participating communities used their personal networks. In addition, participants were asked if they would like to recommend a friend or family member to the study. Potential participants were screened for eligibility when they contacted the researcher on the telephone. The researcher explained the study and gained informed consent over the telephone initially and again in person before the start of the focus group or interview. Participants were reimbursed with a shopping voucher worth Aus $30 (US $21.6) for attending focus groups and interviews and provided with refreshments. Health professionals were recruited using a snowball methodology through the participating services. Health professionals were not reimbursed.

Mothers and health professionals participated in separate focus groups and interviews. Focus groups and interviews were cofacilitated by a combination of Aboriginal researchers (NS and BH), a PhD student (SJP), and an app developer. Interviews and focus groups were 20 to 90 minutes in length. The number of participants in focus groups ranged from 2 to 6. Focus groups and interviews were recorded and transcribed, and field notes were taken.

Measures

Different surveys and discussion guides were used with mothers and health professionals. Discussions and activities were used to identify (1) design characteristics, (2) content modules, and (3) features and functions.

Mothers

Survey

The survey comprised 16 items, including demographic, cultural, and socioeconomic items. The items were selected from a previous study [32], with all items having been tested with Aboriginal and Torres Strait Islander mothers previously.

Discussion Guide

In all focus groups and interviews with mothers, 3 main questions were asked. Follow-up questions were asked depending on responses. Additional questions about mobile phone use to inform features and functions were asked in focus groups cofacilitated by the app developer. The three main questions were as follows:

How would an mHealth intervention designed for healthy living for Aboriginal and Torres Strait Islander people differ from other mHealth interventions?

Are you more interested in mHealth for your own health or your child's health? What topics and features interest you?

What do you think stops or prevents some women from accessing health information and services for themselves and their children?

Activities

Card-sorting activities were used to identify current mobile phone use (functions used, frequency of use, and reasons for use). Storyboarding activity was used to elicit creative descriptions of the mHealth intervention using drawings and words on what the intervention should include. Design activity was used to gain feedback on potential designs.

Health Professionals

Survey

The survey comprised 5 items related to demographic and professional practice characteristics.

Discussion Guide

In all focus groups and interviews with health professionals, 3 main questions were asked. Additional follow-up questions were asked depending on the response. The three main discussion questions were as follows:

1. What do you think are the most important health and well-being topics to include for Aboriginal or Torres Strait Islander women, children, and family?
2. What are the barriers for Aboriginal or Torres Strait Islander families to having good health?
3. What types of mobile technology do you think could support Aboriginal or Torres Strait Islander women's and children's health?

Co-design Analysis

A generalized thematic analysis was completed. An Aboriginal researcher (BH) and a PhD student (SJP) independently coded themes. NVivo software (version 12.0; QSR International) was used to complete independent coding and comparison by the 2 coders. In total, 3 predetermined codes were used based on a similar co-design study [29]. These codes included (1) design characteristics, (2) content modules, and (3) features and functions. The coders met to agree on subcodes and definitions. Survey findings are presented using descriptive statistics.

Intervention Development

The findings from the co-design stage were subsequently used to develop a prototype intervention incorporating an app, SMS text messaging, social media, and videos. The intervention development was an iterative process, with meetings held among the team members to decide the final features and functionalities. Not all ideas could be adopted because of various reasons, such as time, funding, and technology constraints. We used a combination of building new functions (app) and using existing functions (Facebook page and SMS text messaging).

The intervention was grounded in behavior change theory. The Health Belief Model was used to underpin the app portion of the intervention. The Health Belief Model is considered to be well suited to mHealth interventions with use of the *cue to action* component [33]. The basic constructs are perceived threat of illness, perceived benefits of health behavior change, perceived barriers to change, cues to action, and self-efficacy [34]. Behavior change techniques were used to formulate SMS text messages. The SMS text messages were coded for behavior change techniques by 2 coders (Sam McCrabb and SJP) using behavior change technique taxonomy (version 1) [35] and the process outlined by Michie et al [36]. Of the 2 coders, 1 was experienced in coding behavior change techniques (Sam McCrabb) and the other was a PhD student (SJP). Disagreements were resolved through discussion and key messages adapted to include further effective behavior change techniques.

Key messages were developed on health topics identified from the focus groups and interviews. Content was formulated from publicly available evidence-based health resources. Key messages were adapted to SMS text messages, small pieces of written information for the app, and Facebook posts.

The prototype intervention included an app, videos, Facebook page, and SMS text messaging (Textbox 1).

Textbox 1. Components of the prototype intervention.

App

- A web-based prototype app was developed. Rapid iterative cycles between the app developer and research team were used to refine the design. An Aboriginal graphic designer developed graphics for each module and logo.

Videos

- A total of 12 short videos were captured on a Canon camera. All presenters were health professionals from participating sites or contacts of the research team. Short scripts were provided to health professionals based on key messages. Staff were encouraged to use their own knowledge and expertise on each topic. Videos were filmed by a videographer and professionally edited. Captions were completed by Rev, and voiceovers were completed by 2 Aboriginal researchers (BH and NS). The videos ranged from 112 to 300 seconds in length. Vimeo was used as the platform to host the videos.

Facebook page

- A Facebook group was developed and administrated by 2 Aboriginal researchers (BH and NS). Both researchers were regular Facebook users and had significant networks and knowledge of Aboriginal and Torres Strait Islander organizations, events, and health services. Key messages were predeveloped in text and video format. Other content shared was decided by the administrators, including sharing posts from their personal accounts if they were suited to the broad aim of the intervention.

SMS text messaging

- SMS text messages were developed based on the processes described by Abroms et al [37]. Steps include choosing a behavior change goal, choosing communication objectives and behavioral techniques, designing a framework, and writing an SMS text message library [37]. SMS text messages were written to allow tailoring using the mother's and child's names, child's age, and topic interest of the mother. Tailoring SMS text messages around the timing of key behaviors, such as after a baby is born, can improve saliency and likelihood of behavior change [38]. SMS text messages were written by an Aboriginal researcher (BH) and a PhD student (SJP). A web-hosted SMS text messaging server (SMS Express) will be used to send all SMS text messages.

Results

Overview

A total of 42 participants were recruited to the study: 31 mothers and 11 health professionals. Demographics and cultural characteristics of mothers are presented in Table 1, and demographics of health professionals in Table 2.

Table 1. Demographic and cultural characteristics of mothers (N=31).

Characteristics	Values
Age (years), mean (SD; range)	31.17 (7.69; 19-50)
Indigenous status, n (%)	
Aboriginal	21 (68)
Torres Strait Islander	2 (7)
Nonidentified	7 (23)
Did not answer	1 (3)
Identified with an Indigenous community, n (%)	
Yes	25 (81)
No	1 (3)
Unknown	4 (13)
Did not answer	1 (3)
Maintain cultural connections at home, yes, n (%)	25 (81)
Ways of connecting to culture, n (%)	
Music or dance	19 (61)
Storytelling	19 (61)
Indigenous television	18 (58)
Art	15 (48)
Food	14 (45)
Indigenous internet sites	10 (32)
Indigenous newspapers	7 (23)
Traditional medicine	6 (19)
Indigenous radio	5 (16)
Other	1 (3)
Family members from Stolen Generations[a], n (%)	
Yes	6 (19)
No	12 (39)
Unknown	13 (42)
Education of mother, n (%)	
Did not finish high school	6 (19)
High school	6 (19)
Certificate	10 (32)
Diploma	2 (7)
Bachelor’s degree	4 (13)
Postgraduate degree	1 (3)
Did not answer	2 (7)
Currently pregnant, yes, n (%)	1 (3)
Partner, yes, n (%)	16 (52)
Number of people living in household, mean (SD; range)	4 (1.31; 2-7)
Number of children (aged <18 years) living in household, mean (SD; range)	2.39 (1.41; 1-5)
Smoking status of mother, n (%)	
Nonsmoker	21 (68)
Yes, daily	5 (16)

Characteristics	Values
Yes, at least once a week	2 (7)
Yes, less often than once a week	1 (3)
Did not answer	2 (7)
Number of cigarettes smoked per day (on the days smoking), mean (SD; range)	8.5 (3.21; 4-12)
Number of smokers in household, n (%)	
0	14 (45)
1	10 (32)
2 to 3	4 (13)
>3	1 (3)
Child exposure to indoor tobacco smoke, yes, n (%)	1 (3)
Child exposure to outdoor tobacco smoke, yes, n (%)	15 (48)
Child exposure to tobacco smoke in the car, yes, n (%)	0 (0)

[a]The Stolen Generations refers to a period in Australia's history when Aboriginal children were removed from their families through government policies. This happened during the period from the mid-1800s to the 1970s [39].

Table 2. Demographics of health professionals (N=11).

Characteristics	Values
Health service type, n (%)	
Aboriginal medical service	6 (55)
NSW[a] Health service	5 (45)
Sex: female	11 (100)
Indigenous status, n (%)	
Aboriginal	4 (36)
Torres Strait Islander	0 (0)
Nonidentified	7 (64)
Role at health service, n (%)	
Registered nurse	7 (64)
Aboriginal health worker	3 (27)
Senior family health practitioner	1 (9)
Number of years at service, mean (SD; range)	12 (8.7; 3-32)

[a]NSW: New South Wales.

Design Characteristics

We identified six main design characteristics: (1) credibility, (2) Aboriginal and Torres Strait Islander designs and cultural safety, (3) family centeredness, (4) supportive, (5) simple to use, and (6) confidential.

Credibility

Mothers talked about the difficulty of finding information on the web that was evidence based. Most of the mothers said that they used Google to find real-time health information for themselves and for their children: "Literally, I Google everything." Many of the mothers said that it can be difficult to know which websites are most up to date and accurate and that it is difficult to find information: "The biggest thing I find on Google, you get everything. You don't get the ones that are reputable." Another mother said, "I'm finding you're having to like scroll, scroll, and scroll to try and find that information." Mothers said that they want current health information from reputable health professionals and organizations, including "useful websites links." Health professionals talked about the importance of credible health information to improve health literacy: "I think lack of knowledge that they are so sick. Recognizing the signs of illness that can lead to them being really, really [sick]." This highlighted why it is important that all content included in the prototype intervention be sourced from credible evidence-based health resources and broken down into palatable small chunks with links to further information.

Aboriginal and Torres Strait Islander Designs and Cultural Safety

Most of the mothers said that Aboriginal designs, language, and representation were important for engagement. A mother said, "I think if it had Aboriginal designs that would be really good because if I download an app and it doesn't have the look, like being culturally aware [I don't use it]." Another mother said, "Don't make it black and white, it's got to be like colorful." A mother spoke about the intervention needing Aboriginal representation in images and videos: "If it's going to be an Aboriginal app, I think you have to have Aboriginal people." Another mother discussed using an app for quitting smoking that was not representative of Aboriginal people: "It was easy to use, but I couldn't relate to it...didn't seem like it was aimed at Blackfellas even though we thought it was."

It was evident from the mothers' experiences of racism that the intervention needed to be centered in culturally safety. Some mothers talked about feeling fearful and judged when seeking health care. A mother said, "Being an Aboriginal mum especially, I was just worried about DoCS [Department of Child Services]. Like whether they could see if I was handling having two children on top of my own family breakdown. Like my mum's kids are in DoCS. So that's what my biggest fear was." Other mothers expressed feeling judged about certain health behaviors and topics, and a mother said, "The biggest thing is why people do hide it [smoking], because they don't want to be judged. They don't want to hear all that stuff."

To center cultural safety in the intervention, all aspects of the intervention were codeveloped by Aboriginal people: the research was governed by an Aboriginal advisory board and coled by an Aboriginal academic (KH); 4 of the 8 members of the research team are Aboriginal; an Aboriginal graphic designer designed the module icons and logo; Aboriginal researchers were administrators of the Facebook page and shared cultural links, events, activities, affirmations, and images; an Aboriginal videographer filmed all the videos; Aboriginal health professionals presented in the videos; an acknowledgment of Country and a *welcome* message by an Aboriginal researcher was placed on the main page; and all content was cowritten by Aboriginal researchers.

Family Centeredness

It was decided unanimously that the intervention should include content for both mother and child. A mother said, "Is this just for children's health? Because I feel like it should incorporate the mother's health too." The mothers asked for information on "things to do with our kids," and "stuff for us women too. Pap smears and stuff like that." Many of the mothers and health professionals suggested that the intervention needed to encompass the entire family, including the extended family. A health professional said, "Put the main focus on the child and then how their [family] health affects the baby's health," and a mother said, "I think a family app would be really good. Like, I know my husband, he's never been around babies." Some participants talked about how other family members help bring up children: "It's nothing to see an aunt bringing up a child, or a grandparent or a sister" [health professional]. Family centeredness in the intervention was therefore conveyed through messaging that families are the most important role models for jarjums (an Aboriginal word meaning children) across modules and functions. Links to websites, events, and health information for partners and other family members were included.

Supportive

Most of the mothers and health professionals indicated that it was important that the intervention promoted positive self-esteem and well-being of mothers. A health professional said that the intervention should give new mothers "understanding [of] how tired you are going to be, and it's okay, ask for help, everyone feels like that but you're not failing or not doing something wrong." A mother suggested that we include "some sanity sayings or something like that, or some little sage advice from mums that have been there, done that before, that'd be really helpful," and another mother said that the intervention could be "like a reassurance type thing." Mothers and health professionals recognized that motherhood can be "totally exhausting" [health professional] and challenging at times. A mother described the initial period after coming home from hospital: "I didn't know what to do with him. What do I do with this kid? I was lost." To create an intervention that was supportive of motherhood and of Aboriginal and Torres Strait Islander women, positive and affirming messages were posted on Facebook, sent through SMS text messages, and included in the app. Links on where to seek help for mental health concerns were included.

Simple to Use

Mothers and health professionals recommended that the intervention be intuitive, use simple language, and have few technical barriers. Some of the mothers talked about trying to use other health apps; however, they were unable to do so because of technological challenges. For example, a mother said, "It was just too hard to log in and get started so I gave up or just called someone." Many of the mothers and health professionals emphasized that the language used in the intervention needed to be nonjargon. A mother said, "Don't put it in a textbook. Because I'm telling you, if my family member downloaded that and it was a textbook way, they would be like—No." Another mother said that the content should be "just little pieces of information...then links to the bigger pieces." We aimed for simple, intuitive app design and used other mobile functions commonly used by mothers (Facebook and SMS text messaging). To ensure that the intervention was simple and easy to use, health information was presented in short key messages with links to websites for further information. All key messages were written to be at an 8th grade reading level using the Flesch-Kincaid Grade Level Test as recommended by Abroms et al [37].

Confidential

Mothers and health professionals talked about the importance of confidentiality. Health professionals focused on confidentiality in the health care setting and the complexities for some staff regarding knowing patient health details. A health professional said, "There are big things surrounding our health services confidentiality. People don't know or want to know what other people's business is." Some of the mothers spoke

about confidentiality; regarding being anonymous when communicating with other mothers or health professionals in a hypothetical mHealth intervention, a mother said, "Oh God, yeah. I'd ask an anonymous person on a phone. Rather than ask the doctor face to face." Other mothers were happy to not be anonymous: "It wouldn't bother me having my name because it would just be, this is my experience, and it is what it is. But I would understand if some women didn't." To ensure that women can choose to remain anonymous and keep their information confidential, the intervention design meant that no personal data were collected in any part of the intervention, other than a mobile number for the SMS text messaging component. Joining the Facebook group is an optional part of the intervention.

Content Modules

Most of the mothers and health professionals suggested that the intervention needed to cover a wide range of health topics for both the mother and child. Health topics identified in the data included *pains after birth*, *breastfeeding*, *normal speech for toddlers*, *signs of autism*, *earaches*, *behavior*, *rashes*, *high temperatures*, and *coughs*. Similar topics were grouped by the research team and combined into 6 key content modules for women's health and 6 key content modules for children's health. For example, *birth*, *reproductive health*, *urinary leaking*, and *pap smears* became *Women's business*. All health topics captured in the interviews and focus groups were included in the intervention within a module on the app, SMS text messages, or through Facebook posts. Health modules for women included Smoke-free families, Safe drinking, Feeling good, Women's business, Eating, and Exercising. Health modules for children were Breathing well; Sleeping; Milestones; Feeding and eating; Vaccinations and medicines; and Ears, eyes, and teeth.

Features and Functions

We identified eight features and functions: (1) content feed, (2) social connection, (3) diary and storage of health information, (4) local context, (5) reminders, (6) rewards, (7) talk with health professionals, and (8) use of videos.

Content Feed

A content feed was chosen to be a feature of the intervention based on the mothers' current mobile phone use. During the card-sorting activity, most of the mothers reported scrolling the content feed on Facebook numerous times per day. Of the 13 women who were asked how many hours per day they used Facebook, 12 (92%) reported using it >4 hours per day. When asked what kept them going back to Facebook, a mother responded, "The content keeps changing." Mothers frequently talked about watching photo and video stories that were uplifting, funny, or motivating on Facebook. They talked about using Instagram and Snapchat, too, although less frequently. The intervention was therefore designed to include a Facebook page with daily posts covering a variety of health content.

Social Connection

Mothers talked about the social connection and learning from other women when becoming a mother, including from their "mum," "mother-in-law," and "girlfriends." The importance of positive relationships when first becoming a mother was well recognized by health professionals as well as mothers. It was acknowledged by many of the mothers that some new mothers "don't have a big support network." A mother described mothers at playgroup being "more like a family to each other." Some of the mothers said that connecting to other mothers would be helpful because they may be going through the same situation or challenge: "Yeah [I would like to chat with mums in the intervention] because they might have experienced something that I'm starting to experience." Some of the mothers talked about the possibility of meeting up with mothers outside of the intervention: "It's hard to meet people...[could there be] like a mums and bubs [babies] thing [as part of the intervention]," and another mother said, "Say, if I needed to ask them a question or something that I wouldn't want to write on Facebook [I would like to meet up with them in person]." Another mother identified that connection is important for mental health: "When they [new mothers] don't have anybody, depression kicks in." The Facebook page was designed to make it easy for mothers to connect and share stories and ideas. Discussion points were created to be posted on the Facebook page to facilitate discussion; for example, "Tell us how you engage your jarjums in cooking or take a pic or video of your deadly (great or excellent) li'l chef in the kitchen."

Diary and Storage of Health Information:

A feature that enabled users to store specific information about a child's health received mixed responses. Some of the mothers thought that having their child's health information on hand would be of practical benefit when attending medical appointments: "Like a diary section...I found, when [my child] was sick I started recording when I gave the medication, those sorts of things. That'd be good to have an app when you go into the hospital, you go, this is his recordings." Another mother said, "So they [health professionals] could just add in medication, add in reports...it'd be good because like [the health service] is only open during the week. Usually, like on the weekend, I'd have to go up to the hospital...So it would be good if there was information like after the visit. Because you don't always take everything in. It goes right over your head." Other mothers and health professionals thought there would be confidentiality concerns. Because of the confidentiality concerns raised in the co-design process, a diary feature was not included, although it may be considered as an optional feature in future iterations.

Local Context

Many of the mothers and health professionals spoke about the uniqueness of their community and said that the intervention needed to be relevant to each community, including language and environment (eg, coastal and desert), as well as health services and other resources. A mother suggested, "You could put in your postcode, location, or area or something and then it could be localized," and a health professional said, "The contact numbers, if they can't get into emergency, the [local] health line numbers where they can get a bit of advice would be handy on there as well." The intervention included phone numbers of local health services for each community in the app, and Facebook posts were designed promoting local health services, events, organizations, and languages.

Reminders

Many of the mothers talked about the usefulness of SMS text messaging reminders from their health services for appointments, and they said that reminders for other areas of health care would be useful too. A mother said, "I would probably like all of them [milestone reminders]. I'd like the whole lot, make sure I'm not missing anything." Another mother said, "If someone notified me on this app that I'm due for a [pap smear] or something like that, I would like being reminded of things like that." Most of the mothers said that they would prefer reminders through SMS text messages rather than a push notification from an app because they could go back to the message and reread it. For the intervention, SMS text messages were developed covering a range of reminders, including vaccinations, developmental milestones, check-ups, smoking quit date, exercise, and eating well. Reminders about local health initiatives and events were also created for posting on the Facebook page.

Rewards

The mothers talked about rewards and incentives from health programs and services increasing their motivation. They talked about material rewards such as "shirts," "caps," and "supermarket vouchers," as well as social rewards, including "comments" and "likes" on social media and "clapping" and "cheers" on health apps. The mothers who were asked about receiving rewards for a variety of health behaviors were unanimous in their opinion that rewards were enjoyable and motivating. In the intervention, weekly competitions were created for posting on the Facebook page involving mothers sharing a picture of a health activity; for example, active play or exercising with their children. Prize draws were also incorporated into the intervention for those who participated in the competitions.

Talk With Health Professionals

Some of the mothers suggested that being able to communicate with health professionals using SMS text messages or a live chat function would be beneficial. Some of the mothers said that this function would be useful to confirm whether they required face-to-face health care and for reassurance. A mother said, "Sometimes you don't know if you should go up there [health service] or not, so you could kind of message and say, 'Hey, this is what's happening...is it worth coming up or is it just a viral thing going around?'" Another mother said, "I know a lot of women are just like, 'What do I do?' So just having that reassurance I suppose online." Another mother suggested that it would be helpful to be able to ask health questions anonymously: "The option to be anonymous or not known by people [health professionals] would be handy I guess for more embarrassing health concerns." Mothers living in rural areas mentioned being anonymous more often in the discussions. Although it was suggested, facilitating a chat with health professionals directly was out of the scope of the current prototype because of cost and resources. Telephone numbers for national, state, and local health services were listed in the app to enable users to connect with health professionals, if needed, regarding the questions they might have.

Use of Videos

Most of the mothers reported during the card-sorting activity that they frequently watched short videos on social media and YouTube. A number of mothers and health professionals advised us that videos and images may be more accessible and preferable for some mothers. A health professional said, "Videos, everyone can watch a video and understand." Therefore, a video for each health module was developed for the intervention. Each video was stored in the app and added to the Facebook page. Additional health videos from external sources were also able to be shared on the Facebook page.

Final Prototype

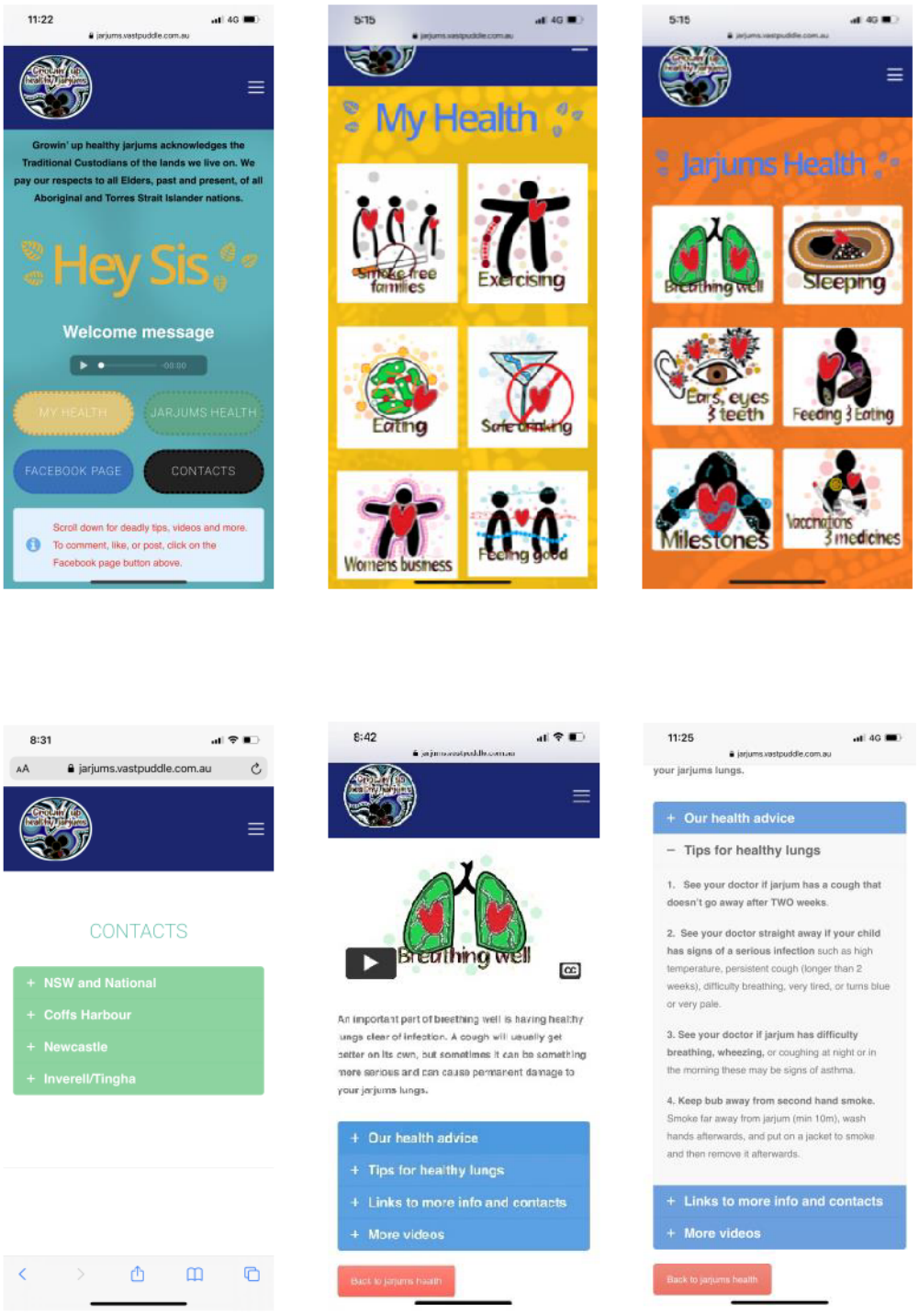
The final mHealth intervention, named Growin' Up Healthy Jarjums, aimed to improve health knowledge and health behaviors, along with providing access to health services. The intervention comprises 3 delivery modalities: app, SMS text messaging, and Facebook page.

App

The app is a central place for users to access all content. The app is primarily for the user who wants in-depth information and has the necessary digital device, internet connection, and literacy skills to access it. It is designed to allow the user to navigate to the topic of interest; for example, *exercise*, where they will find small amounts of written information, videos, links to websites, and useful contacts. The user may choose to access any topic, in any order, and consume as much information as they like.

The app has four menu screens: (1) home screen, (2) women's health, (3) children's health, and (4) contacts (Figure 1). The home screen includes four buttons: (1) *My Health*, (2) *Jarjums Health*, (3) *Facebook Page*, and (4) *Contacts*. The user may click on a button to move to the next screen or scroll down to access the embedded Facebook content feed. The embedded Facebook content feed allows the user to remain in the app and read the posts, but to comment or *like* a post, the user needs to access the Growin' Up Healthy Jarjums Facebook page. An acknowledgment of Country and a spoken welcome message are also included on the home screen. The women's health (*My Health*) menu page includes six buttons, one for each of the women's health modules: (1) Smoke-free families, (2) Safe drinking, (3) Feeling good, (4) Women's business, (5) Eating, and (6) Exercising. The *Jarjum's Health* menu page has the same layout, including six buttons for the children's health modules: (1) Breathing well; (2) Sleeping; (3) Milestones; (4) Feeding and eating; (5) Vaccinations and medicines; and (6) Ears, eyes, and teeth. Each module, for example, *Breathing well*, includes (1) *Key messages* incorporating perceived threat of illness and benefits of changing health behavior; (2) *Tips* to address barriers to change through reassurance and credible advice; (3) cues to action; for example, "Each time jarjum sees a nurse or GP ask them to have a quick look in bub's ears to check if there is any infection"; and (4) links to further information, including skills and activities; for example, exercises and healthy recipes to support self-efficacy. The information is presented using small chunks of written information and videos using the same layout in each module.

Figure 1. Examples of Growin' Up Healthy Jarjums app screens: (top, from left) home, women's menu, and children's menu; (above, from left) contacts, Breathing well, and Our health advice (accessed from Breathing well).



SMS Text Messaging

Alongside the app, the prototype included an SMS text messaging library comprising 112 SMS text messages (Table 3). The SMS text messaging component allows users access to health information regardless of mobile phone type, Wi-Fi access, or digital literacy. The SMS text messages covered the content topics identified by the participants. The SMS text messaging portion of the program is 1-way (unidirectional), other than 3 SMS text messages developed for users who indicate that they want to quit smoking when registering for the program. In total, 23 behavior change techniques from 15 behavior change clusters were incorporated in the SMS text messages (Multimedia Appendix 1).

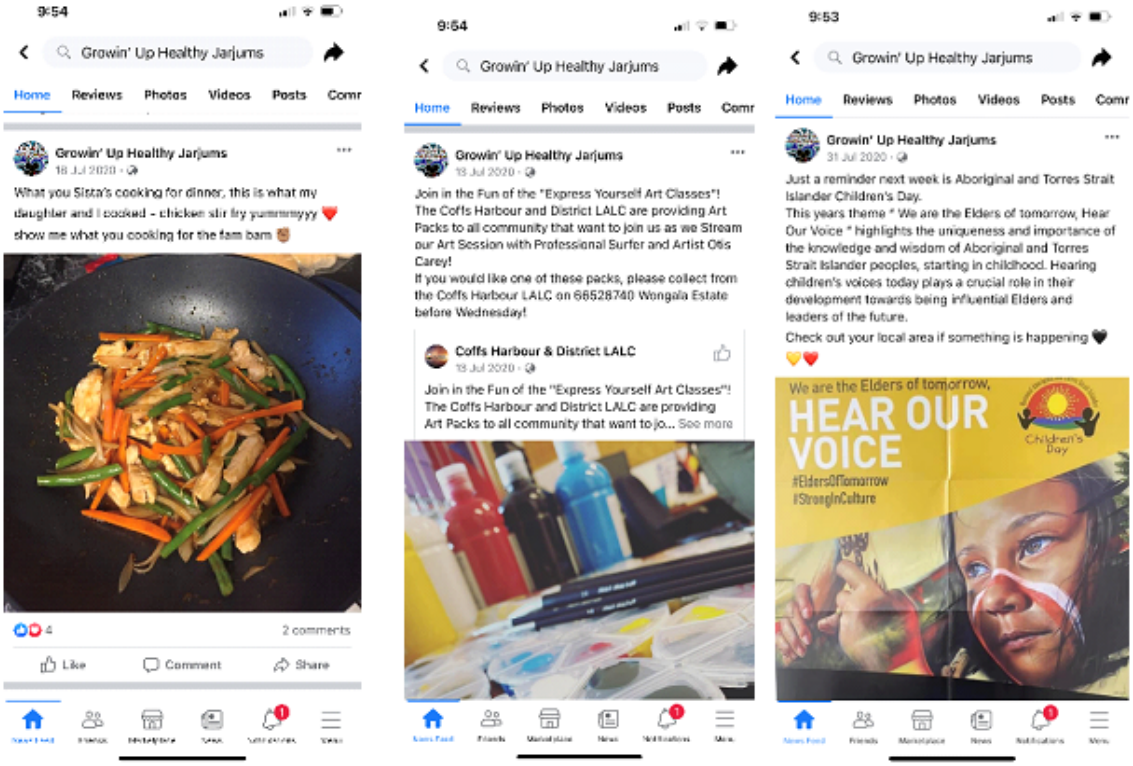
Table 3. Example SMS text messages developed for the Growin' Up Healthy Jarjums modules.

Module	Example SMS text message
Women's health	
Smoke-free families	Text4jarjum: Giving up the smokes is the best thing you can do for your health. Be a role model and be smoke free. Get support from Quitline 13 78 48 or a doctor and quit for good!
Safe drinking	Text4jarjum: While under the influence of alcohol, people can make less safe decisions about their jarjums. Check out 'Safe drinking' for tips to set limits.
Feeling good	Text4jarjum: You're probably not getting much sleep right now. Try to make time for yourself, ask for support from family & friends, and nap when bub does. If you feel that you are not coping, talk to your doctor or midwife. There is help.
Women's business	Text4jarjum: Be kind to yourself. Your body has gone through some big changes during and after birth. It will take time to bounce back. Whether you had a caesarean or vaginal birth, both may require rest & time for recovery. Here's what to expect after birth.
Eating	Text4jarjum: The Australian Breastfeeding Association has some useful tips on nutritional needs for breastfeeding mums.
Exercising	Text4jarjum: Any amount of movement is good for you. Start by doing a little, and gradually build up. You could start with a walk around the block a few times a week and then gradually increase.
Children's health	
Breathing well	Text4jarjum: A cough is often caused by a cold. Usually, a cough gets better on its own and is not serious, but if your child has a cough that doesn't go away after TWO weeks, or if you are concerned sooner – see your doctor or child health nurse.
Sleeping	Text4jarjum: A routine that includes relaxing time like bath, book, a gentle song before bed and a regular bedtime each night can help your child settle better.
Milestones	Text4jarjum: Playgroups, day care and pre-school are great places for jarjums to play and develop. Contact your AMS (Aboriginal Medical Service) or health nurse and find out what's on.
Ears, eyes, and teeth	Text4jarjum: Ear infections are really common and can cause long term hearing loss if not treated. Often there are no signs. Ask your doctor to have quick look in [insert child name] ears each visit to make sure there is no infection.
Vaccinations and medicines	Text4jarjum: Immunising [insert child name] is a safe and easy way to keep jarjums healthy and prevent disease. To check that [insert child name] is up to date with immunisations click here.
Feeding and eating	Text4jarjum: It's recommended you breastfeed exclusively until [insert child name] starts solid foods at around 6 months of age. Keep breastfeeding until at least 12 months and beyond.

Facebook Page

The final modality included in the prototype was the Facebook page. The purpose of the Facebook page was to create community and connection, allow 2-way communication, and use a platform that is highly popular among users. Daily content was designed to be added to the Facebook page, including (1) links to reliable health websites, (2) activities for families, (3) weekly competitions, (4) key messages (written and video), (5) events in the community, and (6) supportive affirmative posts. The page was administrated by 2 Aboriginal team members (NS and BH), who shared posts relevant to their community and region. The Facebook page was embedded into the main screen of the app; it could also be accessed through Facebook. Examples of posts are presented in Figure 2.

Figure 2. Examples of the content feed shared on the Growin' Up Healthy Jarjums Facebook page.



Discussion

Principal Findings

We codeveloped a prototype mHealth intervention focused on the knowledge of mothers of young Aboriginal and Torres Strait Islander children. The aim of the intervention was to improve health knowledge, health behaviors, and access to health services. The final prototype incorporates 3 modalities—app, SMS text messaging, and Facebook page—and includes a range of health topics. In addition, it is centered on being supportive of mothers and culturally safe.

The modality choices were based on a few factors: (1) early discussions with mothers and health services about the need for an app that is culturally relevant and safe, (2) evidence suggesting that SMS text messaging is the most effective mHealth function for health behavior change, and (3) findings from focus groups and interviews indicating that Aboriginal and Torres Strait Islander women were high users of Facebook and SMS text messaging. As suggested in a recent pilot study of a smartphone app with Aboriginal Australians, a *one app fits all* approach is unlikely to be successful [24]. Using mHealth modalities commonly used by the target group to deliver a health intervention may appeal to more families.

Strengths and Limitations

The first limitation of this research is that it was initiated by a research institution rather than by the community itself. True co-design should begin with completing a needs assessment with communities to see what the health priorities and potential solutions are for that community [38]. This is well described in a New Zealand co-design study [29,40]. To ensure that adequate time and resources are available for relationship building and needs assessment, both should be specified in protocols and funding applications so that sufficient budgets and time frames are allocated. Second, although the intervention covers a range of topics in brief, it does not cover any topic in depth. Although an mHealth intervention with wide-ranging topics seems to be preferred by participants, this may dilute the impact of the intervention on any one risk behavior. Providing links within the *Growin' Up Healthy Jarjums* intervention to specific mHealth interventions for target behaviors may overcome this limitation by providing tips for more intense behavior change for those people who are *ready* to change. Third, because the participants were from only 3 NSW communities, the intervention may have limited generalizability in other Aboriginal and Torres Strait Islander communities. Aboriginal and Torres Strait Islander communities are made up of >250 language groups in which there is great diversity. If this intervention is to expand to other communities, systematic adaptation of the intervention would need to be carried out to ensure that the intervention is suitable to the context of each community [41].

A key strength of this study is that Aboriginal researchers (BH, NS, and BL) led engagement with participants and community organizations. Understanding the importance of trusted and strong cultural relationships, we only engaged with communities that the Aboriginal researchers had a relationship with, which likely resulted in trust as well as interest in participating in this study. Another strength of this study is the thorough reporting of the co-design processes. Inadequate reporting of intervention development was identified as a weakness in a recent systematic review on mHealth development (33). An additional strength is the involvement of primary health services and professionals. A recent review on health promotion programs in Aboriginal

communities highlighted that an important consideration is to partner with primary health care services because they are well placed with frequent patient contact, health expertise, and often intricate knowledge of the community [42]. A final and important strength is that we developed a flexible portal for ongoing development and enhancement. The COVID-19 experience has reinforced how important it is to have alternatives to face-to-face health care. Useful additions in future iterations of this mHealth intervention might include development of a flexible platform suitable for inclusion of initiatives inspired by the COVID-19 pandemic, such as subsidized telehealth and videoconferencing. There are also opportunities to develop content on this platform in Aboriginal and Torres Strait Islander languages to better suit users.

Comparison With Prior Work

Design characteristics identified in this study, including *social connection* and *family centeredness*, reflect Aboriginal and Torres Strait Islander perspectives of health. Connection to family, community, and culture, among other factors, are understood to be equal contributors to health [43]. Arabena et al [44] suggest that community and social connection can ultimately be the health promotion intervention for Aboriginal and Torres Strait Islander communities.

The finding that Aboriginal and Torres Strait Islander women were high users of social media, in particular Facebook, was unsurprising. Aboriginal and Torres Strait Islander health organizations have capitalized on the popularity of Facebook among Aboriginal and Torres Strait Islander people and have been early and adept users of social media for health promotion [25]. An Aboriginal-led social marketing campaign for health promotion, *Deadly Choices*, has 94,035 Facebook followers, 19,300 Instagram followers, and 9000 TikTok Followers [26,45].

As stated earlier, the methodologies used in this study were based on a co-design study for a health app with Māori and Pacific Islander people [29,40]. There were a number of similar co-design findings. In both studies, participants expressed a holistic view of health and connections to people and place as being central components of health. Participants in both studies talked about a family approach to health, rather than an individual approach, as well as accessible healthy activities in the community. Social support was found to be an important strategy in both studies.

Culture was also identified as important in both studies, although cultural representation may have been a more nuanced finding in the New Zealand study. In our Australian-based study, participants expressed the importance of Aboriginal and Torres Strait Islander representation in terms of designs, colors, images, people, organizations, and safety. Participants in the New Zealand study [29,40] expressed the need to include Māori knowledge, Whakapono (faith and spirituality), and Whakataukī (traditional proverbs), which were to be woven throughout the intervention; for example, the app depicts the completion of challenges as colored footsteps, which is analogous to the journey that the participants' tūpuna (ancestors) embarked on. There may be differences in participants' connection to culture. In Australia, up to 1 in 3 Aboriginal and Torres Strait Islander children were removed from their families during the period from the mid-1800s to the 1970s. These children are known as the Stolen Generations [39]. Of the 31 mothers in this study, 6 (19%) reported that they had family members from the Stolen Generations, whereas 13 (42%) were unsure. The effect of the Stolen Generations on the loss of culture is profound [39] and is likely reflected in the findings of this study. This intervention may, in a small way, help to promote culture through links to Aboriginal and Torres Strait Islander organizations, connection to mothers of Aboriginal and Torres Strait Islander children, and culturally safe health information.

Conclusions

An mHealth intervention that included app, SMS text messaging, and Facebook page modalities was developed based on co-design findings. The intervention incorporates health behavior change theory, evidence-based information, and the preferences of Aboriginal and Torres Strait Islander women and health professionals. The next step of this research is to assess the acceptability and feasibility of the intervention in a pilot study. The pilot study will be conducted with the Aboriginal Health Services and NSW Health sites that participated in this co-design study. Participating mothers will also be invited to participate in the pilot study. If the *Growin' Up Healthy Jarjums* intervention is shown to have adequate acceptability and feasibility, the next phase will be to measure its effectiveness in improving health knowledge and changing health behaviors. Assessing the effectiveness of this intervention will provide valuable evidence for the use of mHealth in improving the health and well-being of Aboriginal and Torres Strait Islander populations and contribute to the evidence for using co-design methodologies, both of which have been highlighted as gaps in the literature [46].

Acknowledgments

The authors would like to acknowledge and pay their respects to the Aboriginal and Torres Strait Islander women who shared their views and knowledge in the study. The authors would also like to acknowledge the Aboriginal Community Controlled Health Services that participated in the study (Armajun Aboriginal Medical Service [AMS], Galambila AMS, Awabakal AMS, Kulai Aboriginal preschool, and Muloobinba Aboriginal Corporation) as well as the New South Wales Health sites that participated (Aboriginal Maternal and Infant Health Services located in Inverell, Coffs Harbour, and Newcastle). The authors wish to thank researchers from the National Institute for Health and Innovation, Auckland University, Auckland, New Zealand, for sharing expertise on mobile health and Sam McCrabb for coding behavior change techniques.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Behavior change techniques in SMS text messages.
[DOCX File , 15 KB - formative_v6i7e33541_app1.docx]

References

1. National Aboriginal and Torres Strait Islander Health Plan 2013–2023. Australian Government Department of Health. URL: https://www.health.gov.au/resources/publications/national-aboriginal-and-torres-strait-islander-health-plan-2013-2023 [accessed 2022-04-29]
2. Indigenous life expectancy and deaths 2019. Australian Institute for Health and Welfare. URL: https://www.aihw.gov.au/reports/australias-health/indigenous-life-expectancy-and-deaths [accessed 2022-04-29]
3. National Aboriginal and Torres Strait Islander health survey 2018-19. Australian Bureau of Statistics. 2018. URL: https://www.abs.gov.au/statistics/people/aboriginal-and-torres-strait-islander-peoples/national-aboriginal-and-torres-strait-islander-health-survey/2018-19 [accessed 2022-04-29]
4. 1.20 Infant and child mortality. Australian Institute of Health and Welfare: National Indigenous Australians Agency. 2020. URL: https://www.indigenoushpf.gov.au/measures/1-20-infant-child-mortality [accessed 2022-04-29]
5. Profile of Indigenous Australians 2020. Australian Government: Australian Institute for Health and Welfare. URL: https://www.aihw.gov.au/reports/australias-health/profile-of-indigenous-australians [accessed 2022-04-29]
6. Australian Government: Australian Institute for Health and Welfare. URL: https://www.aihw.gov.au/reports/rural-remote-australians/rural-remote-health/contents/access-to-health-care [accessed 2022-04-29]
7. Rennie E, Thomas J, Wilson C. Aboriginal and Torres Strait Islander people and digital inclusion: what is the evidence and where is it? Commun Res Pract 2019 Jun 10;5(2):105-120. [doi: 10.1080/22041451.2019.1601148]
8. Thomas J, Barraket J, Wilson CK, Holcombe-James I, Kennedy J, Rennie E, Royal Melbourne Institute of Technology, Telstra Corporation, Roy Morgan Research Centre. Measuring Australia's digital divide: The Australian Digital Inclusion Index 2020. RMIT University, Centre for Social Impact, Telstra. URL: https://apo.org.au/node/308474 [accessed 2022-04-29]
9. Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low- and middle-income countries--what has been shown to work? Glob Health Action 2014 Oct 27;7(1):25606 [FREE Full text] [doi: 10.3402/gha.v7.25606] [Medline: 25361730]
10. Be He@lthy, Be Mobile (BHBM). World Health Organisation. URL: https://www.who.int/activities/Addressing-mobile-health [accessed 2022-04-29]
11. Be Healthy, Be Mobile Annual Report 2018. Geneva, Switzerland: World Health Organization; 2019.
12. Media usage amongst Aboriginal and Torres Strait Islander People (infographic). McNair Yellow Squares. URL: https://mcnair.com.au/media-release-media-usage-amongst-aboriginal-and-torres-strait-islander-people-infographic/ [accessed 2022-04-29]
13. Kirkham R, MacKay D, Barzi F, Whitbread C, Kirkwood M, Graham S, et al. Improving postpartum screening after diabetes in pregnancy: results of a pilot study in remote Australia. Aust N Z J Obstet Gynaecol 2019 Jun;59(3):430-435. [doi: 10.1111/ajo.12894] [Medline: 30276799]
14. Phillips JH, Wigger C, Beissbarth J, McCallum GB, Leach A, Morris PS. Can mobile phone multimedia messages and text messages improve clinic attendance for Aboriginal children with chronic otitis media? A randomised controlled trial. J Paediatr Child Health 2014 May;50(5):362-367. [doi: 10.1111/jpc.12496] [Medline: 24612007]
15. Fletcher R, Hammond C, Faulkner D, Turner N, Shipley L, Read D, et al. Stayin' on Track: the feasibility of developing Internet and mobile phone-based resources to support young Aboriginal fathers. Aust J Prim Health 2017 Sep;23(4):329-334. [doi: 10.1071/PY16151] [Medline: 28449728]
16. Dobson R, Whittaker R, Bartley H, Connor A, Chen R, Ross M, et al. Development of a culturally tailored text message maternal health program: TextMATCH. JMIR Mhealth Uhealth 2017 Apr 20;5(4):e49 [FREE Full text] [doi: 10.2196/mhealth.7205] [Medline: 28428159]
17. Broom MA, Ladley AS, Rhyne EA, Halloran DR. Feasibility and perception of using text messages as an adjunct therapy for low-income, minority mothers with postpartum depression. JMIR Ment Health 2015;2(1):e4 [FREE Full text] [doi: 10.2196/mental.4074] [Medline: 26543910]
18. Marcolino MS, Oliveira JA, D'Agostino M, Ribeiro AL, Alkmim MB, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
19. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y, Dobson R. Mobile phone text messaging and app-based interventions for smoking cessation. Cochrane Database Syst Rev 2019 Oct 22;10:CD006611 [FREE Full text] [doi: 10.1002/14651858.CD006611.pub5] [Medline: 31638271]

20. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. JMIR Mhealth Uhealth 2020 Mar 18;8(3):e17046 [FREE Full text] [doi: 10.2196/17046] [Medline: 32186518]
21. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can smartphone apps increase physical activity? systematic review and meta-analysis. J Med Internet Res 2019 Mar 19;21(3):e12053 [FREE Full text] [doi: 10.2196/12053] [Medline: 30888321]
22. Hobson GR, Caffery LJ, Neuhaus M, Langbecker DH. Mobile health for first nations populations: systematic review. JMIR Mhealth Uhealth 2019 Oct 07;7(10):e14877 [FREE Full text] [doi: 10.2196/14877] [Medline: 31593537]
23. Ni Mhurchu C, Te Morenga L, Tupai-Firestone R, Grey J, Jiang Y, Jull A, et al. A co-designed mHealth programme to support healthy lifestyles in Māori and Pasifika peoples in New Zealand (OL@-OR@): a cluster-randomised controlled trial. Lancet Digit Health 2019 Oct;1(6):e298-e307 [FREE Full text] [doi: 10.1016/S2589-7500(19)30130-X] [Medline: 33323252]
24. Peiris D, Wright L, News M, Rogers K, Redfern J, Chow C, et al. A smartphone app to assist smoking cessation among aboriginal Australians: findings from a pilot randomized controlled trial. JMIR Mhealth Uhealth 2019 Apr 02;7(4):e12745 [FREE Full text] [doi: 10.2196/12745] [Medline: 30938691]
25. Sweet MA. Social media: new links for Indigenous health. Med J Aust 2013 Jul 08;199(1):18. [doi: 10.5694/mja13.10429] [Medline: 23829246]
26. McPhail-Bell K, Appo N, Haymes A, Bond C, Brough M, Fredericks B. Deadly choices empowering indigenous Australians through social networking sites. Health Promot Int 2018 Oct 01;33(5):770-780. [doi: 10.1093/heapro/dax014] [Medline: 28387801]
27. Finlay S, Wenitong M. Aboriginal community controlled health organisations are taking a leading role in COVID-19 health communication. Aust N Z J Public Health 2020 Aug;44(4):251-252 [FREE Full text] [doi: 10.1111/1753-6405.13010] [Medline: 32583534]
28. Petkovic JD, Duench S, Trawin J, Dewidar O, Pardo Pardo J, Simeon R, et al. Behavioural interventions delivered through interactive social media for health behaviour change, health outcomes, and health equity in the adult population. Cochrane Database Syst Rev 2021 May 31;5:CD012932. [doi: 10.1002/14651858.CD012932.pub2] [Medline: 34057201]
29. Verbiest ME, Corrigan C, Dalhousie S, Firestone R, Funaki T, Goodwin D, et al. Using codesign to develop a culturally tailored, behavior change mHealth intervention for indigenous and other priority communities: a case study in New Zealand. Transl Behav Med 2019 Jul 16;9(4):720-736. [doi: 10.1093/tbm/iby093] [Medline: 30388262]
30. Ethical guidelines: key principles (2020). Aboriginal Health and Medical Research Council of NSW. URL: https://www.ahmrc.org.au/ethics-main-page/ [accessed 2022-04-29]
31. Bratteteig T, Bødker K, Dittrich Y, Holst MP, Simonsen J. Methods: Organising principlesgeneral guidelines for participatory design projects. In: Simonsen J, Robertson T, editors. International Handbook of Participatory Design. London, UK: Routledge; 2012.
32. Hall KK, Chang AB, Anderson J, Arnold D, Goyal V, Dunbar M, et al. The incidence and short-term outcomes of acute respiratory illness with cough in children from a socioeconomically disadvantaged urban community in Australia: a community-based prospective cohort study. Front Pediatr 2017 Oct 31;5:228 [FREE Full text] [doi: 10.3389/fped.2017.00228] [Medline: 29164080]
33. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
34. Champion V, Skinner C. The health belief model. In: Health Behavior and Health Education: Theory, Research, and Practice; 4th edition. Hoboken, NJ, US: John Wiley & Sons; Sep 12, 2008.
35. 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]
36. Michie S, Free C, West R. Characterising the 'Txt2Stop' smoking cessation text messaging intervention in terms of behaviour change techniques. J Smok Cessat 2012 Aug 01;7(1):55-60. [doi: 10.1017/jsc.2012.12]
37. 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]
38. Boyd H, McKernon S, Mullin B, Old A. Improving healthcare through the use of co-design. N Z Med J 2012 Jun 29;125(1357):76-87. [Medline: 22854362]
39. Who are the stolen generations? Healing Foundation. 2021. URL: https://healingfoundation.org.au/who-are-the-stolen-generations/ [accessed 2022-04-29]
40. Te Morenga L, Pekepo C, Corrigan C, Matoe L, Mules R, Goodwin D, et al. Co-designing an mHealth tool in the New Zealand Māori community with a "Kaupapa Māori" approach. AlterNative Int J Indigenous People 2018 Jan 23;14(1):90-99. [doi: 10.1177/1177180117753169]

41. Moore G, Campbell M, Copeland L, Craig P, Movsisyan A, Hoddinott P, et al. Adapting interventions to new contexts-the ADAPT guidance. BMJ 2021 Aug 03;374:n1679 [FREE Full text] [doi: 10.1136/bmj.n1679] [Medline: 34344699]
42. Canuto KJ, Aromataris E, Burgess T, Davy C, McKivett A, Schwartzkopff K, et al. A scoping review of Aboriginal and Torres Strait Islander health promotion programs focused on modifying chronic disease risk factors. Health Promot J Austr 2021 Jan 09;32(1):46-74 [FREE Full text] [doi: 10.1002/hpja.307] [Medline: 31724783]
43. Calma T, Dudgeon P, Bray A. Aboriginal and Torres Strait Islander social and emotional wellbeing. In: Working Together: Aboriginal and Torres Strait Islander Mental Health and Wellbeing Principles and Practice. Canberra: Commonwealth Government of Australia; 2014.
44. Arabena K, Rowley K, MacLean S. Building evidence about effective health promotion in Aboriginal and Torres Strait Islander communities. Aust J Prim Health 2014;20(4):317-318. [doi: 10.1071/PYv20n4_ED] [Medline: 25354980]
45. Deadly numbers 2018. Deadly Choices. URL: https://deadlychoices.com.au/about/deadly-numbers/ [accessed 2022-04-29]
46. Eyles H, Jull A, Dobson R, Firestone R, Whittaker R, Te Morenga LT, et al. Co-design of mHealth delivered interventions: a systematic review to assess key methods and processes. Curr Nutr Rep 2016 Jul 4;5(3):160-167. [doi: 10.1007/s13668-016-0165-7]

Abbreviations

AMS: Aboriginal Medical Service
mHealth: mobile health
NSW: New South Wales

Edited by A Mavragani; submitted 28.10.21; peer-reviewed by U Bork, C Baxter; comments to author 02.01.22; revised version received 16.03.22; accepted 18.03.22; published 08.07.22.

Please cite as:
Perkes SJ, Huntriss B, Skinner N, Leece B, Dobson R, Mattes J, Hall K, Bonevski B
Development of a Maternal and Child mHealth Intervention With Aboriginal and Torres Strait Islander Mothers: Co-design Approach
JMIR Form Res 2022;6(7):e33541
URL: https://formative.jmir.org/2022/7/e33541
doi:10.2196/33541
PMID:35802404

Original Paper

Developing Educational Animations on HIV Pre-exposure Prophylaxis (PrEP) for Women: Qualitative Study

Anna Marie Young[1], MD, MPH; Timothee Fruhauf[1], MD, MPH; Obianuju Okonkwo[1], BSc; Erin Gingher[1], RN, MS, MHS; Jenell Coleman[1], MD, MPH

Department of Gynecology and Obstetrics, Johns Hopkins University School of Medicine, Baltimore, MD, United States

Corresponding Author:
Jenell Coleman, MD, MPH
Department of Gynecology and Obstetrics
Johns Hopkins University School of Medicine
600 N Wolfe Street, Phipps Room 254A
Baltimore, MD
United States
Phone: 1 410 614 4496
Email: colemanj@jhmi.edu

Abstract

Background: Despite experiencing the second-highest rate of HIV incidence in the United States, pre-exposure prophylaxis (PrEP) use remains low among Black women due, in part, to a lack of patients' awareness and providers' knowledge.

Objective: Our goal was to design animated educational tools informed by patients and women's health providers to address these barriers, specifically for women at risk for HIV.

Methods: Two animation storyboards about PrEP for women were created by academic stakeholders (eg, HIV clinical experts, educators, and HIV peer counselors), one for patients and one for providers. Four focus groups with community members from Baltimore, Maryland and four with women's health providers (eg, obstetrician/gynecologists, midwives, nurse practitioners, and peer counselors) at an academic center were conducted to discuss the storyboards. Transcripts were analyzed using conventional content analysis, and themes were incorporated into the final versions of the animations.

Results: Academic stakeholders and 30 focus group participants (n=16 female community members and n=14 women's health providers) described important themes regarding PrEP. The themes most commonly discussed about the patient animation were understandability of side effects, HIV risk factors, messaging, PrEP access, and use confidence. Provider animation themes were indications for PrEP, side effects, and prescribing confidence.

Conclusions: We created two PrEP animations focused on women. Stakeholder feedback highlighted the importance of ensuring the understandability and applicability of PrEP educational materials while including necessary information to facilitate use or prescribing confidence. Both community members and women's health providers reported greater use confidence after viewing the animations.

(JMIR Form Res 2022;6(7):e33978) doi:10.2196/33978

KEYWORDS

PrEP; animations; education; HIV; prevention; women

Introduction

Black women in the United States experience the second-highest incidence of HIV behind men who have sex with men and account for over half of new HIV diagnoses among women [1]. Many population-level factors contribute to this risk, including social and economic inequities that influence sexual networks, stigma, discrimination, and inadequate access to HIV care [2,3].

Oral HIV pre-exposure prophylaxis (PrEP) with antiretroviral drugs emtricitabine and tenofovir is a Food and Drug Administration–approved medication that prevents HIV in women up to 90% when taken daily [4-6]. The Centers for Disease Control and Prevention recommends that all sexually active adolescents and adults be informed about PrEP [7]. Furthermore, the American College of Obstetrics and Gynecology recommends PrEP for women who are at substantial risk of acquiring HIV, including those who have an HIV-positive

or unknown status sexual partner, a recent sexually transmitted infection, a high number of sexual partners, report inconsistent or no condom use, participate in commercial sex work, live in a high HIV prevalent area, or inject drugs [4].

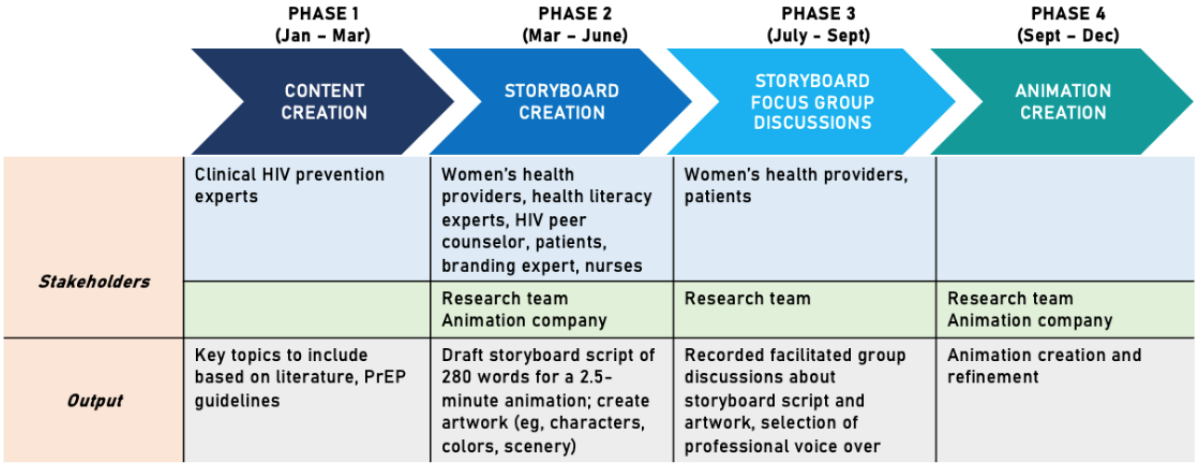
Despite the availability of PrEP, uptake has been poor among women at high risk [8,9]. Patient-level barriers to PrEP include low self-perceived risk of HIV, limited knowledge, and high perceived cost [10,11]. Provider-level barriers include poor support and infrastructure to provide PrEP, inadequate education, and underestimating patients' risk [10,11]. Additionally, few PrEP campaigns specifically target women or women's health providers [10,11].

Computer-based interventions, such as animations, have been associated with decreased high-risk behaviors leading to HIV acquisition [12]. Furthermore, creating multimedia tools for health education with stakeholder involvement has been encouraged to identify specific community needs and ensure effective dissemination [13,14]. Additionally, animations can decrease cognitive overload and increase attention retention and long-term recall [14,15]. Therefore, to address commonly cited barriers to PrEP uptake among women, we sought to create two women-centered PrEP animations, one for providers and one for patients, grounded in Mayer's [15] Cognitive Theory of Multimedia Learning that posits combined auditory text and visual pictures deepens understanding more than either alone. Below, we describe the animation development process with the participation of community members and women's health providers from Baltimore, Maryland.

Methods

Animations were created and iteratively refined in four phases between January 2020 and December 2020 (Figure 1).

Figure 1. Study phases. PrEP: pre-exposure prophylaxis.



Ethics Approval

This study was approved by the Johns Hopkins University School of Medicine Institutional Review Board (approval number: IRB00252170).

Phase 1: Animation Content Creation

Clinical HIV prevention experts (eg, obstetrician/gynecologists with fellowship training in HIV care and HIV peer counselors) created outlines for the animations based on published literature and existing animations about PrEP [11,16-18]. Proposed content included background information about HIV and PrEP among women, risk factors for HIV, indications for PrEP, side effects of PrEP, and a step-by-step guide to taking PrEP or prescribing PrEP.

Phase 2: Storyboard Creation

Two storyboards (ie, scripts and 2D slides) were created by an animation company (Science Animated, Cotswolds, United Kingdom) based on the outlines written in phase 1. The patient storyboard used simple language to aid low–health literacy populations, had relatable characters, and used positive framing. The provider storyboard assumed prior knowledge of women's sexual health and was framed for practicality using medical terminology. In addition, we held formative discussions with our academic stakeholders (ie, lay individuals, a patient education professional, a branding director, nurses, and clinical HIV prevention experts), and storyboards were revised based on these initial discussions.

Phase 3: Focus Groups

Feedback about the storyboards from a larger audience was gathered. English-speaking women from the community and women's health providers (eg, obstetrician/gynecologists, midwives, nurse practitioners, and peer counselors) were recruited for focus groups. Fliers were placed in all general gynecology clinics, all academic specialist obstetrician/gynecologists in an academic hospital were emailed about participating, and prior research participants who had agreed to be contacted for future studies were recruited. Purposive sampling was conducted.

After obtaining written consent and collecting demographic information, the focus groups were scheduled. A trained facilitator and logistical coordinator conducted focus groups virtually using a secure videoconferencing platform (Zoom 5.5.4; Zoom Video Communications, Inc). Participants were offered the option to turn off their video and remove their names

to allow partial anonymity, as all still heard voices. There were 8 focus groups, 4 for community members and 4 for women's health providers. Each focus group contained 3 to 6 participants and lasted 60 to 90 minutes. A semistructured focus group guide was used to frame the discussions. Near the end of each focus group, participants were asked to rank 6 female-sounding professional voice-over actors who read the same script but may have had differences in inflection, cadence, pitch, or articulation. Participants received a US $25 compensatory gift card.

Phase 4: Focus Group Analysis and Animation Creation

The focus group audio was recorded and transcribed verbatim. The transcripts were coded using ATLAS.ti 9.0.3 (ATLAS.ti Scientific Software Development GmbH), and the findings were organized and analyzed using conventional content analysis [19]. Two research team members read the transcripts twice, line-by-line, and prepared memos summarizing their preliminary findings. Next, preliminary codes were derived inductively by highlighting recurring words or statements through an iterative process. Research team members convened multiple times to discuss and compare memos, and revisit emerging themes iteratively. Discrepancies were solved by the principal investigator. Final codes were then assessed for broader concepts to generate themes through subsequent rounds of team discussion. The generation of themes was guided by an adaptation of the Model of Communication and Health Behavior Change by Kincaid [20]. The themes identified in the storyboards (phase 2) and focus group discussions were incorporated into animation prototypes. The highest-ranked voice-over options were chosen. Finally, two 120-second 2D animations were created and iteratively refined by the research team.

Results

Storyboard Creation (Phases 1 and 2)

Themes that were presented in the initial creation of the storyboard included accurate PrEP information, ensuring an appropriate health literacy level to reflect the target population (eg, proficient level for the provider animation and basic to below basic level for patient animation) and representative graphics/artwork (eg, multicultural characters and scenery). The scripts for the storyboards and animation graphics were refined numerous times by stakeholders.

Focus Groups (Phase 3)

A total of 30 participants enrolled in the focus groups (n=16 female community members and n=14 providers, Table 1). Some themes pertained to both animations (messaging, design, background, side effects and risk factors, and perceptions of PrEP access and barriers), while others were only relevant to the patient animation (relatability and applicability of characters and storyline, and PrEP use confidence) or the provider animation (understandability of PrEP indications and prescribing confidence).

Table 1. Sociodemographic characteristics of focus group participants (N=30).

Characteristic	Community member (n=16)	Women's health provider (n=14)
Age (years), mean (SD)	26 (6)	42 (13)
Sex, n (%)		
Male	0 (0)	2 (14)
Female	16 (100)	12 (86)
Self-reported race, n (%)		
Black	12 (75)	3 (21)
Asian	0 (0)	3 (21)
White	1 (6)	7 (50)
Other	3 (19)	1 (7)
Highest education level, n (%)		
High school	7 (44)	0 (0)
College	6 (38)	1 (7)
Graduate school	2 (13)	13 (93)
Other	1 (6)	0 (0)
Marital status, n (%)		
Single	8 (50)	6 (43)
Married/union	7 (44)	7 (50)
Divorced	1 (6)	1 (7)

Provider Feedback

Providers' most common themes included prescribing confidence, indications for PrEP, and side effects. Additional themes and respective quotes are highlighted in Multimedia Appendix 1.

Indications

Providers were surprised by a few of the indications for prescribing PrEP, specifically in relation to the area in which their patients were living. Several of them were surprised to learn that the HIV prevalence in a location was included in evaluating HIV risk and PrEP indications.

> *So, automatically living in Baltimore, it puts you at a higher risk for HIV, so really understanding that kinda stood out to me.*

Side Effects

The level of details on side effects and their timing was a concern for some providers. Providers asked for more clarification on which side effects to expect and a given time frame for each to be included in the animation.

> *I wanna like know [...] how long is it [resolution of initial side-effects] gonna take? Are we talking about the next day? Are we talking about a month?*

Prescribing Confidence

Providing a more detailed step-by-step guide applicable to the flow of a typical clinic encounter, including follow-up steps after prescribing PrEP, was recommended.

> *I'd first ask myself like, "Hmm, this person seems like they're high risk for HIV." And then, I would ask myself, "Do I – do I have testing that allows me to firmly confirm or deny the fact that they actually have HIV right now?" [...] And then, I would say to myself, like, "Okay, so I think if they like don't have a test on file or it's not that recent, are there any signs that I think that they're nonetheless actively infected with HIV and I need to test them for that before, you know, starting a conversation about preventing HIV."*

Providers were, overall, confident in their ability to prescribe PrEP based on these steps but expressed doubt about their patient's desire and ability to comply with extensive follow-up and regular lab draws.

> *You're gonna have people who are not gonna wanna come in for a HIV test every three months...I mean, that's just gonna be a deterrent for people.*

Community Member Feedback

Community members' most common themes were understandability of side effects and risk factors, messaging, PrEP use, PrEP access, and use confidence.

Risk Factors

Community members were surprised by the prevalence of HIV in their community and that they themselves would qualify for PrEP based on the listed risk factors.

> *I would say it related to me because before then, I never knew about PrEP [...] And I think that it will be not only a big eye-opener for me but for everyone else.*
>
> *I didn't know that – um, that Baltimore City, HIV was as high as it is. They had shared it on the news, I think like a week or two ago, and liked it just kinda like caught me off-guard...*

Side Effects

Community members were concerned about serious side effects and wanted more information about drug-drug interactions, specifically interactions between PrEP and contraceptive methods.

> *The stomach and the headaches, [...] that's kinda common. But like generally, kidney and bone density, that's not like average things.*
>
> *Um, you said that it doesn't affect pregnancy or anything like that, but is there any risk – Like, if I'm on birth control, and I supplement with PrEP, is there any effect there, or they don't affect each other whatsoever?*

PrEP Use

Community members highlighted confusing concepts, including PrEP's ability to prevent HIV and the logistics of PrEP follow-up that would need to be clarified to encourage use.

> *How long does it last? Is it like a shot? Well, I know it's like a pill, but like how long does it last? Like do you have to take one every day, every week, once a month?*

Messaging

Overall, community members reported that the language and the messaging were appropriate for all education levels.

> *So, if we can clear it up that – Yeah, condoms alone do prevent HIV acquisition, but it's much more effective if you use PrEP. And if you do both together that's even better. Um, so maybe there is a way that we could kind of, like, get that message across.*

PrEP Access

Community members thought use confidence would be impacted by information about PrEP access, especially insurance coverage. They wanted the animation to convey that it was easy to take PrEP via the step-by-step guide, which motivated participants to recommend PrEP to their peers.

> *I think the most surprising thing for me is that there is an option for people without insurance [...] I don't know how true that is. Because y'all always say that, but they'll be, like, "Yeah. There's an option. You can take off 10 percent." That's not enough.*

Use Confidence

Overall, there was a potential for greater use confidence after viewing the storyboard. Community members expressed that

the storyboard motivated them to read more about PrEP and initiate a discussion with their provider.

> *But now, after this focus group, I'm more interested because it was kinda well-explained. I will do my own research on like the bone density and the kidneys and the side-effects, but I think after this focus group that, uh, it's something that I will have a conversation with my doctor about.*
>
> *...But, yes, I would. It's very simple, it's appealing. Um, if it's 90% accurate plus on top of a condom, um, especially if you have multiple partners. Why not?*

Animation Creation (Phase 4)

Stakeholder feedback from all phases led to clarifying language modifications and additional detail to describe the background, risk factors, indications, and steps for prescribing or accessing PrEP. Specifically, the provider animation was modified to clearly delineate prescribing steps. Details about side effects and interactions were added to the patient animation. The importance of evaluating HIV risk and PrEP eligibility according to risk factors, including geographic HIV prevalence, was explained better. Finally, design changes to the characters were made to make them more relatable (Multimedia Appendixes 2 and 3).

Discussion

Principal Findings

Two educational animations to facilitate learning about HIV prevention and PrEP for female patients and providers were created using a user-centered approach. There were some similar themes both community members and women's health providers wanted to highlight in the animations that included a clear demonstration about indications for PrEP, addressing barriers to PrEP use, and providing step-by-step guides to accessing or prescribing PrEP. These themes were considered the most important for both patients and providers to increase PrEP awareness and uptake among at-risk women. In addition, there was greater use confidence after viewing the storyboards.

Comparison to Prior Work

Although we did not test the "real-world" effectiveness of the final PrEP animations in this formative study, in general, animations have been found to be effective in increasing health information recall [12,16,21]. One study used a 2 × 2 factorial design among patients with different health literacy levels to determine which features of animations improved health information recall and attitudes [21]. They found that spoken animation significantly improved recall of health information compared to written messages among low-literacy participants (P=.02). Additionally, there was no differences in health information recall between high- and low-literacy participants after exposure to spoken animation (P=.12). Furthermore, a meta-analysis demonstrated that technology-based HIV prevention interventions have been proven to be at least as efficacious as human-delivered interventions in reducing high-risk sexual behaviors [12].

Strengths and Limitations

A notable strength of our study is the user-centered approach with key stakeholders, which has been proven to foster stronger relationships between researchers and the community [22]. This may allow greater dissemination among at-risk women and women's health providers. However, there are limitations. First, the generalizability of these findings to other locations may be limited. We recruited providers from a single tertiary care center, who do not represent all health care providers in different settings. However, our community members reflect women most impacted by the HIV epidemic and that share similar characteristics. Second, although we collected data until we thought that theme saturation was reached, the sample of participants was small and additional themes might have been missed. Third, the animations did not provide exhaustive information and were not tested for effectiveness. However, the purpose of the animations will be to facilitate a discussion between female patients and women's health providers as an adjunct to routine sexual health care. Additionally, we do not expect our animations to be less effective than other human-delivered interventions, as existing data has shown that technological interventions are effective [12].

Conclusion

To increase the use of PrEP in women who live in communities with high HIV risk, the dissemination of information regarding its use in a relatable, accessible, and applicable way is vital for both patients and providers. Therefore, we included stakeholders in creating short educational PrEP animations. Stakeholders highlighted important issues to them, which included identifying individuals that qualified for PrEP, delineating key steps in accessing or prescribing PrEP, and addressing barriers to PrEP. As a result, there was greater use confidence for community members and women's health providers after viewing the storyboards. Future research is planned to evaluate the effectiveness of the animations to increase PrEP awareness and uptake among women who are at substantial risk for HIV.

Acknowledgments

The authors thank the focus group participants and Rosemary Ramroop, Janelle Ramroop, Gregoria Valdez, Tom Bauer, and Tanique Bennett for reviewing early versions of the focus group guides and animation storyboards.

Authors' Contributions

AMY was responsible for facilitating the focus groups, data analysis, data interpretation, modification of scripts for the animations, and drafting of all manuscript sections. TF was responsible for facilitating the focus groups, data analysis, data interpretation,

creating and modifying scripts for the animations, and drafting all manuscript sections. OO was responsible for facilitating and recording the focus groups, study coordination, participant follow-up, and editing and reviewing the manuscript. EG was responsible for analyzing and interpreting the data, modifying the animation scripts, and editing and reviewing the manuscript. JC posed the research question for this manuscript and was responsible for data interpretation, analysis, editing, and manuscript review.

Conflicts of Interest

JC's institution receives research funding from Gilead Sciences.

Multimedia Appendix 1
Supplemental Tables 1 and 2.
[DOCX File , 19 KB - formative_v6i7e33978_app1.docx]

Multimedia Appendix 2
Provider animation.
[MP4 File (MP4 Video), 6960 KB - formative_v6i7e33978_app2.mp4]

Multimedia Appendix 3
Patient animation.
[MP4 File (MP4 Video), 7944 KB - formative_v6i7e33978_app3.mp4]

References

1. HIV and women. Centers for Disease Control and Prevention. 2018. URL: https://www.cdc.gov/hiv/group/gender/women/index.html [accessed 2021-06-11]
2. Adimora AA, Schoenbach VJ. Contextual factors and the black-white disparity in heterosexual HIV transmission. Epidemiology 2002 Nov;13(6):707-712. [doi: 10.1097/00001648-200211000-00016] [Medline: 12410013]
3. Adimora A, Schoenbach V. Social context, sexual networks, and racial disparities in rates of sexually transmitted infections. J Infect Dis 2005 Feb 01;191 Suppl 1:S115-S122. [doi: 10.1086/425280] [Medline: 15627221]
4. Committee on Gynecologic Practice. ACOG Committee Opinion no 595: Committee on Gynecologic Practice: preexposure prophylaxis for the prevention of human immunodeficiency virus. Obstet Gynecol 2014 May;123(5):1133-1136. [doi: 10.1097/01.AOG.0000446855.78026.21] [Medline: 24785877]
5. Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. World Health Organization. 2015. URL: http://apps.who.int/iris/bitstream/handle/10665/186275/9789241509565_eng.pdf [accessed 2021-06-14]
6. Flash C, Dale S, Krakower D. Pre-exposure prophylaxis for HIV prevention in women: current perspectives. Int J Womens Health 2017;9:391-401. [doi: 10.2147/IJWH.S113675] [Medline: 28615975]
7. Tanner MR, Miele P, Carter W, Valentine SS, Dunville R, Kapogiannis BG, et al. Preexposure prophylaxis for prevention of HIV acquisition among adolescents: clinical considerations, 2020. MMWR Recomm Rep 2020 Apr 24;69(3):1-12 [FREE Full text] [doi: 10.15585/mmwr.rr6903a1] [Medline: 32324724]
8. Seidman D, Weber S. Integrating preexposure prophylaxis for human immunodeficiency virus prevention into women's health care in the United States. Obstet Gynecol 2016 Jul;128(1):37-43. [doi: 10.1097/AOG.0000000000001455] [Medline: 27275793]
9. Fruhauf T, Coleman J. A missed opportunity for U.S. perinatal human immunodeficiency virus elimination: pre-exposure prophylaxis during pregnancy. Obstet Gynecol 2017 Oct;130(4):703-709 [FREE Full text] [doi: 10.1097/AOG.0000000000002258] [Medline: 28885420]
10. Johnson AK, Fletcher FE, Ott E, Wishart M, Friedman EE, Terlikowski J, et al. Awareness and intent to use pre-exposure prophylaxis (PrEP) among African American women in a family planning clinic. J Racial Ethn Health Disparities 2020 Jun;7(3):550-554. [doi: 10.1007/s40615-019-00683-9] [Medline: 31848943]
11. Ojikutu BO, Bogart LM, Higgins-Biddle M, Dale SK, Allen W, Dominique T, et al. Facilitators and barriers to pre-exposure prophylaxis (PrEP) use among Black individuals in the United States: results from the National Survey on HIV in the Black Community (NSHBC). AIDS Behav 2018 Nov;22(11):3576-3587 [FREE Full text] [doi: 10.1007/s10461-018-2067-8] [Medline: 29468493]
12. Noar SM, Black HG, Pierce LB. Efficacy of computer technology-based HIV prevention interventions: a meta-analysis. AIDS 2009 Jan 02;23(1):107-115. [doi: 10.1097/QAD.0b013e32831c5500] [Medline: 19050392]
13. Institute of Medicine (US) Forum on the Science of Health Care Quality Improvement and Implementation, Institute of Medicine (US) Roundtable on Health Disparities, Institute of Medicine (US) Roundtable on Health Literacy. Toward Health Equity and Patient-Centeredness: Integrating Health Literacy, Disparities Reduction, and Quality Improvement: Workshop Summary. Washington, DC: National Academies Press (US); 2009.
14. Gagliano ME. A literature review on the efficacy of video in patient education. J Med Educ 1988 Oct;63(10):785-792. [doi: 10.1097/00001888-198810000-00006] [Medline: 3050102]

15. Mayer RE, Moreno R. Animation as an aid to multimedia learning. Educ Psychol Rev 2002;14(1):87-99. [doi: 10.1023/A:1013184611077]
16. Bond KT, Ramos SR. Utilization of an animated electronic health video to increase knowledge of post- and pre-exposure prophylaxis for HIV Among African American women: nationwide cross-sectional survey. JMIR Form Res 2019 May 29;3(2):e9995 [FREE Full text] [doi: 10.2196/formative.9995] [Medline: 31144667]
17. Chapman Lambert C, Marrazzo J, Amico K, Mugavero M, Elopre L. PrEParing women to prevent HIV: an integrated theoretical framework to PrEP Black women in the United States. J Assoc Nurses AIDS Care 2018;29(6):835-848 [FREE Full text] [doi: 10.1016/j.jana.2018.03.005] [Medline: 29685648]
18. Calabrese SK, Earnshaw VA, Underhill K, Hansen NB, Dovidio JF. The impact of patient race on clinical decisions related to prescribing HIV pre-exposure prophylaxis (PrEP): assumptions about sexual risk compensation and implications for access. AIDS Behav 2014 Feb;18(2):226-240 [FREE Full text] [doi: 10.1007/s10461-013-0675-x] [Medline: 24366572]
19. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
20. Kincaid DL. Mass media, ideation, and behavior: a longitudinal analysis of contraceptive change in the Philippines. Commun Res 2016 Jun 30;27(6):723-763. [doi: 10.1177/009365000027006003]
21. Meppelink CS, van Weert JCM, Haven CJ, Smit EG. The effectiveness of health animations in audiences with different health literacy levels: an experimental study. J Med Internet Res 2015 Jan 13;17(1):e11 [FREE Full text] [doi: 10.2196/jmir.3979] [Medline: 25586711]
22. Newman PA. Towards a science of community engagement. Lancet 2006 Jan 28;367(9507):302. [doi: 10.1016/S0140-6736(06)68067-7] [Medline: 16443036]

Abbreviations

PrEP: pre-exposure prophylaxis

Edited by T Leung; submitted 10.10.21; peer-reviewed by C Kakalou, C Hemingway; comments to author 13.01.22; revised version received 03.04.22; accepted 11.05.22; published 08.07.22.

Please cite as:
Young AM, Fruhauf T, Okonkwo O, Gingher E, Coleman J
Developing Educational Animations on HIV Pre-exposure Prophylaxis (PrEP) for Women: Qualitative Study
JMIR Form Res 2022;6(7):e33978
URL: https://formative.jmir.org/2022/7/e33978
doi:10.2196/33978
PMID:35802416

Original Paper

Adaptation of a Problem-solving Program (Friendship Bench) to Treat Common Mental Disorders Among People Living With HIV and AIDS and on Methadone Maintenance Treatment in Vietnam: Formative Study

Ha V Tran[1], MSc, MD; Ha T T Nong[1], PhD; Thuy T T Tran[2], PhD; Teresa R Filipowicz[3], MPH; Kelsey R Landrum[3], MPH; Brian W Pence[3], PhD; Giang M Le[4], PhD; Minh X Nguyen[5], PhD; Dixon Chibanda[6,7], PhD; Ruth Verhey[6], PhD; Vivian F Go[8], PhD; Hien T Ho[9], PhD; Bradley N Gaynes[10], MPH, MD

[1]The University of North Carolina, Vietnam Office, Hanoi, Vietnam

[2]Faculty of Environmental and Occupational Health, Hanoi University of Public Health, Hanoi, Vietnam

[3]Department of Epidemiology, University of North Carolina, Chapel Hill, NC, United States

[4]Center for Research and Training in HIV/AIDS, Hanoi Medical University, Hanoi, Vietnam

[5]Faculty of Epidemiology Department, Hanoi Medical University, Hanoi, Vietnam

[6]Department of Community Medicine & Research Support Centre, University of Zimbabwe, Harare, Zimbabwe

[7]London School of Hygiene & Tropical Medicine, London, United Kingdom

[8]Department of Health Behavior, Gillings, School of Global Public Health, University of North Carolina, Chapel Hill, NC, United States

[9]Faculty of Clinical Medicine, Hanoi University of Public Health, Hanoi, Vietnam

[10]Department of Psychiatry, University of North Carolina, Chapel Hill, NC, United States

Corresponding Author:
Ha V Tran, MSc, MD
The University of North Carolina
Vietnam Office
Room 407 – 408, Block A2
Van Phuc Diplomatic Compound 298 Kim Ma, Ba Dinh
Hanoi, 10000
Vietnam
Phone: 84 24 3211 5839 ext 108
Email: vietha@live.unc.edu

Abstract

Background: The prevalence of common mental disorders (CMDs) among people living with HIV and people who inject drugs is high worldwide and in Vietnam. However, few evidence-informed CMD programs for people living with HIV who inject drugs have been adapted for use in Vietnam. We adapted the *Friendship Bench* (FB), a problem-solving therapy (PST)–based program that was successfully implemented among patients with CMDs in primary health settings in Zimbabwe and Malawi for use among people living with HIV on methadone maintenance treatment (MMT) with CMDs in Hanoi, Vietnam.

Objective: This study aimed to describe the adaptation process with a detailed presentation of 4 phases from the third (adaptation) to the sixth (integration) of the Assessment-Decision-Adaptation-Production-Topical Experts-Integration-Training-Testing (ADAPT-ITT) framework.

Methods: The adaptation phase followed a qualitative study design to explore symptoms of CMDs, facilitators, and barriers to conducting FB for people living with HIV on MMT in Vietnam, and patient, provider, and caretaker concerns about FB. In the production phase, we revised the original program manual and developed illustrated PST cases. In the topical expert and integration phases, 2 investigators (BNG and BWP) and 3 subject matter experts (RV, DC, and GML) reviewed the manual, with reviewer comments incorporated in the final, revised manual to be used in the training. The draft program will be used in the training and testing phases.

Results: The study was methodologically aligned with the ADAPT-ITT goals as we chose a proven, effective program for adaptation. Insights from the adaptation phase addressed the who, where, when, and how of FB program implementation in the

MMT clinics. The ADAPT-ITT framework guided the appropriate adaptation of the program manual while maintaining the core components of the PST of the original program throughout counseling techniques in all program sessions. The deliverable of this study was an adapted FB manual to be used for training and piloting to make a final program manual.

Conclusions: This study successfully illustrated the process of operationalizing the ADAPT-ITT framework to adapt a mental health program in Vietnam. This study selected and culturally adapted an evidence-informed PST program to improve CMDs among people living with HIV on MMT in Vietnam. This adapted program has the potential to effectively address CMDs among people living with HIV on MMT in Vietnam.

Trial Registration: ClinicalTrials.gov NCT04790201; https://clinicaltrials.gov/ct2/show/NCT04790201

(JMIR Form Res 2022;6(7):e37211) doi:10.2196/37211

KEYWORDS

Friendship Bench; Vietnam; Assessment-Decision-Adaptation-Production-Topical Experts-Integration-Training-Testing; ADAPT-ITT; common mental disorders; people living with HIV; PWH; people who inject drugs; PWID; methadone maintenance treatment; MMT; depression; anxiety; stress disorder

Introduction

Injection drug use is the main cause of the HIV epidemic in Vietnam, resulting in a high HIV prevalence among people who inject drugs, 15% in 2017 [1] and 12.7% in 2020 [2]. Antiretroviral therapy (ART) and methadone maintenance treatment (MMT) have been implemented in Vietnam since 2004 and 2008, respectively, and people living with HIV who inject drugs are commonly treated with simultaneous ART and MMT [3,4]. HIV and drug addiction are chronic stressors on mental health, and when combined with many other barriers related to HIV and drug addiction such as stigma and discrimination, the prevalence of common mental disorders (CMDs) among people living with HIV who inject drugs is higher than in the general population [5,6]. CMD is a collective term that refers to a range of depressive, anxiety, and stress-related disorders [7]. Depression is characterized by sadness, low self-esteem, tiredness, lack of concentration, lack of interest, and sleeping difficulties. Anxiety disorders refer to the feeling of anxiety, fear, panic including posttraumatic disorder, and social anxiety [7]. Stress is a state of physical and emotional tension caused by reactions to stressful stimuli from life and the environment [8]. The prevalence of depression and anxiety disorder among MMT clients in China in 2017 was 12.8% and 19.5% [9], compared with only 4.4% and 3.6% in the general population in 2015, respectively [7]. In a study by Levintow et al [10], people living with HIV who inject drugs in North Vietnam had a mild depression rate of 25% and moderate depression rate of 44%. Mughal et al [11] reported that people living with HIV on MMT in Hanoi had CMDs of depression and anxiety and posttrauma stress rate of 17% [11]. Prior research has demonstrated the effectiveness of behavioral and mental health programs among people living with HIV [12], including an evidence-informed program (EIP) to decrease levels of distress among people living with HIV in Thai Lan [13], an adapted life step intervention to effectively improve depression among people living with HIV in Zimbabwe [14], and an adapted CMD program for people living with HIV proved effective in 2020 in Malawi [15]. In Vietnam, we found evidence that family group counseling improved depression symptoms among people who inject drugs in 2017 [16]. Other studies have also recommended the need for CMD programs for people who inject drugs in Vietnam [17]. Problematically, there have been no mental health programs for this population and setting.

Friendship Bench (FB), developed by Chibanda Dixon in 2006 [18], is a brief evidence-informed psychological program using problem-solving therapy (PST) and implemented by lay counselors. More than 50,000 people have received counseling in the FB program, making it the largest mental health program integrated into primary health care in Africa, supported by the government and the Ministry of Health of Zimbabwe [18]. FB has been validated through a clinical trial to reduce CMDs among community members in Zimbabwe. In this study, people receiving FB sessions had lower depression and anxiety scores on the Shona Symptoms Questionnaire (SSQ-14) than those who did not receive FB sessions (3.81 and 8.90, respectively). Moreover, the percentage of people in the intervention arm who had depression symptoms on the Patient Health Questionnaire-9 was lower than that in the control arm (13.7% in the intervention arm compared with 49.9% in the control arm) [19]. FB was used in the provision of ART adherence and depression counseling sessions to people living with HIV in Malawi. A total of 501 people living with HIV who had just started ART and had mild depression (Patient Health Questionnaire-9 scores from 5 to 9) were randomized into 2 groups: the intervention group and the control group. The study indicated that people living with HIV in the intervention group had a higher rate of adherence to ART than those in the control group [15]. Recently, FB had been adapted to reduce depression and anxiety among people living with HIV adolescents in a clinical trial in Zimbabwe. The study of this adaptation is ongoing [20]. FB is also currently being adapted for use among people living with HIV who are pregnant in Malawi [21].

FB consists of up to 6 structured 45-minute sessions, and it has been effective in treating CMDs among community members in Zimbabwe [19,22]. The sessions were conducted outside of the primary clinics. Each session followed a PST approach and included steps to determine a manageable problem, choose a problem to solve, make a plan to solve the problem, create a follow-up of plan implementation, and encourage work on new problems. After completing an individual session, participants are invited to take part in group activities, where people facing similar life challenges share their stories, spend time together,

and work on income-generating activities [22]. The original FB manual has 23 chapters, covering the topics of counseling techniques and session structure, core competencies of counselors, mental illness, HIV and mental health, substance use disorders (SUD), and PST. However, to implement this program for a new population, we first simplified and restructured the program to meet local norms [23], the new target population, and the new study settings [24] via the Assessment-Decision-Adaptation-Production-Topical Experts-Integration-Training-Testing (ADAPT-ITT) model [25].

The adaptation method ADAPT-ITT of the Centers for Disease Control and Prevention (CDC) was developed by Wingood and DiClemente in 2008 as a framework to guide the adaptation of a proven EIP to the specific goals and study participants of HIV prevention [25]. The ADAPT-ITT model has been used to adapt HIV prevention EIP for HIV key populations such as African American couples in the United States [26], people living with HIV in Zimbabwe [27], African American women in the United States [28], and men who have sex with men in Thailand [29] and to address mental and sexual health for young people living with HIV in sub-Saharan Africa [30].

This study is nested within the parent study "Adaptation of the Friendship Bench counseling programme to improve mental health and HIV care engagement outcomes among people living with HIV who inject drugs in Vietnam," which aimed to test the feasibility, acceptability, and fidelity of the adapted FB in Vietnam [31]. This paper describes the process of FB adaptation to address CMDs among people living with HIV on MMT in some MMT clinics in Hanoi, Vietnam, in 2021, in preparation for a randomized controlled trial (RCT) following the ADAPT-ITT framework and specifically focuses on ADAPT-ITT phases 3 to 6.

Methods

Overview

The ADAPT-ITT model [25] comprises 8 sequential phases. Phase 1, the assessment, was conducted in a previous study [17]. Phase 2, the decision, was conducted by the parent study [31]. This paper describes in detail from the third to the sixth phases of the model: phase 3 (adaptation), phase 4 (production), phase 5 (topical experts), and phase 6 (integration). Phase 7 (training) and phase 8 (testing) will be presented in future papers.

Phase 1: Assessment

This phase involved needs assessment of people living with HIV on MMT for reduction in CMDs. The study conducted 28 in-depth interviews (IDIs) with people living with HIV who inject drugs (n=16, 57%), HIV and MMT providers (n=8, 29%), and health care providers (n=4, 14%) in private clinic rooms in Hanoi, Vietnam in May 2018.

Results of the study showed that both health care providers and people living with HIV who inject drugs believed that people living with HIV who inject drugs were particularly susceptible to CMDs, especially depression, and had a high need for mental health treatment. The study also recommended integrating mental health care into MMT clinics to increase mental health access to people living with HIV on MMT [17].

Phase 2: Decision

This phase involved reviewing evidence-informed CMD programs and deciding which EIP to select for the new target population. Few programs addressing CMDs for people living with HIV exist in low- and middle-income countries, which have limited mental health resources. One of the few evidence-informed CMD counseling programs is the "FB" model, which has extensive validation in low- and middle-income countries [19]. FB helped to reduce CMDs, which was the aim of this study. In addition, FB has also been adapted for people living with HIV in Zimbabwe and Malawi [15,20,21] and may be suitable for people living with HIV on MMT in Vietnam.

Phase 3: Adaptation

In phase 3, stakeholders including people living with HIV on MMT, their family members, and health care providers reviewed the original FB content and implementation methods to answer questions related to who, when, where, and how to deliver FB in Vietnam for appropriate implementation in MMT clinics. The study conducted IDIs with 12 people living with HIV on MMT, 5 family members and caretakers of the patients, and 4 clinic directors. In all, 2 focus group discussions (FGDs) were conducted with MMT health care providers (each group had 5 people).

Inclusion criteria for patients were being aged ≥18 years, willing to participate in the study, and having at least one of the symptoms: depression, anxiety, stress disorders as indicated by the Depression, Anxiety, and Stress Scale-21 items (DASS-21), consisting of 21 questions and 3 items measuring levels of the emotional states of depression, anxiety, and stress. The DASS-21 is based on a dimensional conception of depression, anxiety, and stress as part of the full human experience rather than the categorical approach of either having a disorder or not (as used in the Diagnostic and Statistical Manual of Mental Disorders or International Classification of Diseases) [32]. Although FB was initially tested using the SSQ-14 to screen for CMDs, this tool has not been validated in Vietnam. The DASS-21, however, has been validated in this setting. Originally developed by Lovibond [32] in Australia in 1983 and tested to screen for CMDs in the community, the DASS-21 has been translated and successfully validated for a variety of study participants in Vietnam, including adolescents [33], rural women [34], health care workers [35], and people living with HIV on MMT [17]. The DASS-21 is aligned with the SSQ-14 in terms of screening CMDs, including depression and anxiety disorders in the general population, but is distinguished by having a stress disorder scale. Thus, the study team chose the DASS-21 as a research tool, as it has the advantages of validity in this population, brevity, ease of implementation, and measurement of depression, anxiety disorder, and stress disorder relevant to the purpose of the study. The study will also use suicide risk assessment that was effectively used in the HPTN074 study for people living with HIV who inject drugs in Vietnam [36].

Inclusion CMDs scores were at least a moderate level in which depression subscale score was ≥14, anxiety subscale score was ≥10, and stress subscale score was ≥19. People who were aged <18 years, did not meet the DASS-21 moderate score, had cognitive impairments, or had suicidal thoughts (exclusion, therefore, included referral for standard clinical follow-up care) were excluded from the study. Patients who met the inclusion criteria were invited to participate in an IDI. Family members and caretakers were the patients' parents, spouses, or main caretakers who already knew the patients' HIV and MMT status and who were introduced by the patients. Only family members and caretakers who had patients' consent to approach and who already knew the patients' HIV and MMT status were invited for IDIs. The inclusion criteria for health care providers were having at least two years of counseling and clinical experience with people living with HIV and people who inject drugs.

IDI and FGD guides were developed in English in accordance with previous FB interviews and the objectives of this study. The interview guides were reviewed for linguistic and cultural appropriateness. The guides were translated into Vietnamese, discussed, and revised by the research team, who had knowledge and experience on the study topic and participants. The guides were piloted among 5 patients who met the study inclusion criteria but received treatment from another MMT clinic and did not participate in the study. The IDIs and FGDs were facilitated in Vietnamese by 2 interviewers who were trained in qualitative research methods and had extensive experience working with people living with HIV who inject drugs. Each IDI had an average duration of 50 (SD 18) minutes. The average FGD duration was 1.5 (SD 25) hours. All IDIs and FGDs were conducted in Vietnamese, audio recorded, and transcribed verbatim. The interview and FGD transcriptions were deidentified and translated into English. The analysis was conducted according to the applied thematic analysis approach [37]. Thematic analysis was performed using NVivo (version 12.0; QSR International). All qualitative research was guided by the consolidated criteria for reporting qualitative studies), a 32-item checklist [38]. In all, 4 study members with experience in qualitative data analysis were involved in the data analysis process. The list of parent and child codes was developed and agreed upon by the coders. Parent and child codes and definitions were created a priori and were derived from the interview guide and literature review. Furthermore, 20% of the transcripts were double coded. Intercoder reliability was assessed after all coders coded the first 2 transcripts. A total of 2 meetings were held among the coders to discuss the coding results item by item. The official coding process began after all 4 coders agreed on the coding method and procedure. The codebook was refined throughout the analysis process and the findings were presented, discussed, and agreed upon by the study team.

Phase 4: Production

In all, 3 external content experts reviewed the phase 3 results. The original FB manual study procedures and content were modified according to expert comments. On the basis of the content experts' review, the manual was tailored for use by both health care providers and lay counselors. In addition, the experts' feedback informed the study team to include relevant matters regarding injecting drugs and HIV transmission in Vietnam. Finally, all the case studies were revised to make them suitable for participants in Vietnam.

Phases 5 and 6: Topical Experts and Integration

We solicited and integrated input and feedback on the first draft of the adapted FB program from 2 investigators, 2 creators of FB, and 1 topical expert on behavioral programs for people living with HIV and SUD from Hanoi Medical University (HMU). This feedback informed the development of the next draft of the FB manual. Culturally, Vietnamese illustrations were designed to replace Zimbabwean illustrations and were added to the second manual draft in English. A total of 2 translators of the study team had experience with people living with HIV who inject drugs and mental health and were responsible for translating the second draft. The Vietnamese topical expert revised the translated draft to ensure conceptual equivalency. The team then revised the translation again and agreed on the second FB manual version.

Phases 7 and 8: Training and Testing

Using the second adapted version, the creators of FB led a Training of Trainers session for 6 study team members in Vietnam and 3 people from HMU via a 5-day virtual training course. Trainers in Vietnam then provided in-person 5-day training for lay counselors and health care counselors in MMT clinics. The second draft was then piloted with 5 eligible patients to further refine the session content. The process culminated in a final version of the adapted FB manual that would be ready for formal evaluation in an RCT study (Table 1).

Table 1. The Assessment-Decision-Adaptation-Production-Topical Experts-Integration-Training-Testing (ADAPT-ITT) model [25] applied in the study.

Phase	Method	Version
1: Assessment of new study population	Assessed needs of people living with HIV on MMT[a] for CMDs[b]	Completed in 2019
2: Decision on choosing EIP[c] for adaptation	Reviewed EIPs for CMDs then decided to choose FB[d] counseling CMDs for community people of Zimbabwe to adapt for counseling CMDs for people living with HIV on MMT	Completed in 2019
3: Adaptation	Conducted IDI[e] with 12 people living with HIV on MMT, 5 family members and caretakers, 4 clinic directors, and 2 FGDs[f] with MMT health care providers (each group had 5 people); analyzed results by themes and reported results	Original manual
4: Production	Presented results of phase 3 to 2 creators of FB and 2 investigators and 1 expert in HMU[g]; recorded and reported comments of 2 investigators and topical experts for production; reviewed original FB manual to modify per results above and received investigators' agreement on adaptation into the first draft manual	First draft
5: Topical experts	Sent the first draft manual to 3 topical experts and 2 investigators for comments and revision	First draft
6: Integration	Integrated experts' and investigators' comments to make the second draft, added Vietnamese illustrations to make the third draft in English, and sent the third draft to experts and investigators again for comments; translated the third draft to Vietnamese; sent the Vietnamese version of the third draft to a Vietnamese topical expert for revision and to ensure conceptual equivalency	Second draft
7: Training	Creators of FB trained for trainers in Vietnam; trainers trained for counselors	Third draft
8: Testing	Piloted the third draft for 5 people living with HIVs on MMT and having CMDs; revised third draft to make final manual	Final

[a]MMT: methadone maintenance treatment.

[b]CMD: common mental disorder.

[c]EIP: evidence-informed program.

[d]FB: Friendship Bench.

[e]IDI: in-depth interview.

[f]FGD: focus group discussion.

[g]HMU: Hanoi Medical University.

Ethics Approval

The study protocol is available at ClinicalTrials.gov (NCT04790201). The study protocol, interview guides, and informed consent forms were approved by the institutional review boards at University of North Carolina at Chapel Hill on August 24, 2020 (study 20-1689), and HMU on June 19, 2020 (decision 119 ĐHYHN). All the study participants provided written informed consent in Vietnamese.

Results

Phase 3: Adaptation

Overview of the Study Participants

We approached 67 patients using purposive sampling representing people living with HIV on MMT in 4 MMT clinics in Hanoi to answer the DASS-21 questions. Less than one-third of the patients approached (19/67, 28%) met the DASS-21 threshold CMD scores. Of 19 eligible patients, 12 (63%) agreed to be interviewed. A patient was female. The average age of the patients was 44 years. In all, 83% (10/12) of the patients had symptoms of depression, 83% (10/12) had symptoms of anxiety, and 33% (4/12) had symptoms of stress at screening (Table 2).

The age of family members and caretakers ranged from 60 to 73 years. All participants were either retired or unemployed. Health care providers included 2 MMT clinic directors, 1 associate director of CDC Hanoi, and 1 deputy director of the Hanoi Department of Health. A total of 10 health care workers, including MMT physicians and nurses in MMT clinics aged 35 to 66 years, participated in 2 FGDs.

The adaptation phase in our study aimed to explore the most culturally appropriate way to adapt FB in Vietnam. Using IDIs with patients, their family members or caretakers, clinic directors, and FGDs with health care providers, we addressed the who, where, when, and how regarding the administration of the FB program in the MMT clinics.

Table 2. Demographic characteristics of participants (N=12).

Demographic characteristics	Value
Gender, n (%)	
Male	11 (92)
Female	1 (8)
Age (years), mean (SD; range)	44 (6; 35-56)
Employment, n (%)	
Unemployed	6 (50)
Working full-time or part-time	6 (50)
Marital status, n (%)	
Single	5 (42)
Married	3 (25)
Divorced or separated	4 (33)
Years on ART[a], mean (SD)	8 (6)
Years on MMT[b], mean (SD)	5 (2)
DASS-21[c], n (%)	
Symptom of depression	10 (83)
Symptom of anxiety	10 (83)
Symptom of stress	4 (33)
Number of types of different CMD[d] symptoms, n (%)	
One symptom	3 (25)
Two symptoms	6 (50)[e]
Three symptoms	3 (25)

[a]ART: antiretroviral therapy.

[b]MMT: methadone maintenance treatment.

[c]DASS-21: Depression, Anxiety, and Stress Scale-21 items.

[d]CMD: common medical disorder.

[e]Overall, 5 had depression and anxiety 1 had anxiety and stress.

Who Should Deliver the Program

When being asked who (health staff person or peer) should provide the FB program for people living with HIV on MMT, the study respondents gave mixed reviews of who would be best for the FB program. Broadly, it appeared that directors and focus group respondents saw more advantages with a staff member as the program provider. Among family members, caretakers, and patient respondents, there was no distinct, overarching opinion on a preferred program provider. A total of 2 directors and 2 patient respondents shared that the professionalism of health care providers would be beneficial. A family member stressed that staff members should be purposively selected and need certain personality traits to be as effective as lay counselors, including honesty and willingness to share their experiences with patients in FB. Clinic providers noted that clinical experience with ART and MMT management are also helpful skills for a potential program provider. They shared that their mental health counseling experience may not be at the level needed for this type of program, as exemplified in focus group FGD_01:

> *It's certainly good. We haven't been trained in depression. We haven't had a chance to attend any depression courses. That's why we can't help our patients. We just told the patient's family to take the patient to the psychiatric clinic. Without experience, we can't support more than that. I have worked here for many years, but there has been no training course for depression.*

A director emphasized key aspects to improve and that staff members as program providers may need further training and understanding, enthusiasm, and gentleness:

> *Knowledge is always along with skills. They must be trained in skills and knowledge, when they understand, they can do it. Their attitude and their speech towards the patients must be from their hearts. If it's a real feeling, the patients will believe in the health workers. Patients are sensitive, sometimes only a glance could affect them.* [Director_IDI_D03]

For the program led by peers, participants from different groups (1 director, 4 FGD participants, 2 family members, and 7

patients) agreed that peers could understand patients very well. Peers and patients had many similarities in terms of life, health, and living conditions. Peers understood the patient's language and experienced similar things, both good and bad. In addition, formality was unnecessary between them, making it easier for patients to talk and share their opinions with peers, which could not be achieved with staff members:

> *I think it's feasible because they are the persons same as us, so they can understand the mindset of sick people like me, you are not sick so you cannot understand our thoughts.* [Patient_IDI_102]

Moreover, peers had a better understanding of patients' health conditions from their perspectives and experiences. All of these built more trust between peers and patients, facilitating the counseling sessions. However, according to clinic providers and directors, to successfully implement the FB program, peers had to have a serious attitude and commitment to the task. In doing so, they need training on how to provide counseling and care for patients to fulfill the tasks of a counselor. In addition to their practical experience, peers should update their knowledge about the topic and improve communication skills to effectively transfer what they have learned to the patients. Only 1 clinic health care provider and 1 patient were concerned that peers might not be able to fully understand their patients' conditions. Because of their addiction and medication and their mental state, it might be difficult to find enthusiastic and suitable peers to lead the program:

> *Because a peer can change differently day by day, it's difficult to find an appropriate methadone peer, but it is easier to find an ART peer. In my clinic, there is a peer who also counsels other patients...But we should think carefully about selecting methadone peers, we should consider medical staff doing the programme because we don't know if they [peers] are stable or not. We may have to change to others if we choose them.* [FGD_01]

In addition, although peers might receive training to do the job, they lack practical experience in caring for patients with different types of mental problems. A dominant theme among the patient respondents was the importance of the personality traits and counseling approach of the program provider. Patient respondents stressed the need for counselors who are gentle, empathetic, and understanding:

> *Someone like me needs affection, gentle speech, even a slightly unappropriated attitude will lead to my rejection. I need them to be considerate and gentle.* [Patient_IDI_303]

Considering the insights from interviews and FGDs in the adaptation phase, the study selected health care providers and people living with HIV or community members who were trusted by the study participants to work as study counselors in the RCT.

When to Deliver the Program

All clinic directors, health care providers, family members, and 9 participants (14 references from clinic directors, 15 references from FGDs, 10 references from family members, and 46 references from patients) reported their concerns about what time to implement the FB program. They all shared that as many patients work, it might be challenging to find a suitable time for them to attend the FB program. Most of the patients suggested that the counseling sessions should be in the morning when they come to take methadone:

> *Because we usually take the medicine in the morning and after that, we can do activities. In the afternoon, we can't wait, at that time we have to go home, so the most convenient time is the morning.* [Patient_IDI_301]

A patient shared that many patients could not sit for a long time, which would affect their participation in the counseling sessions. The patient believed that the counseling time should be arranged to suit the patient's employment and examination time, such as on weekends. For clinic directors and health care providers, apart from patients' employment, the timetable for clinic providers is very strictly scheduled; thus, time arrangement for the clinic providers to join the FB program is problematic. Providers thought that it is best to conduct FB during office hours to ensure convenience and safety for counselors and patients:

> *If we do it outside of office hours, we should consider the safety of people here. If something happens, we are also affected. So we should do counseling in the daytime, not nighttime because we can't control if they do something bad or not.* [FGD_01]

However, it is necessary to arrange counseling time in advance so that clinic staff do not have many overlapping tasks and have enough time for counseling patients:

> *So if we have such a programme then we have to have a better management of our time so that we can do it.* [FGD_01]

Where to Deliver the Program

All directors and most health care providers said that the place of program should be at the MMT clinic, as this is the place MMT patients visit every day. At MMT clinics, there were examination rooms; therefore, confidentiality and privacy would be ensured during the counseling sessions. Counseling outside the MMT clinic might be noticed by community members, which patients feared would cause them to be disliked. In addition, it is illegal for drug addicts to congregate in public spaces:

> *It's not possible to gather in the community. It is against the law, they are already drug addicts, if they gather that is against the law, and people in the community don't like that, so the methadone treatment facilities are the most suitable places. In the community, there are also a few places like that, but I only see their family come, I never see them [patients/people on treatment]. So, the medical facilities are still the best place.* [Director_IDI_D03]

Half of the patients (6/12, 50%) and almost all family members (4/5, 80%) preferred to receive counseling in MMT clinics for convenience and safety for both counselors and patients,

ensuring privacy for patients and having timely and relevant health services if needed. The other respondents expressed that they needed to have a private and confidential place for counseling. They also suggested that the program only needs to designate a random room in the clinic as the counseling room and that it would be great if the room could have materials or visual aides to refer to or use.

What and How to Deliver the Program

All 4 directors mentioned that the program should be facilitated by the medical leadership. According to the directors and health care providers, to have the FB program implemented, it is crucial that the program follows all the administrative procedures of the MMT clinics before starting it. It should have a clear and detailed work plan, in which who, how, where, and when to conduct each activity are specified. Importantly, it is necessary to know about the people living with HIV on MMT and the severity of their CMD symptoms to have appropriate approaches and program sessions for them:

> *If the programme is implemented in the clinic, in general, in terms of administrative procedures, I also said earlier that there must be a direction. As for the implementation, first, you must have purposes and goals, then it is necessary to train staff, then find the target group we need to consult, and then plan a schedule for specific activities, jobs.* [Director_IDI_01]

From the perspective of patients and their family members, though they believed that patients would benefit from a program such as FB, they thought the patients would participate in the program if they received compensation for their time and effort in participating in FB sessions (3/31, 10%) and did not have to pay or have any additional restrictions for program participation (2/31, 6%):

> *If only he participates without any condition from the programme, that's very good.* [Family_IDI_1011]

Participants thought that not having any additional challenges related to employment (10/31, 32%), financial issues (12/31, 39%), and transportation (2/31, 6%) would facilitate participation. Patients expected the program to be new and be organized interestingly and attractively:

> *I can't imagine it yet, but I want new and interesting methods. We will love it and if the programme have people like you, we will join.* [Patient_IDI_101]

Family members and patients discussed the following recommendations: the program materials should be easy to understand; fruits, cake, and drinks should be provided; patients who actively participated in the program activities and adhered to treatment should be rewarded and praised; and if possible, the program could assist patients with job opportunities or friend-making.

Both clinic providers and patients agreed that they preferred to delicately, rather than directly, discuss mental health to help patients feel comfortable. A health care provider suggested the use of words and language that is familiar to patients to help them feel more comfortable joining the program:

> *I think their language when talking together is very important, they want to release and use their language to feel comfortable but they can't do that when talking to medical staff.* [FGD_01]

A patient even suggested excluding the name of the program if possible. None of the suggestions for the language surrounding FB included "mental illness" in the program name.

From the perspective of family members, words in Vietnamese that are simple and easy to understand with positive meaning were preferable, and sensitive words that might worsen patients' mental state should be avoided. A health care provider and a patient thought that the name "Friendship Bench" was beautiful and acceptable, as it did not mention the problem and had a pleasant feeling:

> *The Friendship bench name is also very beautiful and good.* [FGD_F01]

Phase 4: Production

The results of the qualitative research in phase 3 were presented in a web-based meeting with 2 investigators, 2 creators of FB, and 1 expert from HMU to reach a consensus on the necessary adaptation of the FB manual. The original FB was adapted for use by both health care providers and lay counselors in Vietnam. Table 3 describes the details of FB adaptation. Overall, the manual illustrations of counselors were changed from Zimbabwean to Vietnamese and included both sexes. The illustrations of case studies have also been redrawn. The title "Lay Health Workers" was changed to "Counselors" to suit both groups of health care counselors and lay counselors for the study in Vietnam. The training curriculum was shortened from 8 to 5 days, and the new training program was written in a web-based format. The preface cited additional information about HIV infection and substance use patterns in Vietnam and provided a brief introduction to the adaptation of the FB manual to the study in Vietnam.

We retained the content of 8 out of the 23 chapters in the original manual. The 8 chapters included PST, counseling skills, FB cards, emotions, stabilization, strong emotional reactions, men's health-seeking behaviors, and self-care. We completely removed 6 of the 23 chapters, namely epilepsy, belief in supernatural powers, psychosis, home visits, group circle, and proverbs. We edited the contents in 8 of the 23 chapters. Key edits included using the words *Common Mental Disorders* to refer to a collective term of depression, anxiety, and stress-related disorders instead of "kufungisisa" (Zimbabwean for depression and anxiety), replacement of SSQ-14 to DASS-21, and suicide risk assessment and management procedures, which were used in the study in Vietnam. In addition, we excluded sessions on abuse of bronchodilators, alcohol and pregnancy, and information about antipsychotics, as they were not related to the study population. Information about the use of FB on tablets was removed, as the Vietnam FB program used paper forms to record counseling session information. Information on HIV transmission through injection, substance abuse disorders caused by injecting drugs, and amphetamine-type stimulants, which were common issues of people who inject drugs—the study population in the RCT—was added.

Table 3. Summary of changes in Friendship Bench (FB) manual.

Chapter	Name of chapter	Revision
N/A[a]	Cover page	• Changed the illustration from Zimbabwean to Vietnamese • Changed the word "Lay health workers" to "Counselors" to suit both groups of counselors of the study in Vietnam
N/A	Introduction	• Rewrote the preface to fit the purpose of the program in Vietnam
1	Psychoeducation	• Used the words "Common Mental Disorder" to refer to a collective term of depression, anxiety, and stress-related disorders instead of "kufungisisa" (Zimbabwean for depression and anxiety). • Removed information about tablets
2	Core competencies	• Changed chapter title from "Lay health workers" to "Counselors" • Replaced the Zimbabwean CMD[b] screening SSQ[c] with DASS-21[d] for research in Vietnam
3	Mental illness or mental neurological substance use disorders	• Removed information about postpartum depression, fatherhood, sexual partners, and family
4	Epilepsy	• Removed completely
5	Medication	• Kept information about ART[e] • Added information about MMT[f]
6	Belief in supernatural powers	• Removed completely
7	HIV and mental health	• Added information on HIV transmission through injection • Removed information about misunderstandings about church teachings
8	Substance use disorders	• Added information on substance use disorders caused by injecting drugs • Added information about amphetamine-type stimulants • Removed the substance use chemicals that cause bronchodilators • Removed alcohol and pregnancy
9	Psychosis	• Removed completely
10	Problem-solving therapy	• Retained
11	Questionnaires	• Replaced Screening CMD by SSQ in Zimbabwe by DASS-21 in study
12	Counseling skills	• Retained
13	FB cards	• Retained
14	Emotions	• Retained
15	Stabilization	• Retained
16	Strong emotional reactions	• Retained
17	Men's health-seeking behaviors	• Retained
18	Suicide assessment	• Replaced suicide assessment of SSQ by the suicide assessment of the study in Vietnam
19	Supervision	• Changed monitor titles to the study staffs in Vietnam
20	Home visits	• Removed completely
21	Group circle	• Removed completely
22	Self-care	• Retained
23	Proverbs	• Removed completely

Chapter	Name of chapter	Revision
N/A	Training overview	• Changed training curriculum from 8 days to 5 days. Rewrote new training program according to the virtual format
N/A	Shona training material	• Removed completely
N/A	Others	• Rewrote the illustrative examples taking information from the qualitative research of step 3 and making it specific to the study population and Vietnamese culture • Removed the instructions for using the tablet from the original document because the study in Vietnam used paper-based manual materials

[a]N/A: not applicable.

[b]CMD: common mental disorder.

[c]SSQ: Shona Symptoms Questionnaire.

[d]DASS-21: Depression, Anxiety, and Stress Scale-21 items.

[e]ART: antiretroviral therapy.

[f]MMT: methadone maintenance treatment.

All cases were revised, and the PST case was rewritten. In the original FB, the PST case featured a common problem in the Zimbabwe community, and the main character was a person with an unknown HIV status, a person with SUD, or a woman experiencing domestic violence. The new character in the adaptation was people living with HIV who inject drugs with short-term, solvable problems about employment, family matters, finances, and relationships with neighbors.

Phase 5: Topical Experts

The first English version was sent to 2 investigators, 2 creators of FB, and 1 expert from HMU for viewing and editing. All of them are associate professors with intensive expertise in the areas of mental health, HIV or AIDS, and substance use. They have led clinical studies and trials on a global scale. Experts advised correcting cases of PST to adjust them to Vietnamese culture. They recommended not changing the in-person counseling procedures and having up to 6 individual counseling sessions for each patient. An FB creator wrote scripts for the counselors to be included in the draft. The creator also advised maintaining the PST structure, with 4 main parts in each counseling session. The 4 main parts were the following: open your mind, make a list of problems, plan to solve the problem according to the Specific, Measurable, Achievable, Realistic, and Timely method, and encourage the clients to implement the plan. The 4 main parts of the PST were implemented into 7 small steps: (1) How does the client deal with problems? (2) How to recognize a problem? (3) How to select a problem, find the goal, and define the problem? (4) How to brainstorm for solutions? (5) How to select a solution? (6) How to make a Specific, Measurable, Achievable, Realistic, and Timely action plan? and (7) Did it work?

Phase 6: Integration

On October 20, 2021, a meeting was held on the web with 3 content experts, the study team in Vietnam, and 2 investigators. The investigators agreed on continuing to revise the case example of a typical people who inject drugs who encounters common mental problems in their life. These problems should be specific, simple, and solvable within 1 week. For example, “I want to quit using drugs for the next 1 year because next week is my son’s birthday.” The study team agreed to remove the group discussion portion, as it would be ineffective where MMT clinics are far from each other, without finding additional study resources. The study team then synthesized the expert revision to create the second draft in English. Illustrations of Vietnamese people were added to the second English version to make the third version which was sent to 3 content experts and 2 investigators to review and finalize. The third English version was translated into Vietnamese with the same content and images. A Vietnamese topical expert reviewed the translated third version to ensure conceptual equivalency.

Discussion

Principal Findings

This study successfully illustrates the process of operationalizing the ADAPT-ITT framework to adapt to a mental health program in Vietnam. The outcome of this study was an adapted FB manual to be used for training and piloting to create the final program manual.

We note that our study departs from the standard ADAPT-ITT framework by selecting an EIP that does not come from the CDC database of EIPs [26] and by adapting a program addressing CMDs rather than HIV prevention per the usual ADAPT-ITT goals [25]. Nevertheless, our study was methodologically aligned with the ADAPT-ITT goals, as we chose a proven, effective program for adaptation. FB improved CMDs among people living with HIV in Zimbabwe and Malawi, and we hope to use FB to address CMDs among people living with HIV on MMT in Vietnam.

The application of the ADAPT-ITT framework for the systematic adaptation of EIP can vary depending on the program, population, and resources [28,30,39]. A key strength of this study is that it used a linear approach to adapt the ADAPT-ITT framework on a small scale [28]. Results of the previous phases provided informative contributions for the next phase [40] in terms of the program procedures and content of the program manual. Other study and adaptation strengths were the translation into Vietnamese and the creation of local

illustrations to increase the clarity of the concepts while maintaining the original meaning of the English version.

The results from phase 3 informed the program procedures built upon the existing resources and administrative strength of MMT clinics [40]. As a result of adaptation, the study will choose health care providers and people living with HIV or community members who are trusted by the study participants to work as study counselors. The counselors will be trained and managed closely to ensure the quality of counseling. The counseling locations must be private and safe in MMT clinics. Counseling appointments will be scheduled between counselors and patients in parallel. Standardized criteria of counselors to deliver the program are defined as being understanding and having the trust of the patients but not necessarily having a high level of counseling techniques before training. Instead of using tools and procedures in the original FB program, the study will use questionnaires and protocols relevant to the study participants.

Regarding the content of the manual, the ADAPT-ITT framework guided the appropriate adaptation of the program while maintaining the core components of the PST of the original program throughout the counseling techniques in all program sessions [25,26]. The strength of our adaptation process was that the PST of the original FB fits the goals of reducing CMDs for people living with HIV on MMT in Vietnam; therefore, we did not have to change PST counseling as well as other chapters mentioning CMDs, HIV, and SUD in the original program that address common problems that global and Vietnamese people living with HIV and people who inject drugs encounter [41,42]. Feedback from topical experts informed the study team to alter some contents of the original program to ensure that it was suitable for Vietnamese people living with HIV on MMT and in MMT settings [27,40]. Specifically, unlike the original program conducted in Zimbabwe, we cut content that is not common in Vietnamese culture such as a church, beliefs in supernatural power, and praying together. The study also changed the illustrations from Zimbabwean to Vietnamese. Information from a qualitative study in phase 3 about common health problems faced by people living with HIV on MMT (the results are not reported within the scope of this paper) helped the study team rewrite sample PST cases to reflect typical problems in the key population. In addition, because of the new study with people living with HIV on MMT in Vietnam, we added additional guidelines on HIV, injecting drugs, and SUD to the original program to make it more culturally relevant and relevant to the study population. As recommended by topical experts, we dropped group activities that were proven more adaptive to the Zimbabwean setting and chapters less relevant to the target population (fatherhood, alcohol and pregnancy, epilepsy, and psychosis).

Limitations

This study has some limitations, the first being that the process of adaptation and production has not yet tested the feasibility and acceptability of the adapted manual. In addition, the semistructured interviews had not presented the original manual to the study participants. Therefore, we did not have comments from them on the content of the manual. Instead, the content of the manual was reviewed and revised in detail according to the study team's revisions and experts' comments. There might be challenges such as following the structure of FB counseling sessions, applying relevant counseling skills required in FB, following PST methods for health care, and lay counselors using the adapted manual as a resource material in conducting PST counseling sessions as they first work with this approach. As such, the adapted manual will be used to train counselors and piloted with the study participants of the future RCT to receive feedback to finalize the manual. Study participants and counselors can provide their feedback during the pilot phase. Future RCTs will test the acceptability and fidelity of the finalized adapted program.

Conclusions

This initial exploratory study demonstrates a successful process of following the ADAPT-ITT to adapt a proven mental health program for people living with HIV on MMT in Vietnam. This study selected and culturally adapted an evidence-informed PST program to improve CMDs among people living with HIV on MMT in Vietnam. The adaptation of the program through qualitative interviews and discussions with stakeholders and study participants and the use of feedback to tailor the program procedures allowed us to identify and preemptively address potential barriers to implementation. The production, integration, and expert input phases were used to tailor the manual to reflect typical manageable problems that people living with HIV on MMT encounter daily. If the adapted FB manual is acceptable and feasible, it may be used in MMT clinics in Vietnam to reduce CMDs for patients on MMT, which would contribute to the effectiveness of drug treatment and ART.

Acknowledgments

The authors are grateful for the support of the leaders of the Hanoi Department of Health and the Centers of Disease Control Hanoi and the 4 methadone maintenance treatment clinics in Hanoi for the implementation of this study. The authors would like to thank interviewers Nguyen Nhu Trang and Vu Thi Thanh Nhan who conducted the interviews for this study. Finally, the authors would like to thank the participants for their invaluable contribution to this research.

Authors' Contributions

BNG and BWP obtained funding and designed the study. HVT and HTTN oversaw training and data collection. HVT collected the data. TTTT, HTTN, TRF, and KRL coded and analyzed all data. KRL created the analysis plan, reviewed the coding, and reviewed memos to measure consistency and reliability. HVT drafted the manuscript. BNG, BWP, KRL, TRF, VFG, RV, HTH,

MXN, DC, and GML commented on, reviewed, and revised the manuscript for important intellectual content. All authors contributed to and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. A review of HIV prevention in Vietnam in 2019. Viet Nam Administration of HIV/AIDS Control, Joint United Nations Programme on HIV/AIDS. 2019 Jan. URL: https://unaids.org.vn/wp-content/uploads/2020/03/Prevention-Review-report_ENG_Final-14102019.pdf [accessed 2021-03-21]
2. HIV and AIDS Estimates - Country Factsheets – Viet Nam. Joint United Nations Programme on HIV/AIDS. 2020. URL: https://www.unaids.org/en/regionscountries/countries/vietnam [accessed 2021-10-13]
3. USAID/Vietnam Fact Sheet: HIV/AIDS. United States Agency for International Development. 2013 Aug 16. URL: https://www.usaid.gov/vietnam/documents/1861/fsvietnamhivaidsengpdf [accessed 2021-10-15]
4. Nguyen TT, Nguyen LT, Pham MD, Vu HH, Mulvey KP. Methadone maintenance therapy in Vietnam: an overview and scaling-up plan. Adv Prev Med 2012;2012:732484 [FREE Full text] [doi: 10.1155/2012/732484] [Medline: 23227351]
5. Colledge S, Larney S, Peacock A, Leung J, Hickman M, Grebely J, et al. Depression, post-traumatic stress disorder, suicidality and self-harm among people who inject drugs: a systematic review and meta-analysis. Drug Alcohol Depend 2020 Feb 01;207:107793. [doi: 10.1016/j.drugalcdep.2019.107793] [Medline: 31874449]
6. Remien RH, Stirratt MJ, Nguyen N, Robbins RN, Pala AN, Mellins CA. Mental health and HIV/AIDS: the need for an integrated response. AIDS 2019 Jul 15;33(9):1411-1420 [FREE Full text] [doi: 10.1097/QAD.0000000000002227] [Medline: 30950883]
7. Depression and Other Common Mental Disorders: Global Health Estimates. World Health Organization. Geneva, Switzerland: World Health Organization; 2017. URL: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf [accessed 2021-10-15]
8. Assessment and Management of Conditions Specifically Related to Stress: mhGAP Intervention Guide Module (version 1.0). World Health Organization and United Nations High Commissioner for Refugees. Geneva, Switzerland: World Health Organization; 2013. URL: https://apps.who.int/iris/bitstream/handle/10665/85623/9789241505932_eng.pdf [accessed 2021-06-14]
9. Jiao M, Gu J, Xu H, Hao C, Lau JT, Mo P, et al. Resilience associated with mental health problems among methadone maintenance treatment patients in Guangzhou, China. AIDS Care 2017 May;29(5):660-665 [FREE Full text] [doi: 10.1080/09540121.2016.1255705] [Medline: 27825278]
10. Levintow SN, Pence BW, Ha TV, Minh NL, Sripaipan T, Latkin CA, et al. Prevalence and predictors of depressive symptoms among HIV-positive men who inject drugs in Vietnam. PLoS One 2018 Jan 24;13(1):e0191548 [FREE Full text] [doi: 10.1371/journal.pone.0191548] [Medline: 29364928]
11. Mughal AY, Stockton MA, Bui Q, Go V, Pence BW, Ha TV, et al. Examining common mental health disorders in people living with HIV on methadone maintenance therapy in Hanoi, Vietnam. Harm Reduct J 2021 Apr 23;18(1):45 [FREE Full text] [doi: 10.1186/s12954-021-00495-3] [Medline: 33892743]
12. van Luenen S, Garnefski N, Spinhoven P, Spaan P, Dusseldorp E, Kraaij V. The benefits of psychosocial interventions for mental health in people living with HIV: a systematic review and meta-analysis. AIDS Behav 2018 Jan;22(1):9-42 [FREE Full text] [doi: 10.1007/s10461-017-1757-y] [Medline: 28361453]
13. Li L, Lee SJ, Jiraphongsa C, Khumtong S, Iamsirithaworn S, Thammawijaya P, et al. Improving the health and mental health of people living with HIV/AIDS: 12-month assessment of a behavioral intervention in Thailand. Am J Public Health 2010 Dec;100(12):2418-2425 [FREE Full text] [doi: 10.2105/AJPH.2009.185462] [Medline: 20966372]
14. Abas M, Nyamayaro P, Bere T, Saruchera E, Mothobi N, Simms V, et al. Feasibility and acceptability of a task-shifted intervention to enhance adherence to HIV medication and improve depression in people living with HIV in Zimbabwe, a low income country in Sub-Saharan Africa. AIDS Behav 2018 Jan;22(1):86-101 [FREE Full text] [doi: 10.1007/s10461-016-1659-4] [Medline: 28063075]
15. Stockton MA, Udedi M, Kulisewa K, Hosseinipour MC, Gaynes BN, Mphonda SM, et al. The impact of an integrated depression and HIV treatment program on mental health and HIV care outcomes among people newly initiating antiretroviral therapy in Malawi. PLoS One 2020 May 6;15(5):e0231872 [FREE Full text] [doi: 10.1371/journal.pone.0231872] [Medline: 32374724]
16. Li L, Tuan NA, Liang LJ, Lin C, Farmer SC, Flore M. Mental health and family relations among people who inject drugs and their family members in Vietnam. Int J Drug Policy 2013 Nov;24(6):545-549 [FREE Full text] [doi: 10.1016/j.drugpo.2013.06.007] [Medline: 23910167]
17. Le TA, Le MQ, Dang AD, Dang AK, Nguyen CT, Pham HQ, et al. Multi-level predictors of psychological problems among methadone maintenance treatment patients in difference types of settings in Vietnam. Subst Abuse Treat Prev Policy 2019 Sep 18;14(1):39 [FREE Full text] [doi: 10.1186/s13011-019-0223-4] [Medline: 31533764]

18. Wallén A, Eberhard S, Landgren K. The experiences of counsellors offering problem-solving therapy for common mental health issues at the Youth Friendship Bench in Zimbabwe. Issues Ment Health Nurs 2021 Sep;42(9):808-817. [doi: 10.1080/01612840.2021.1879977] [Medline: 33555957]
19. Chibanda D, Weiss HA, Verhey R, Simms V, Munjoma R, Rusakaniko S, et al. Effect of a primary care-based psychological intervention on symptoms of common mental disorders in Zimbabwe: a randomized clinical trial. JAMA 2016 Dec 27;316(24):2618-2626. [doi: 10.1001/jama.2016.19102] [Medline: 28027368]
20. Chinoda S, Mutsinze A, Simms V, Beji-Chauke R, Verhey R, Robinson J, et al. Effectiveness of a peer-led adolescent mental health intervention on HIV virological suppression and mental health in Zimbabwe: protocol of a cluster-randomised trial. Glob Ment Health (Camb) 2020 Aug 28;7:e23 [FREE Full text] [doi: 10.1017/gmh.2020.14] [Medline: 32963795]
21. Pence BW. Adaptation of the Friendship Bench Intervention for HIV-infected Perinatal Women in Lilongwe (Periscope). Chapel Hill, NC, USA: University of North Carolina at Chapel Hill; 2019.
22. Chibanda D, Bowers T, Verhey R, Rusakaniko S, Abas M, Weiss HA, et al. The Friendship Bench programme: a cluster randomised controlled trial of a brief psychological intervention for common mental disorders delivered by lay health workers in Zimbabwe. Int J Ment Health Syst 2015 May 23;9:21 [FREE Full text] [doi: 10.1186/s13033-015-0013-y] [Medline: 27408619]
23. Manh Than H, Minh Nong V, Trung Nguyen C, Phu Dong K, Ngo HT, Thu Doan T, et al. Mental health and health-related quality-of-life outcomes among frontline health workers during the peak of COVID-19 outbreak in Vietnam: a cross-sectional study. Risk Manag Healthc Policy 2020 Dec 8;13:2927-2936 [FREE Full text] [doi: 10.2147/RMHP.S280749] [Medline: 33324126]
24. Escoffery C, Lebow-Skelley E, Haardoerfer R, Boing E, Udelson H, Wood R, et al. A systematic review of adaptations of evidence-based public health interventions globally. Implement Sci 2018 Sep 26;13(1):125 [FREE Full text] [doi: 10.1186/s13012-018-0815-9] [Medline: 30257683]
25. Wingood GM, DiClemente RJ. The ADAPT-ITT model: a novel method of adapting evidence-based HIV interventions. J Acquir Immune Defic Syndr 2008 Mar 01;47 Suppl 1:S40-S46. [doi: 10.1097/QAI.0b013e3181605df1] [Medline: 18301133]
26. Sullivan PS, Stephenson R, Grazter B, Wingood G, Diclemente R, Allen S, et al. Adaptation of the African couples HIV testing and counseling model for men who have sex with men in the United States: an application of the ADAPT-ITT framework. Springerplus 2014 May 16;3(1):249 [FREE Full text] [doi: 10.1186/2193-1801-3-249] [Medline: 24877036]
27. Bere T, Nyamayaro P, Magidson JF, Chibanda D, Chingono A, Munjoma R, et al. Cultural adaptation of a cognitive-behavioural intervention to improve adherence to antiretroviral therapy among people living with HIV/AIDS in Zimbabwe: Nzira Itsva. J Health Psychol 2017 Sep;22(10):1265-1276 [FREE Full text] [doi: 10.1177/1359105315626783] [Medline: 26893295]
28. Davis T, DiClemente RJ, Prietula M. Using ADAPT-ITT to modify a telephone-based HIV prevention intervention for SMS delivery: formative study. JMIR Form Res 2020 Oct 06;4(10):e22485 [FREE Full text] [doi: 10.2196/22485] [Medline: 32831178]
29. Khumsaen N, Stephenson R. Adaptation of the HIV/AIDS self-management education program for men who have sex with men in Thailand: an application of the ADAPT-ITT framework. AIDS Educ Prev 2017 Oct;29(5):401-417 [FREE Full text] [doi: 10.1521/aeap.2017.29.5.401] [Medline: 29068714]
30. Libous JL, Montañez NA, Dow DE, Kapetanovic S, Buckley J, Kakhu TJ, et al. IMPAACT 2016: operationalizing HIV intervention adaptations to inform the science and outcomes of implementation. Front Reprod Health 2021 May 28;3:18. [doi: 10.3389/frph.2021.662912]
31. Gaynes B. Adaptation of the Friendship Bench counseling intervention to improve mental health and HIV care engagement outcomes among people living with HIV who inject drugs in Vietnam. University of North Carolina at Chapel Hill. Hanoi, Vietnam: National Institute of Health; 2020. URL: https://reporter.nih.gov/search/qae-BdjXdk6SfJst8IaAHw/project-details/10219229 [accessed 2021-02-11]
32. Lovibond SH, Lovibond PF. Manual for the Depression Anxiety Stress Scales. 2nd edition. Sydney, Australia: Psychology Foundation of Australia; 1995.
33. Le MT, Tran TD, Holton S, Nguyen HT, Wolfe R, Fisher J. Reliability, convergent validity and factor structure of the DASS-21 in a sample of Vietnamese adolescents. PLoS One 2017 Jul 19;12(7):e0180557 [FREE Full text] [doi: 10.1371/journal.pone.0180557] [Medline: 28723909]
34. Tran TD, Tran T, Fisher J. Validation of the depression anxiety stress scales (DASS) 21 as a screening instrument for depression and anxiety in a rural community-based cohort of northern Vietnamese women. BMC Psychiatry 2013 Jan 12;13:24 [FREE Full text] [doi: 10.1186/1471-244X-13-24] [Medline: 23311374]
35. Tran TT, Nguyen NB, Luong MA, Bui TH, Phan TD, Tran VO, et al. Stress, anxiety and depression in clinical nurses in Vietnam: a cross-sectional survey and cluster analysis. Int J Ment Health Syst 2019 Jan 3;13:3 [FREE Full text] [doi: 10.1186/s13033-018-0257-4] [Medline: 30622629]
36. Miller WC, Hoffman IF, Hanscom BS, Ha TV, Dumchev K, Djoerban Z, et al. A scalable, integrated intervention to engage people who inject drugs in HIV care and medication-assisted treatment (HPTN 074): a randomised, controlled phase 3 feasibility and efficacy study. Lancet 2018 Sep 01;392(10149):747-759 [FREE Full text] [doi: 10.1016/S0140-6736(18)31487-9] [Medline: 30191830]

37. Guest G, MacQueen KM, Namey EE. Applied Thematic Analysis. Thousand Oaks, CA, USA: Sage Publications; 2014.
38. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
39. Abubakari GM, Turner D, Nelson LE, Odhiambo AJ, Boakye F, Manu A, et al. An application of the ADAPT-ITT model to an evidence-based behavioral HIV prevention intervention for men who have sex with men in Ghana. Int Health Trends Persp 2021 Apr 04;1(1):1-16 [FREE Full text] [doi: 10.32920/ihtp.v1i1.1412]
40. Craig Rushing S, Gardner W. Native VOICES: adapting a video-based sexual health intervention for American Indian teens and young adults using the ADAPT-ITT model. Am Indian Alsk Native Ment Health Res 2016;23(1):24-46. [doi: 10.5820/aian.2301.2016.24] [Medline: 28562841]
41. HIV/AIDS and Mental Health. National Institute of Mental Health. URL: https://www.nimh.nih.gov/health/topics/hiv-aids/index.shtml [accessed 2020-06-16]
42. Ahmed T, Long NT, Huong PT, Stewart DE. HIV and injecting drug users in Vietnam: an overview of policies and responses. World Med Health Policy 2014 Dec;6(4):395-418. [doi: 10.1002/wmh3.122]

Abbreviations

ADAPT-ITT: Assessment-Decision-Adaptation-Production-Topical Experts-Integration-Training-Testing
ART: antiretroviral therapy
CDC: Centers for Disease Control and Prevention
CMD: common mental disorder
DASS-21: Depression, Anxiety, and Stress Scale-21 items
EIP: evidence-informed program
FB: Friendship Bench
FGD: focus group discussion
HMU: Hanoi Medical University
IDI: in-depth interview
MMT: methadone maintenance treatment
PST: problem-solving therapy
RCT: randomized controlled trial
SSQ-14: Shona Symptoms Questionnaire
SUD: substance use disorders

Edited by A Mavragani; submitted 16.02.22; peer-reviewed by R Ho, ASW Chan , A Doukani, S Miller; comments to author 30.05.22; revised version received 14.06.22; accepted 22.06.22; published 08.07.22.

Please cite as:
Tran HV, Nong HTT, Tran TTT, Filipowicz TR, Landrum KR, Pence BW, Le GM, Nguyen MX, Chibanda D, Verhey R, Go VF, Ho HT, Gaynes BN
Adaptation of a Problem-solving Program (Friendship Bench) to Treat Common Mental Disorders Among People Living With HIV and AIDS and on Methadone Maintenance Treatment in Vietnam: Formative Study
JMIR Form Res 2022;6(7):e37211
URL: https://formative.jmir.org/2022/7/e37211
doi:10.2196/37211
PMID:35802402

Original Paper

Preliminary Real-World Evidence Supporting the Efficacy of a Remote Neurofeedback System in Improving Mental Health: Retrospective Single-Group Pretest-Posttest Study

Jocelyne C Whitehead[1,2], PhD; Ron Neeman[2], MA; Glen M Doniger[2], PhD

[1]Integrated Program in Neuroscience, McGill University, Montreal, QC, Canada

[2]Myndlift Ltd, Tel Aviv, Israel

Corresponding Author:
Jocelyne C Whitehead, PhD
Integrated Program in Neuroscience
McGill University
845 Rue Sherbrooke O
Montreal, QC, H3A 0G4
Canada
Phone: 1 866 994 3210
Email: jocelyne.whitehead@mail.mcgill.ca

Abstract

Background: Neurofeedback training (NFT) has been shown to be effective in treating several disorders (eg, attention-deficit/hyperactivity disorder [ADHD], anxiety, and depression); however, little is currently known regarding the effectiveness of remote NFT systems.

Objective: This retrospective study provides real-world data (N=593) to assess the efficacy of app-based remote NFT in improving brain health and cognitive performance.

Methods: Improvement was measured from pre- to postintervention of in-app assessments that included validated symptom questionnaires (the 12-item General Health Questionnaire, the ADHD Rating Scale IV, the Adult ADHD Self-Report Scale, the 7-item Generalized Anxiety Disorder scale, and the 9-item Patient Health Questionnaire), a cognitive test of attention and executive functioning (ie, continuous performance task), and resting electroencephalography (EEG) markers. Clinically significant improvement was evaluated using standard approaches.

Results: The greatest improvement was reported for the anxiety questionnaire, for which 69% (68/99) of participants moved from abnormal to healthy score ranges. Overall, adult and child participants who engaged in neurofeedback to improve attention and executive functions demonstrated improved ADHD scores and enhanced performance on a cognitive (ie, response inhibition) task. Adults with ADHD additionally demonstrated elevated delta/alpha and theta/alpha ratios at baseline and a reduction in the delta/alpha ratio indicator following neurofeedback.

Conclusions: Preliminary findings suggest the efficacy of app-based remote neurofeedback in improving mental health, given the reduced symptom severity from pre- to postassessment for general psychological health, ADHD, anxiety, and depression, as well as adjusted resting EEG neural markers for individuals with symptoms of ADHD. Collectively, this supports the utility of the in-app assessment in monitoring behavioral and neural indices of mental health.

(JMIR Form Res 2022;6(7):e35636) doi:10.2196/35636

KEYWORDS

EEG biofeedback; remote care; neurofeedback; attention-deficit/hyperactivity disorder; delta/alpha ratio

Introduction

Background

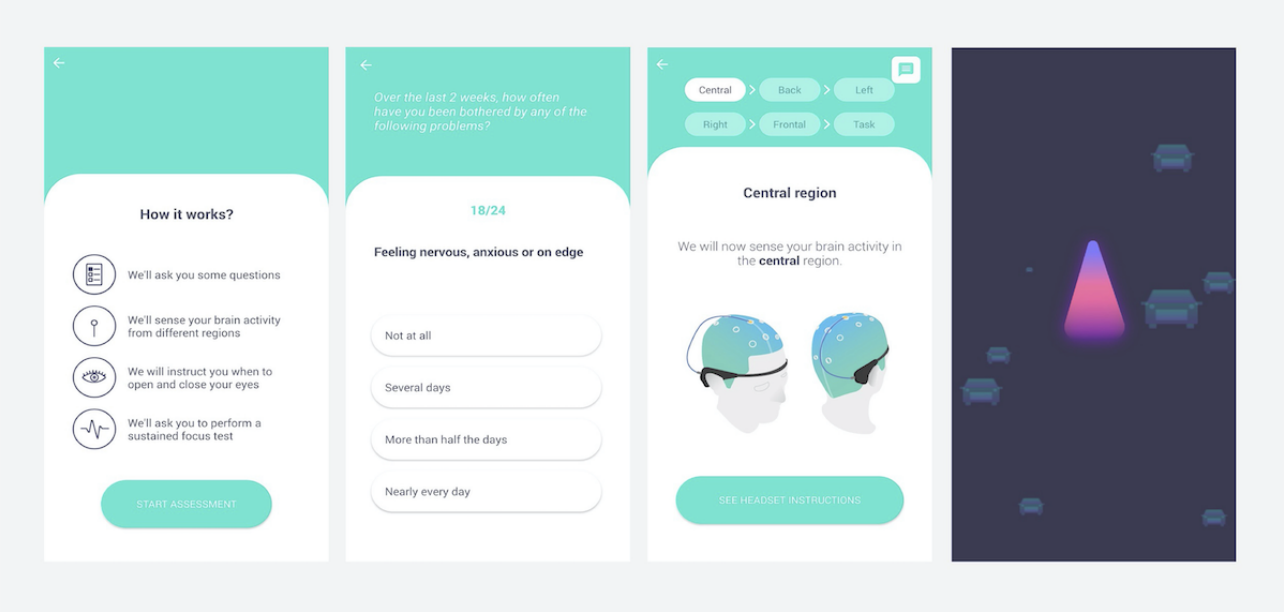
Neurofeedback training (NFT) is considered a primary or supplementary treatment for a number of disorders, including attention-deficit/hyperactivity disorder (ADHD) [1-5], anxiety [6-9], and depression [7,8,10]. The American Academy of Pediatrics [11] provided a “level 1 best support” rating of NFT as a safe and effective evidence-based therapy for childhood ADHD. Nonetheless, several significant barriers prevent patients from receiving quality neurofeedback therapeutics; for example,

electroencephalography (EEG) systems are expensive, complex, and often only accessible at health care clinics. A recent pilot study [12] provided encouraging evidence for the efficacy of therapist-guided NFT, suitable for remote home-based use. Findings showed improved ADHD symptomatology in a small cohort of children after 9 weeks of NFT. The system was designed as an affordable convenient wireless alternative to clinic-based EEG. NFT users regulate neural activity through operant conditioning, which can lead to morphological changes in the brain [13,14] and calmer, more focused cognitive, affective, and physical functioning. Currently, little is known about the effectiveness of NFT systems in the field [15]; therefore, this retrospective open-label pilot study offers real-world data supporting the efficacy of remote NFT in improving brain health.

Mental Health Improvement in Real-World Settings

Unlike standard EEG systems, Myndlift is an easy-to-use tool for patients and clinicians (Figure 1). While wearing the validated EEG headband (Muse; InteraXon [16,17]) containing four dry recording electrodes (ie, anterior frontal [AF] 7, AF8, temporal pole [TP] 9, and TP10), one ground electrode, and one auxiliary wet electrode, the patient trains with an Android or iOS app linked to the headset by Bluetooth, which delivers visual and auditory feedback during YouTube videos or specialized games. When patients' brain waves are in the desired range, positive feedback is delivered. A therapist can set or adjust the training protocol and monitor progress remotely via a cloud-based web service. The device incorporates an app-based assessment, lasting approximately 40 minutes, completed prior to NFT (ie, baseline) and periodically over the intervention period for longitudinal tracking of improvement.

Figure 1. Myndlift in-app assessment screens. From left to right: introduction, symptom questionnaires, resting electroencephalography assessment, and cognitive task.



Real-world studies provide external validity and accurately represent the heterogeneity of a patient population [18]. From the app, real-world data were collected from more than 500 participants on outcome measures, including pre- and postintervention assessments of validated symptom questionnaires, a cognitive test of attention and executive functioning (ie, continuous performance task [CPT]), and resting EEG markers. An efficacious system could serve as a reliable, cost-effective solution for users. In-clinic NFT costs approximately US $150 to $200 per session, with a minimum of 30 to 40 sessions typically recommended. In contrast to a cost-per-session model, remote NFT could offer monthly charges, ranging from US $200 to $500.

EEG Neuromarkers of ADHD

Given the success of neurofeedback for child ADHD, more adults with ADHD are turning to NFT for treatment. Currently, 6.76% of adults worldwide—translating to 366.3 million people—are affected [19]. ADHD is commonly recognized as a hypoaroused brain state [20]. In recent years, EEG measures have provided supporting evidence for popular theoretical models of hypoactivation [21] related to core symptoms of hyperactivity, inattention, and impulsivity [22]. The hypoarousal state is best localized to frontal and posterior regions [23] (ie, neuroanatomical structures subserving attentional networks [24,25]). EEG patterns of ADHD in children are characterized by elevated low-frequency power (ie, primarily theta) and reduced *relative* high-frequency power (ie, alpha and beta) [23,26-28], or an elevated ratio of the two (ie, low to high frequency). The theta/beta ratio (TBR) is the most common form of NFT in treating ADHD [29,30]; however, inconsistencies in the literature suggest that TBR [31,32] may not be reliable as a diagnostic measure [33]. This may reflect EEG heterogeneity across ADHD-diagnosed individuals (eg, Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition [DSM-5] subtypes; psychiatric comorbidities; age; and sex) [34,35]. For instance, although theta and beta power differences are evident in child ADHD [31,36], a recent review [37] suggested that theta and alpha frequencies may be more reliable markers for adults. Notably, most adult studies

emphasize group differences in alpha power during eyes-closed conditions [38-43], while more recent work has identified elevated theta and delta power in adults with ADHD [44,45]. Given this evidence, this study investigates whether TBR versus the delta/alpha ratio (DAR) or the theta/alpha ratio (TAR) are biomarkers for adult ADHD. Overall, the study evaluates evidence for improvement in mental health via symptom questionnaires, a CPT, and hypothesized EEG markers. Findings have implications for the benefits of NFT and efficacy of a remote home-based system.

Methods

Participants

Participants, 13 years of age or older, signed up through their clinician or a clinician suggested by Myndlift and completed NFT at home or in clinic in a clinical care context. Informed consent was provided through the app, allowing participants' anonymized data to be used for research. Data were included for analysis if baseline (ie, preintervention) assessment was conducted after 7 or fewer NFT sessions (ie, attributed to in-app NFT tutorial). For analyses of improvement, postintervention sessions occurred 30 to 180 days after baseline with 20 or more NFT sessions completed [46]. An average of 1 or more NFT sessions per week was required for inclusion, given that effective neurofeedback requires consistency [47,48], irrespective of the neurofeedback protocol used. Data were collected via the app.

Ethical Considerations

Procedures were reviewed by an independent Institutional Review Board (IRB)—Pearl IRB—who permitted IRB exemption for analyses of data previously collected and deidentified, following the guidelines of the Declaration of Helsinki.

Neurofeedback Protocol

Participants performed neurofeedback protocols (Multimedia Appendix 1) that were customized by their clinicians and consistent with current literature [49].

Procedure and Outcome Measures

Symptom Questionnaires

The in-app assessment includes 14 brief standardized questionnaires commonly used to screen for mental health conditions. In this study, data were reported for the following five questionnaires completed at baseline and follow-up by at least 25 participants: the 12-item General Health Questionnaire (GHQ-12) [50], the ADHD Rating Scale IV (ADHD-RS-IV) [51], the Adult ADHD Self-Report Scale (ASRS) for DSM-5 [52], the 7-item Generalized Anxiety Disorder scale (GAD-7) [53], and the 9-item Patient Health Questionnaire (PHQ-9) [54]. For each questionnaire, participants filled out self-report measures based on frequency of symptom occurrence using a 4- or 5-point Likert-style scale. Total scores were calculated for use in improvement analyses. Participants engaging in neurofeedback for ADHD completed the ASRS [52] if they were 18 years of age or older; otherwise, they completed the ADHD-RS-IV. The GHQ-12, GAD-7, and PHQ-9 were completed by participants of all ages [55-58].

Continuous Performance Task

The assessment contained an 8-minute CPT, a behavioral test of response inhibition, in which participants are instructed to tap the screen when the target object (ie, an arrow-like shape pointing upward) is shown, but not when other stimuli appear. The interstimulus interval and presence of audiovisual distracter stimuli were varied throughout the task. Outcomes included average response time (RT) and response time variability (ie, the SD of RT [SDRT]), as well as omission and commission errors related to inattention and impulsivity, respectively [59]. This type of test is commonly used as an objective measure of attention and executive function [60-62] and has become a standard assessment tool for attentional difficulties [59,63,64].

Resting EEG

Resting EEG was recorded from 9 electrodes (ie, AF7, AF8, TP9, TP10, central [C] zero [z], frontal [F] z, F3, F4, and occipital [O] 1). The EEG assessment was divided into five sequential (ie, "sensing") phases; in each phase, the auxiliary electrode was placed at a different scalp location: central (Cz), frontal (Fz), left (F3), right (F4), and back or posterior (O1). Each phase was split into eyes-closed and eyes-open blocks. A block continued until 30 seconds of clean EEG—sampled at 256 Hz—had been recorded, which typically took up to 45 seconds.

Statistical Analysis

Symptom Questionnaires

Questionnaire results were analyzed in terms of improvement in total score from pre- to postintervention, including mean change in points, effect size (ie, Cohen *d*), and percent of users with clinically significant improvement, defined as 20% improvement [65,66]. Results are presented separately for participants scoring in healthy and abnormal ranges at baseline, as per conventional clinical cutoff values. The percent of participants who shifted from abnormal to normal (ie, healthy) ranges after the intervention is also reported. Paired-samples *t* tests (2-tailed) evaluated statistically significant improvement for each clinical measure ($P<.05$). By convention, small, medium, and large effects correspond to $d=0.2$, $d=0.5$, and $d=0.8$, respectively. For symptom questionnaires, CPT, and resting EEG analyses, multiple comparisons were corrected using the Benjamini-Hochberg (BH) method [67] to maintain a family-wise error at $P=.05$, reported as BH-adjusted *P* values (P_{BH}). The Levene test assessed assumptions of equality of variance and corrected for inhomogeneities.

Continuous Performance Task

CPT results were analyzed for participants who completed child (ie, ADHD-RS-IV) or adult (ie, ASRS) ADHD questionnaires. Results are given in terms of improvement in RT and SDRT for correct responses (ie, shorter and less variable response times, respectively), commission errors, and omission errors. This includes mean change, in milliseconds or errors, and effect size. RT and SDRT scores were standardized by age to minimize age effects on performance [68]. Percent of participants demonstrating clinically significant improvement was reported, defined by a reliable change index (RCI) [69] that accounts for

practice effects [70]. Exceeding a critical value of 95% for a 1-tailed test—equivalent to 1.65 SD units on a standardized *z* scale—indicates a significant reliable change, similar to others [71].

Resting EEG

Participants who completed the adult ADHD questionnaire at baseline were split into groups with "healthy" and "abnormal" ranges of values based on their score. Only participants with clean EEG signals were included (see Multimedia Appendix 2 for EEG preprocessing). Results were reported in terms of EEG amplitude (ie, Hz; relative power) for TAR, DAR, and TBR at baseline. Independent-samples *t* tests were conducted for each power ratio across groups (ie, healthy and abnormal values). Frequency bands were defined as follows: delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (13-30 Hz). These were averaged across frontal electrodes (ie, F3 and F4, based on the frontal nodes of the frontoparietal network [25,72] and the prevalence of a clean EEG signal) during the eyes-closed condition. Improvement analyses were conducted separately for each group and included the mean change in ratio amplitude from pre- to postintervention and associated effect size; paired-samples *t* tests were used to evaluate within-group changes.

Results

Sample Characteristics

Data from 560 participants met the criteria for inclusion in the analysis. Depending on clinical considerations determined by their therapist, subsets of participants completed each symptom questionnaire, CPT, resting EEG, or any combination of the three. Table 1 gives sample characteristics for each assessment component, including the NFT protocols completed by 50% or more of each sample population (Multimedia Appendix 1).

Table 1. Sample characteristics as separated by each outcome measure and analysis.

Measure	Age (years)		Gender, n (%)		Test setting, n (%)		NFT[a] protocols used in ≥50% of sample	No. of sessions, mean (SD)	Treatment duration (days), mean (SD)	Frequency (sessions/wk), mean (SD)
	Mean (SD)	Range	Female	Male	Clinic	Home				
Symptom questionnaire pre-post (n=301[b])	38 (14.5)	13-71	157 (52.7)	141 (47.3)	220 (73.1)	81 (26.9)	Reduce theta; reduce high beta; enhance low beta	53 (38.2)	91 (41.0)	4 (2.6)
CPT[c] pre-post (ADHD[d]; n=203[e])	37 (12.9)	13-69	103 (51.2)	98 (48.8)	112 (55.2)	91 (44.8)	Reduce theta; reduce high beta	53 (29.5)	96 (42.2)	4 (1.9)
Resting EEG[f] baseline (adult ADHD; n=271[g])	38 (10.9)	18-70	94 (35.2)	173 (64.8)	87 (32.1)	184 (67.9)	N/A[h]	N/A	N/A	N/A
Resting EEG pre-post (adult ADHD; n=41[i])	36 (9.3)	19-55	17 (42.5)	23 (57.5)	5 (12.2)	36 (87.8)	Reduce theta; enhance alpha; reduce high beta	55 (30.8)	76 (27.0)	5 (2.2)

[a]NFT: neurofeedback training.

[b]No gender identity was reported by 3 participants (n=298).

[c]CPT: continuous performance task.

[d]ADHD: attention-deficit/hyperactivity disorder.

[e]No gender identity was reported by 2 participants (n=201).

[f]EEG: electroencephalography.

[g]No gender identity was reported by 4 participants (n=267).

[h]N/A: not applicable; intervention details were not reported, as only preintervention values were of interest for baseline analyses.

[i]No gender identity was reported by 1 participant (n=40).

Symptom Questionnaires

Results for participants who completed symptom questionnaires (n=301) were separated into groups with abnormal and healthy scores (Table 2). Most participants engaged in NFT protocols to reduce theta (227/301, 75.4%) and enhance high beta (248/301, 82.4%), while many who completed the PHQ-9 (76/134, 56.7%) and the ASRS (59/112, 52.7%) also performed enhanced alpha, whereas children who completed the ADHD-RS-IV also often included enhanced low beta (21/27, 78%) and enhanced sensorimotor rhythm (SMR; 16/27, 59%). In the groups with abnormal results, all questionnaires had large effect sizes (d=0.99 to 2.41), while the effect sizes for groups with healthy results were large only for child and adult ADHD questionnaires. Improvement in the groups with abnormal results was statistically significant for all questionnaires, with the majority (30/56, 54% to 7/7, 100%) of users demonstrating clinically significant change (ie, ≥20%) [65,66]. The most prominent improvement was observed in participants with abnormal baseline anxiety or child ADHD scores. Nevertheless,

ADHD-RS-IV findings are considered preliminary given the small sample size. Most participants (30/56, 54% to 7/7, 100%) in the groups with abnormal results shifted their values to healthy ranges at postintervention. Improvement of healthy participants was statistically significant for all questionnaires, with the majority (30/66, 45% to 14/20, 70%) demonstrating clinically significant change.

Table 2. Improvement in self-reported subjective symptoms after ≥30 days of Myndlift neurofeedback for users that scored in the healthy range, and separately for those that scored in the abnormal range (per conventional clinical cutoffs) at baseline.

Questionnaire and group at baseline (cutoff value)	No. of sessions, mean (SD)	Treatment duration (days), mean (SD)	Change (points decreased), mean (SD)	Change *T* value	Change *P* value[a]	Effect size, *d*	Users improved by ≥20%, n (%)	Abnormal to healthy results, n (%)
12-item General Health Questionnaire (maximum score = 36)								
Abnormal (≥12; n=197)	53 (39.1)	94 (42.2)	7.8 (7.80)	13.94	<.001	0.99	139 (71)	113 (57)
Healthy (<12; n=66)	52 (34.3)	84 (36.6)	1.0 (4.28)	1.90	.06	0.23	30 (45)	N/A[b]
ADHD[c] Rating Scale IV (for children; maximum score = 54): preliminary								
Abnormal (>36; n=7)	49 (19.7)	75 (32.8)	19.3 (7.99)	6.38	<.001	2.41	7 (100)	7 (100)
Healthy (≤36; n=20)	53 (23.2)	102 (37.9)	7.9 (8.10)	4.36	<.001	0.98	14 (70)	N/A
Adult ADHD Self-Report Scale (maximum score =24)								
Abnormal (≥14; n=56)	48 (25.6)	86 (41.5)	4.0 (3.81)	7.83	<.001	1.05	30 (54)	30 (54)
Healthy (<14; n=56)	63 (35.7)	97 (37.2)	2.1 (2.14)	7.38	<.001	0.99	33 (59)	N/A
7-item Generalized Anxiety Disorder scale (maximum score = 21)								
Abnormal (≥14; n=99)	52 (36.7)	87 (40.2)	6.4 (5.18)	12.39	<.001	1.24	82 (83)	68 (69)
Healthy (<14; n=107)	55 (32.5)	97 (40.4)	1.3 (3.92)	3.43	.001	0.33	63 (59)	N/A
9-item Patient Health Questionnaire (max imum score = 27)								
Abnormal (≥10; n=63)	57 (47.6)	88 (37.7)	6.2 (5.47)	8.94	<.001	1.13	45 (71)	38 (60)
Healthy (<10; n=71)	57 (34.7)	95 (39.8)	1.5 (4.07)	3.04	.004	0.36	49 (69)	N/A

[a]Reported as Benjamini-Hochberg–adjusted *P* values.

[b]N/A: not applicable; healthy subjects are already within the healthy range.

[c]ADHD: attention-deficit/hyperactivity disorder.

Continuous Performance Task

Participants completing CPT and ADHD questionnaires performed primarily reduced theta (76/99, 77%) and enhanced high beta (81/99, 90%) protocols. Most adults also performed enhanced alpha (54/90, 60%), whereas most children also performed enhanced low beta (9/9, 100%) and enhanced SMR (7/9, 78%). Results (n=99) for average RT, SDRT, omission errors, and commission errors were divided by abnormal versus healthy scores for child and adult ADHD combined (Table 3). The greatest improvement observed, irrespective of group (ie, abnormal and healthy ADHD ranges), was in SDRT (*d*=1.02 and *d*=1.24, respectively), where nearly half of the participants (42/99, 43%) demonstrated clinically significant improvement, as indicated by the RCI. Although average RTs improved comparably (42/99, 43%), differences between pre- and postintervention were significant only for the healthy results group (*d*=0.56). At least one-third of users improved in their commission errors (35/99, 35%) and omission errors (45/99, 45%) from pre- to postintervention. Results from a group (n=104) with unknown ADHD assignment were comparable to those of groups with abnormal and healthy results (Multimedia Appendix 3).

Table 3. Improvement in CPT after ≥30 days of Myndlift neurofeedback (n=99) separately for healthy users that scored in the normal range for children or adults at baseline and for those in the abnormal ADHD range (per conventional clinical cutoffs).

CPT[a] outcome and group results at baseline ASRS[b] or ADHD[c]-RS-IV[d]	No. of sessions, mean (SD)	Treatment duration (days), mean (SD)	Change reduction, mean (SD)[e]	Change *T* value	Change *P* value[f]	Effect size, *d*	Users improved (RCI[g] ≥1.65 SD), n (%)
Average response time							
Abnormal (n=46)	48 (26.3)	85 (40.3)	8.9 (33.52)	1.80	.08	0.27	20 (43)
Healthy (n=53)	61 (31.0)	100 (40.0)	15.0 (26.96)	4.05	<.001	0.56	22 (42)
Response time variability (SD of response time)							
Abnormal (n=46)	48 (26.3)	85 (40.3)	10.3 (10.02)	6.95	<.001	1.02	18 (39)
Healthy (n=53)	61 (31.0)	100 (40.0)	10.7 (8.64)	8.99	<.001	1.24	25 (47)
Commission errors (impulsivity)							
Abnormal (n=46)	48 (26.3)	85 (40.3)	4.0 (7.23)	3.75	<.001	0.55	19 (41)
Healthy (n=53)	61 (31.0)	100 (40.0)	2.0 (3.17)	4.51	<.001	0.62	16 (30)
Omission errors (inattention)							
Abnormal (n=46)	48 (26.3)	85 (40.3)	1.5 (3.16)	3.27	.003	0.48	24 (52)
Healthy (n=53)	61 (31.0)	100 (40.0)	0.64 (1.88)	2.48	.02	0.34	21 (40)

[a]CPT: continuous performance task.

[b]ASRS: Adult ADHD Self-Report Scale.

[c]ADHD: attention-deficit/hyperactivity disorder.

[d]ADHD-RS-IV: ADHD Rating Scale IV.

[e]Reported in milliseconds for response time average and variability, and in number of errors for commission and omission errors.

[f]Reported as Benjamini-Hochberg–adjusted *P* values.

[g]RCI: reliable change index.

EEG Indicators of Adult ADHD

Resting EEG Baseline

The DAR, TAR, and TBR were calculated from baseline resting EEG data (n=271) in frontal regions (ie, average of F3 and F4) with eyes closed from participants scoring in abnormal (n=125) or healthy ranges (n=146) on the adult ADHD questionnaire.

Regarding the DAR, an independent-samples *t* test demonstrated that participants in the abnormal results group (mean 1.10, SD 0.61) had significantly greater frontal DAR than healthy participants (mean 0.90, SD 0.48; t_{235}=3.02, P_{BH}=.009, *d*=0.37). The Levene test indicated unequal variances (*F*=5.25, *P*=.02), so degrees of freedom were adjusted from 269 to 235. Post hoc independent-samples *t* tests confirmed that results were driven by participants in the abnormal results group having significantly greater frontal delta power (t_{269}=2.80, P_{BH}=.01, *d*=0.34) and less frontal alpha power (t_{269}=2.61, P_{BH}=.01, *d*=0.34) than healthy participants.

Regarding the TAR, a comparable *t* test reported a significant difference for the frontal TAR (t_{269}=2.46, P_{BH}=.02, *d*=0.30), as participants with abnormal scores (mean 0.64, SD 0.30) had significantly greater ratios than those with healthy scores (mean 0.56, SD 0.26). Post hoc *t* tests confirmed that results were driven by less frontal alpha, as opposed to differences in theta (t_{269}=1.11, P_{BH}=.27, *d*=0.13).

Regarding the TBR, a final *t* test reported no significant difference between participants with abnormal scores (mean 0.66, SD 0.27) and those with healthy scores (mean 0.64, SD 0.31; t_{269}=0.532, P_{BH}=.60, *d*=0.06).

Preliminary Resting EEG Improvement

Changes in the DAR, TAR, and TBR in the frontal regions with eyes closed were reported for participants (n=41) scoring in the abnormal (n=20) or healthy ranges (n=21) of the adult ADHD questionnaire (Table 4). Most participants completed reduced theta (32/41, 78%), enhanced high beta (37/41, 90%), and enhanced alpha protocols (27/41, 66%). After correcting for multiple comparisons, significant improvement was only reported for the DAR in the abnormal results group.

Table 4. Change in resting EEG ratios from frontal (ie, average F3 and F4) electrodes during the eyes-closed condition after ≥30 days of Myndlift neurofeedback (n=41) for healthy users and separately for those that scored in the abnormal adult ADHD range (per conventional clinical cutoffs) at baseline.

EEG[a] pre-post outcome and group at baseline (cutoff value)	No. of sessions, mean (SD)	Treatment duration (days), mean (SD)	Change reduction (Hz), mean (SD)	Change *T* value	Change *P* value[b]	Effect size, *d*
Delta/alpha ratio						
Abnormal (≥14; n=20)	49 (22.1)	77 (26.8)	0.20 (0.284)	3.15	.03	0.70
Healthy (<14; n=21)	61 (36.9)	76 (27.8)	0.08 (0.450)	0.79	.59	0.18
Theta/alpha ratio						
Abnormal (≥14; n=20)	49 (22.1)	77 (26.8)	0.04 (0.171)	1.00	.66	0.22
Healthy (<14; n=21)	61 (36.9)	76 (27.8)	0.01 (0.227)	0.35	.66	0.08
Theta/beta ratio						
Abnormal (≥14; n=20)	49 (22.1)	77 (26.8)	0.04 (0.144)	1.34	.79	0.30
Healthy (<14; n=21)	61 (36.9)	76 (27.8)	0.02 (0.218)	0.44	.73	0.10

[a]EEG: electroencephalography.

[b]Reported as Benjamini-Hochberg–adjusted *P* values.

Discussion

Principal Findings

This retrospective study offers initial evidence of therapist-guided remote neurofeedback as an effective tool for reducing subjective symptoms, improving objective cognitive performance, and adaptively modifying EEG markers. Improvements in attention were evident in children and adults with ADHD, as well as healthy participants. Findings suggest that the TBR is not a reliable marker for adult ADHD, instead demonstrating alternative elevated slow/fast power ratios [37]. Moreover, we provide preliminary evidence for improvement (ie, reduced DAR) in adults with ADHD. These findings offer a promising use for remote NFT as a low-cost alternative to clinic-based EEG.

Efficacy for Improving Mental Health Remotely

Based on real-world data, significant improvement was reported across standardized questionnaires. The greatest improvement was observed in participants with abnormal anxiety scores, where most received reduced theta, enhanced high beta, and enhanced alpha protocols. As anticipated, greater effect sizes were observed for participants with scores in the abnormal versus healthy ranges. Interestingly, healthy participants and those with ADHD, both children and adults, demonstrated significant improvement with large effect sizes after completing primarily reduced theta and enhanced high beta protocols, as well as adults who completed enhanced alpha protocols or children who completed reduced low beta and reduced SMR protocols. Consistent with the literature [33-35], our findings suggest that children and adults may benefit from unique NFT protocols to improve ADHD symptoms, although a larger sample is required to confirm preliminary ADHD-RS-IV results.

Apart from the child ADHD assessment, questionnaire analyses included large total sample numbers (ie, 112 to 263 participants), and after an average of 53 NFT sessions, 57% to 78% of the participants demonstrated significant improvement, depending on the questionnaire. Results were particularly impressive compared to other in-app mental health therapeutics [73-76], such as mobile-enabled text psychotherapy [77] or app-based cognitive behavioral therapy [78]. The majority (61%) of participants scoring in the abnormal ranges moved to the healthy results group over an average of approximately 3 months, a time frame costing less than US $1500 with Myndlift versus US $6000 to $8000 for traditional neurofeedback.

Improved Cognitive Performance for Healthy Participants and Those With ADHD

NFT led to greater consistency in response times on a response inhibition task for subjects scoring in healthy or abnormal ADHD ranges, agreeing with similar reports of subjects with ADHD [79,80]. In addition, the RCI demonstrated that approximately 50% of healthy participants improved their average response time, while similarly, participants in abnormal ranges reduced omission errors. Importantly, CPT findings agree with improved ADHD questionnaire scores, suggesting that NFT provides objective evidence of improved executive function, the primary cognitive domain impacted by attentional difficulties.

Identifying Adult ADHD Neuromarkers

Resting EEG findings demonstrated that elevated DAR and TAR were indicative of adult ADHD at baseline. This translated to significantly higher levels of delta and lower levels of alpha, as previously reported in adults with ADHD [39-42,81,82]. Notably, Liechti and colleagues [35] reported high theta to be less consistent in adults than in children, and that ADHD versus healthy control classification improved having exploratorily included delta waves in the discriminant analysis. Adults with ADHD may present slower theta waves—bordering fast delta waves—than children, although further analysis is required. Together, findings are consistent with the cortical hypoarousal theory, where low-power fast oscillations accompany reduced self-control and executive functioning [83], and high-power slow oscillations are reported with decreased subcortical motivational drive [84]. Preliminary evidence for reduced DAR

in adult ADHD from pre- to postassessment may reflect the improved ADHD symptoms and CPT measures, particularly given the success of protocols inhibiting slow oscillations and enhancing fast oscillation [1], and the high percentage of ADHD participants performing reduced theta (ie, slow) and enhanced alpha (ie, fast) protocols.

In contrast to our work and that of others, several groups reported high alpha power at baseline during eyes-closed conditions in adult ADHD populations [85,86], or rather, no difference across ADHD participants and healthy controls [87,88]. Importantly, variability across the adult ADHD literature may, in part, be due to the heterogeneity of ADHD [34,35] and differences in study designs, sample sizes, analyses, and EEG technology [89]. For example, Loo and colleagues [38] demonstrated that adults with ADHD combined-type (ie, symptoms of inattention and hyperactivity or impulsivity) present reduced alpha power globally, compared to ADHD inattentive-type or non-ADHD controls.

Limitations and Future Directions

Study results are encouraging, but conclusions should be tempered by limitations, including small subgroup sample sizes and lack of control groups. Moreover, subjects may have received alternative treatment in parallel (eg, medication) that could influence symptom improvement as well as alter neuromarkers. For example, two studies administering stimulants (ie, methylphenidate or dexamphetamine) to treat symptoms of ADHD in adults demonstrated altered delta [90,91] and theta waves [90] posttreatment. No changes in alpha or beta waves were reported. Given the evidence in this study for altered delta and alpha waves in adults with abnormal ADHD scores, we would hypothesize that the mechanism of action for stimulants versus NFT may differ, *resulting in influence over varied frequency bands*. Moreover, as this population reflects real-world use, the likelihood of these two forms of treatment to have commenced simultaneously, for treating symptoms of depression, anxiety, and ADHD, would arguably be low. Those seeking treatment with remote neurofeedback most often do so to avoid taking pharmaceuticals [92,93] or, rather, to supplement their current treatment, which alone may not be sufficiently effective [94]. Frank H Duffy [95], a Harvard professor and pediatric neurologist, suggests that “if any medication had demonstrated such a wide spectrum of efficacy it would be universally accepted and widely used.” Further, controlled research studies will be required to facilitate comparison of neurofeedback efficacy with other interventions. Notwithstanding these limitations, the findings are essential as they reflect real-world benefits of remote neurofeedback to actual patients. Follow-up analyses will compare benefits across NFT protocols and will further evaluate the impact on resting EEG outcomes.

Conclusions

Preliminary findings from this retrospective pilot study demonstrate efficacy of remote NFT in improving mental health, particularly for individuals with symptoms of ADHD and anxiety, mainly through reduced theta, enhanced high beta, and enhanced alpha NFT protocols. Moreover, adult ADHD was distinguished from healthy individuals by elevated frontal DARs, where ratios were significantly reduced following NFT. The effectiveness of the system in a real-world population via remote use positions it as an affordable and accessible alternative to clinic-based systems.

Acknowledgments

Funding for this study was provided by McGill University, Montreal, Canada, and Myndlift Ltd. JCW received funding from McGill University for a doctoral internship.

Authors' Contributions

JCW conducted the analyses, interpreted the results, cowrote the paper, and revised the final version. RN co-designed the study, extracted the data, interpreted the results, cowrote the paper, and revised the final version. GMD designed the study, cowrote the paper, and revised the final draft. All authors reviewed and revised the results and approved the final version of the paper.

Conflicts of Interest

JCW is a freelance consultant for Myndlift. RN and GMD are employees of Myndlift.

Multimedia Appendix 1
Neurofeedback protocols implemented in this study.
[DOCX File , 18 KB - formative_v6i7e35636_app1.docx]

Multimedia Appendix 2
Electroencephalography preprocessing.
[DOCX File , 13 KB - formative_v6i7e35636_app2.docx]

Multimedia Appendix 3
Continuous performance task improvement for users without an ADHD assessment. ADHD: attention-deficit/hyperactivity disorder.

[DOCX File , 14 KB - formative_v6i7e35636_app3.docx]

References

1. Arns M, de Ridder S, Strehl U, Breteler M, Coenen A. Efficacy of neurofeedback treatment in ADHD: The effects on inattention, impulsivity and hyperactivity: A meta-analysis. Clin EEG Neurosci 2009 Jul;40(3):180-189. [doi: 10.1177/155005940904000311] [Medline: 19715181]
2. Micoulaud-Franchi J, Geoffroy PA, Fond G, Lopez R, Bioulac S, Philip P. EEG neurofeedback treatments in children with ADHD: An updated meta-analysis of randomized controlled trials. Front Hum Neurosci 2014;8:906 [FREE Full text] [doi: 10.3389/fnhum.2014.00906] [Medline: 25431555]
3. Bussalb A, Congedo M, Barthélemy Q, Ojeda D, Acquaviva E, Delorme R, et al. Clinical and experimental factors influencing the efficacy of neurofeedback in ADHD: A meta-analysis. Front Psychiatry 2019;10:35 [FREE Full text] [doi: 10.3389/fpsyt.2019.00035] [Medline: 30833909]
4. Riesco-Matías P, Yela-Bernabé JR, Crego A, Sánchez-Zaballos E. What do meta-analyses have to say about the efficacy of neurofeedback applied to children with ADHD? Review of previous meta-analyses and a new meta-analysis. J Atten Disord 2021 Feb;25(4):473-485. [doi: 10.1177/1087054718821731] [Medline: 30646779]
5. Lambez B, Harwood-Gross A, Golumbic EZ, Rassovsky Y. Non-pharmacological interventions for cognitive difficulties in ADHD: A systematic review and meta-analysis. J Psychiatr Res 2020 Jan;120:40-55. [doi: 10.1016/j.jpsychires.2019.10.007] [Medline: 31629998]
6. Moore NC. A review of EEG biofeedback treatment of anxiety disorders. Clin Electroencephalogr 2000 Jan;31(1):1-6. [doi: 10.1177/155005940003100105] [Medline: 10638346]
7. Hammond DC. Neurofeedback treatment of depression and anxiety. J Adult Dev 2005 Aug;12(2-3):131-137. [doi: 10.1007/s10804-005-7029-5]
8. Hammond DC. Neurofeedback with anxiety and affective disorders. Child Adolesc Psychiatr Clin N Am 2005 Jan;14(1):105-123, vii. [doi: 10.1016/j.chc.2004.07.008] [Medline: 15564054]
9. Tolin D, Davies C, Moskow D, Hofmann S. Biofeedback and neurofeedback for anxiety disorders: A quantitative and qualitative systematic review. Adv Exp Med Biol 2020;1191:265-289. [doi: 10.1007/978-981-32-9705-0_16] [Medline: 32002934]
10. Micoulaud-Franchi J, McGonigal A, Lopez R, Daudet C, Kotwas I, Bartolomei F. Electroencephalographic neurofeedback: Level of evidence in mental and brain disorders and suggestions for good clinical practice. Neurophysiol Clin 2015 Dec;45(6):423-433. [doi: 10.1016/j.neucli.2015.10.077] [Medline: 26553293]
11. American Academy of Pediatrics. Appendix S2: Evidence-based child and adolescent psychosocial interventions. Pediatrics 2010;125:S128 [FREE Full text] [doi: 10.1542/peds.2010-0788h]
12. Doniger GM, Kaddan A. Pilot study of the efficacy of mobile neurofeedback for attention-deficit/hyperactivity disorder (ADHD). In: Proceedings of the 27th ISNR Conference: Keynotes, Plenary Sessions, and Poster and Presentations. 2019 Presented at: The 27th ISNR Conference; September 19-22, 2019; Denver, CO p. 220 URL: https://www.neuroregulation.org/article/view/19901/13353
13. Marins T, Rodrigues E, Bortolini T, Melo B, Moll J, Tovar-Moll F. Structural and functional connectivity changes in response to short-term neurofeedback training with motor imagery. Neuroimage 2019 Jul 01;194:283-290. [doi: 10.1016/j.neuroimage.2019.03.027] [Medline: 30898654]
14. Ghaziri J, Tucholka A, Larue V, Blanchette-Sylvestre M, Reyburn G, Gilbert G, et al. Neurofeedback training induces changes in white and gray matter. Clin EEG Neurosci 2013 Oct;44(4):265-272. [doi: 10.1177/1550059413476031] [Medline: 23536382]
15. Antle AN, Chesick L, Sridharan SK, Cramer E. East meets west: A mobile brain-computer system that helps children living in poverty learn to self-regulate. Pers Ubiquitous Comput 2018 Jun 12;22(4):839-866. [doi: 10.1007/s00779-018-1166-x]
16. Krigolson OE, Hammerstrom MR, Abimbola W, Trska R, Wright BW, Hecker KG, et al. Using Muse: Rapid mobile assessment of brain performance. Front Neurosci 2021;15:634147 [FREE Full text] [doi: 10.3389/fnins.2021.634147] [Medline: 33584194]
17. Krigolson OE, Williams CC, Norton A, Hassall CD, Colino FL. Choosing MUSE: Validation of a low-cost, portable EEG system for ERP research. Front Neurosci 2017;11:109 [FREE Full text] [doi: 10.3389/fnins.2017.00109] [Medline: 28344546]
18. Blonde L, Khunti K, Harris SB, Meizinger C, Skolnik NS. Interpretation and impact of real-world clinical data for the practicing clinician. Adv Ther 2018 Nov;35(11):1763-1774 [FREE Full text] [doi: 10.1007/s12325-018-0805-y] [Medline: 30357570]
19. Song P, Zha M, Yang Q, Zhang Y, Li X, Rudan I. The prevalence of adult attention-deficit hyperactivity disorder: A global systematic review and meta-analysis. J Glob Health 2021 Feb 11;11:04009 [FREE Full text] [doi: 10.7189/jogh.11.04009] [Medline: 33692893]
20. Satterfield JH, Dawson ME. Electrodermal correlates of hyperactivity in children. Psychophysiology 1971 Mar;8(2):191-197. [doi: 10.1111/j.1469-8986.1971.tb00450.x] [Medline: 5089415]
21. Martella D, Aldunate N, Fuentes LJ, Sánchez-Pérez N. Arousal and executive alterations in attention deficit hyperactivity disorder (ADHD). Front Psychol 2020;11:1991 [FREE Full text] [doi: 10.3389/fpsyg.2020.01991] [Medline: 32903419]

22. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5®). Washington, DC: American Psychiatric Association Publishing; 2013.
23. Clarke AR, Barry RJ, McCarthy R, Selikowitz M. Children with attention-deficit/hyperactivity disorder and comorbid oppositional defiant disorder: An EEG analysis. Psychiatry Res 2002 Aug;111(2-3):181-190. [doi: 10.1016/s0165-1781(02)00137-3]
24. Posner MI. Imaging attention networks. Neuroimage 2012 Jun;61(2):450-456 [FREE Full text] [doi: 10.1016/j.neuroimage.2011.12.040] [Medline: 22227132]
25. Menon V. Large-scale brain networks and psychopathology: A unifying triple network model. Trends Cogn Sci 2011 Oct;15(10):483-506. [doi: 10.1016/j.tics.2011.08.003] [Medline: 21908230]
26. Kuperman S, Johnson B, Arndt S, Lindgren S, Wolraich M. Quantitative EEG differences in a nonclinical sample of children with ADHD and undifferentiated ADD. J Am Acad Child Adolesc Psychiatry 1996 Aug;35(8):1009-1017. [doi: 10.1097/00004583-199608000-00011] [Medline: 8755797]
27. Clarke AR, Barry RJ, McCarthy R, Selikowitz M. Age and sex effects in the EEG: Differences in two subtypes of attention-deficit/hyperactivity disorder. Clin Neurophysiol 2001 May;112(5):815-826. [doi: 10.1016/s1388-2457(01)00487-4]
28. Markovska-Simoska S, Pop-Jordanova N. Quantitative EEG in children and adults with attention deficit hyperactivity disorder: Comparison of absolute and relative power spectra and theta/beta ratio. Clin EEG Neurosci 2017 Jan;48(1):20-32. [doi: 10.1177/1550059416643824] [Medline: 27170672]
29. Monastra VJ. Electroencephalographic biofeedback (neurotherapy) as a treatment for attention deficit hyperactivity disorder: Rationale and empirical foundation. Child Adolesc Psychiatr Clin N Am 2005 Jan;14(1):55-82, vi. [doi: 10.1016/j.chc.2004.07.004] [Medline: 15564052]
30. Arns M, Heinrich H, Strehl U. Evaluation of neurofeedback in ADHD: The long and winding road. Biol Psychol 2014 Jan;95:108-115. [doi: 10.1016/j.biopsycho.2013.11.013] [Medline: 24321363]
31. Snyder SM, Hall JR. A meta-analysis of quantitative EEG power associated with attention-deficit hyperactivity disorder. J Clin Neurophysiol 2006 Oct;23(5):440-455. [doi: 10.1097/01.wnp.0000221363.12503.78] [Medline: 17016156]
32. Snyder SM, Quintana H, Sexson SB, Knott P, Haque A, Reynolds DA. Blinded, multi-center validation of EEG and rating scales in identifying ADHD within a clinical sample. Psychiatry Res 2008 Jun 30;159(3):346-358. [doi: 10.1016/j.psychres.2007.05.006] [Medline: 18423617]
33. Arns M, Conners CK, Kraemer HC. A decade of EEG theta/beta ratio research in ADHD: A meta-analysis. J Atten Disord 2013 Jul;17(5):374-383. [doi: 10.1177/1087054712460087] [Medline: 23086616]
34. Loo SK, Makeig S. Clinical utility of EEG in attention-deficit/hyperactivity disorder: A research update. Neurotherapeutics 2012 Jul;9(3):569-587 [FREE Full text] [doi: 10.1007/s13311-012-0131-z] [Medline: 22814935]
35. Liechti MD, Valko L, Müller UC, Döhnert M, Drechsler R, Steinhausen H, et al. Diagnostic value of resting electroencephalogram in attention-deficit/hyperactivity disorder across the lifespan. Brain Topogr 2013 Jan;26(1):135-151. [doi: 10.1007/s10548-012-0258-6] [Medline: 23053601]
36. Barry RJ, Clarke AR, Johnstone SJ. A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. Clin Neurophysiol 2003 Feb;114(2):171-183. [doi: 10.1016/s1388-2457(02)00362-0]
37. Adamou M, Fullen T, Jones SL. EEG for diagnosis of adult ADHD: A systematic review with narrative analysis. Front Psychiatry 2020;11:871 [FREE Full text] [doi: 10.3389/fpsyt.2020.00871] [Medline: 33192633]
38. Loo SK, Hale TS, Hanada G, Macion J, Shrestha A, McGough JJ, et al. Familial clustering and DRD4 effects on electroencephalogram measures in multiplex families with attention deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 2010 Apr;49(4):368-377. [doi: 10.1016/j.jaac.2010.01.002]
39. Loo SK, Hale TS, Macion J, Hanada G, McGough JJ, McCracken JT, et al. Cortical activity patterns in ADHD during arousal, activation and sustained attention. Neuropsychologia 2009 Aug;47(10):2114-2119 [FREE Full text] [doi: 10.1016/j.neuropsychologia.2009.04.013] [Medline: 19393254]
40. Ponomarev VA, Mueller A, Candrian G, Grin-Yatsenko VA, Kropotov JD. Group independent component analysis (gICA) and current source density (CSD) in the study of EEG in ADHD adults. Clin Neurophysiol 2014 Jan;125(1):83-97. [doi: 10.1016/j.clinph.2013.06.015] [Medline: 23871197]
41. Woltering S, Jung J, Liu Z, Tannock R. Resting state EEG oscillatory power differences in ADHD college students and their peers. Behav Brain Funct 2012 Dec 18;8:60 [FREE Full text] [doi: 10.1186/1744-9081-8-60] [Medline: 23249444]
42. Deiber M, Hasler R, Colin J, Dayer A, Aubry J, Baggio S, et al. Linking alpha oscillations, attention and inhibitory control in adult ADHD with EEG neurofeedback. Neuroimage Clin 2020;25:102145 [FREE Full text] [doi: 10.1016/j.nicl.2019.102145] [Medline: 31911342]
43. Missonnier P, Hasler R, Perroud N, Herrmann F, Millet P, Richiardi J, et al. EEG anomalies in adult ADHD subjects performing a working memory task. Neuroscience 2013 Jun 25;241:135-146. [doi: 10.1016/j.neuroscience.2013.03.011] [Medline: 23518223]
44. Kiiski H, Bennett M, Rueda-Delgado LM, Farina FR, Knight R, Boyle R, et al. EEG spectral power, but not theta/beta ratio, is a neuromarker for adult ADHD. Eur J Neurosci 2020 May;51(10):2095-2109. [doi: 10.1111/ejn.14645] [Medline: 31834950]

45. Kaur S, Singh S, Arun P, Kaur D. Analysis of resting state EEG signals of adults with attention-deficit hyperactivity disorder. In: Chaki R, Cortesia A, Saeed K, Chaki N, editors. Advanced Computing and Systems for Security. Advances in Intelligent Systems and Computing, vol 897. In: Springer, Singapore; 2019:61-72.
46. Neurofeedback Collaborative Group. Double-blind placebo-controlled randomized clinical trial of neurofeedback for attention-deficit/hyperactivity disorder with 13-month follow-up. J Am Acad Child Adolesc Psychiatry 2021 Jul;60(7):841-855. [doi: 10.1016/j.jaac.2020.07.906] [Medline: 32853703]
47. Vernon D, Egner T, Cooper N, Compton T, Neilands C, Sheri A, et al. The effect of training distinct neurofeedback protocols on aspects of cognitive performance. Int J Psychophysiol 2003 Jan;47(1):75-85. [doi: 10.1016/s0167-8760(02)00091-0]
48. Kropotov JD. Functional Neuromarkers for Psychiatry: Applications for Diagnosis and Treatment. London, UK: Academic Press; 2016.
49. Marzbani H, Marateb HR, Mansourian M. Neurofeedback: A comprehensive review on system design, methodology and clinical applications. Basic Clin Neurosci 2016 Apr;7(2):143-158 [FREE Full text] [doi: 10.15412/J.BCN.03070208] [Medline: 27303609]
50. Goldberg DP, Gater R, Sartorius N, Ustun TB, Piccinelli M, Gureje O, et al. The validity of two versions of the GHQ in the WHO study of mental illness in general health care. Psychol Med 1997 Jan;27(1):191-197. [doi: 10.1017/s0033291796004242] [Medline: 9122299]
51. DuPaul GJ, Power TJ, Anastopoulos AD, Reid R. ADHD Rating Scale—IV: Checklists, Norms, and Clinical Interpretation. New York, NY: The Guilford Press; 1998.
52. Ustun B, Adler LA, Rudin C, Faraone SV, Spencer TJ, Berglund P, et al. The World Health Organization Adult Attention-Deficit/Hyperactivity Disorder Self-Report Screening Scale for DSM-5. JAMA Psychiatry 2017 May 01;74(5):520-527 [FREE Full text] [doi: 10.1001/jamapsychiatry.2017.0298] [Medline: 28384801]
53. 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]
54. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
55. Goldberg DPW, Paul DPM. A User's Guide to the General Health Questionnaire. Windsor, UK: NFER-NELSON; 1988.
56. Mossman SA, Luft MJ, Schroeder HK, Varney ST, Fleck DE, Barzman DH, et al. The Generalized Anxiety Disorder 7-item scale in adolescents with generalized anxiety disorder: Signal detection and validation. Ann Clin Psychiatry 2017 Nov;29(4):227-234A [FREE Full text] [Medline: 29069107]
57. Bardhoshi G, Cobb N, Erford BT. Determining evidence-based outcomes in school-aged youth: Free-access instruments for school counselor use. Prof Sch Couns 2019 Mar 07;22(1b):1-10 [FREE Full text] [doi: 10.1177/2156759x19834431]
58. Richardson LP, McCauley E, Grossman DC, McCarty CA, Richards J, Russo JE, et al. Evaluation of the Patient Health Questionnaire-9 item for detecting major depression among adolescents. Pediatrics 2010 Dec;126(6):1117-1123 [FREE Full text] [doi: 10.1542/peds.2010-0852] [Medline: 21041282]
59. Riccio CA, Cohen MJ, Hynd GW, Keith RW. Validity of the Auditory Continuous Performance Test in differentiating central processing auditory disorders with and without ADHD. J Learn Disabil 1996 Sep;29(5):561-566. [doi: 10.1177/002221949602900510] [Medline: 8870525]
60. Greenberg LM, Waldman ID. Developmental normative data on the test of variables of attention (T.O.V.A.). J Child Psychol Psychiatry 1993 Sep;34(6):1019-1030. [doi: 10.1111/j.1469-7610.1993.tb01105.x] [Medline: 8408366]
61. Cassuto H, Ben-Simon A, Berger I. Using environmental distractors in the diagnosis of ADHD. Front Hum Neurosci 2013;7:805 [FREE Full text] [doi: 10.3389/fnhum.2013.00805] [Medline: 24319423]
62. Conners CK. Conners, 3rd Edition (Conners 3). Los Angeles, CA: Western Psychological Services; 2008. URL: https://www.wpspublish.com/conners-3-conners-third-edition [accessed 2022-06-30]
63. Edwards MC, Gardner ES, Chelonis JJ, Schulz EG, Flake RA, Diaz PF. Estimates of the validity and utility of the Conners' Continuous Performance Test in the assessment of inattentive and/or hyperactive-impulsive behaviors in children. J Abnorm Child Psychol 2007 Jun;35(3):393-404. [doi: 10.1007/s10802-007-9098-3] [Medline: 17295064]
64. Shaked D, Faulkner LMD, Tolle K, Wendell CR, Waldstein SR, Spencer RJ. Reliability and validity of the Conners' Continuous Performance Test. Appl Neuropsychol Adult 2020;27(5):478-487. [doi: 10.1080/23279095.2019.1570199] [Medline: 30793982]
65. Kounali D, Button KS, Lewis G, Gilbody S, Kessler D, Araya R, et al. How much change is enough? Evidence from a longitudinal study on depression in UK primary care. Psychol Med 2020 Nov 03:1-8. [doi: 10.1017/s0033291720003700]
66. McGuire JF, Geller DA, Murphy TK, Small BJ, Unger A, Wilhelm S, et al. Defining treatment outcomes in pediatric obsessive-compulsive disorder using a self-report scale. Behav Ther 2019 Mar;50(2):314-324 [FREE Full text] [doi: 10.1016/j.beth.2018.06.003] [Medline: 30824248]
67. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 2018 Dec 05;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
68. Brocki KC, Tillman CM, Bohlin G. CPT performance, motor activity, and continuous relations to ADHD symptom domains: A developmental study. Eur J Dev Psychol 2010 Mar;7(2):178-197. [doi: 10.1080/17405620801937764]

69. Jacobson NS, Truax P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991;59(1):12-19. [doi: 10.1037/0022-006x.59.1.12]
70. Parsons TD, Notebaert AJ, Shields EW, Guskiewicz KM. Application of reliable change indices to computerized neuropsychological measures of concussion. Int J Neurosci 2009;119(4):492-507. [doi: 10.1080/00207450802330876] [Medline: 19229718]
71. Perreau-Linck E, Lessard N, Lévesque J, Beauregard M. Effects of neurofeedback training on inhibitory capacities in ADHD children: A single-blind, randomized, placebo-controlled study. J Neurother 2010 Aug 16;14(3):229-242. [doi: 10.1080/10874208.2010.501514]
72. Rojas GM, Alvarez C, Montoya CE, de la Iglesia-Vayá M, Cisternas JE, Gálvez M. Study of resting-state functional connectivity networks using EEG electrodes position as seed. Front Neurosci 2018;12:235 [FREE Full text] [doi: 10.3389/fnins.2018.00235] [Medline: 29740268]
73. 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]
74. 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]
75. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: A meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
76. Wu A, Scult MA, Barnes ED, Betancourt JA, Falk A, Gunning FM. Smartphone apps for depression and anxiety: A systematic review and meta-analysis of techniques to increase engagement. NPJ Digit Med 2021 Feb 11;4(1):20 [FREE Full text] [doi: 10.1038/s41746-021-00386-8] [Medline: 33574573]
77. Hull TD, Mahan K. A study of asynchronous mobile-enabled SMS text psychotherapy. Telemed J E Health 2017 Mar;23(3):240-247. [doi: 10.1089/tmj.2016.0114] [Medline: 27797646]
78. Venkatesan A, Rahimi L, Kaur M, Mosunic C. Digital cognitive behavior therapy intervention for depression and anxiety: Retrospective study. JMIR Ment Health 2020 Aug 26;7(8):e21304 [FREE Full text] [doi: 10.2196/21304] [Medline: 32845246]
79. Ryoo M, Son C. Effects of neurofeedback training on EEG, continuous performance task (CPT), and ADHD symptoms in ADHD-prone college students. J Korean Acad Nurs 2015 Dec;45(6):928-938. [doi: 10.4040/jkan.2015.45.6.928] [Medline: 26805505]
80. Rossiter T. The effectiveness of neurofeedback and stimulant drugs in treating AD/HD: Part II. Replication. Appl Psychophysiol Biofeedback 2004 Dec;29(4):233-243. [doi: 10.1007/s10484-004-0383-4] [Medline: 15707253]
81. Kitsune GL, Cheung CHM, Brandeis D, Banaschewski T, Asherson P, McLoughlin G, et al. A matter of time: The influence of recording context on EEG spectral power in adolescents and young adults with ADHD. Brain Topogr 2015 Jul;28(4):580-590 [FREE Full text] [doi: 10.1007/s10548-014-0395-1] [Medline: 25200165]
82. Bresnahan SM, Barry RJ. Specificity of quantitative EEG analysis in adults with attention deficit hyperactivity disorder. Psychiatry Res 2002 Oct;112(2):133-144. [doi: 10.1016/s0165-1781(02)00190-7]
83. Rowe DL, Robinson PA, Lazzaro IL, Powles RC, Gordon E, Williams LM. Biophysical modeling of tonic cortical electrical activity in attention deficit hyperactivity disorder. Int J Neurosci 2005 Sep;115(9):1273-1305. [doi: 10.1080/00207450590934499] [Medline: 16048806]
84. Schutter DJ, Leitner C, Kenemans JL, van Honk J. Electrophysiological correlates of cortico-subcortical interaction: A cross-frequency spectral EEG analysis. Clin Neurophysiol 2006 Feb;117(2):381-387. [doi: 10.1016/j.clinph.2005.09.021] [Medline: 16371254]
85. Koehler S, Lauer P, Schreppel T, Jacob C, Heine M, Boreatti-Hümmer A, et al. Increased EEG power density in alpha and theta bands in adult ADHD patients. J Neural Transm (Vienna) 2009 Jan;116(1):97-104. [doi: 10.1007/s00702-008-0157-x] [Medline: 19030776]
86. Poil S, Bollmann S, Ghisleni C, O'Gorman RL, Klaver P, Ball J, et al. Age dependent electroencephalographic changes in attention-deficit/hyperactivity disorder (ADHD). Clin Neurophysiol 2014 Aug;125(8):1626-1638. [doi: 10.1016/j.clinph.2013.12.118] [Medline: 24582383]
87. Hermens DF, Williams LM, Lazzaro I, Whitmont S, Melkonian D, Gordon E. Sex differences in adult ADHD: A double dissociation in brain activity and autonomic arousal. Biol Psychol 2004 Jul;66(3):221-233. [doi: 10.1016/j.biopsycho.2003.10.006] [Medline: 15099695]
88. van Dongen-Boomsma M, Lansbergen MM, Bekker EM, Kooij JJS, van der Molen M, Kenemans JL, et al. Relation between resting EEG to cognitive performance and clinical symptoms in adults with attention-deficit/hyperactivity disorder. Neurosci Lett 2010 Jan 18;469(1):102-106. [doi: 10.1016/j.neulet.2009.11.053] [Medline: 19945506]
89. Saad JF, Kohn MR, Clarke S, Lagopoulos J, Hermens DF. Is the theta/beta EEG marker for ADHD inherently flawed? J Atten Disord 2018 Jul;22(9):815-826. [doi: 10.1177/1087054715578270] [Medline: 25823742]

90. Bresnahan SM, Barry RJ, Clarke AR, Johnstone SJ. Quantitative EEG analysis in dexamphetamine-responsive adults with attention-deficit/hyperactivity disorder. Psychiatry Res 2006 Feb 28;141(2):151-159. [doi: 10.1016/j.psychres.2005.09.002] [Medline: 16343642]
91. Cooper RE, Skirrow C, Tye C, McLoughlin G, Rijsdijk F, Banaschweski T, et al. The effect of methylphenidate on very low frequency electroencephalography oscillations in adult ADHD. Brain Cogn 2014 Apr;86:82-89. [doi: 10.1016/j.bandc.2014.02.001] [Medline: 24594658]
92. Berger I, Dor T, Nevo Y, Goldzweig G. Attitudes toward attention-deficit hyperactivity disorder (ADHD) treatment: Parents' and children's perspectives. J Child Neurol 2008 Sep;23(9):1036-1042. [doi: 10.1177/0883073808317726] [Medline: 18487521]
93. McHugh RK, Whitton SW, Peckham AD, Welge JA, Otto MW. Patient preference for psychological vs pharmacologic treatment of psychiatric disorders. J Clin Psychiatry 2013 Jun 15;74(06):595-602. [doi: 10.4088/jcp.12r07757]
94. Razoki B. Neurofeedback versus psychostimulants in the treatment of children and adolescents with attention-deficit/hyperactivity disorder: A systematic review. Neuropsychiatr Dis Treat 2018 Oct;Volume 14:2905-2913. [doi: 10.2147/ndt.s178839]
95. Duffy FH. The state of EEG biofeedback therapy (EEG operant conditioning) in 2000: An editor's opinion. Clin Electroencephalogr 2000 Jan;31(1):V-VII. [Medline: 10638345]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
ADHD-RS-IV: ADHD Rating Scale IV
AF: anterior frontal
ASRS: Adult ADHD Self-Report Scale
BH: Benjamini-Hochberg
C: central
CPT: continuous performance task
DAR: delta/alpha ratio
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
EEG: electroencephalography
F: frontal
GAD-7: 7-item Generalized Anxiety Disorder scale
GHQ-12: 12-item General Health Questionnaire
IRB: Institutional Review Board
NFT: neurofeedback training
O: occipital
P_{BH}: Benjamini-Hochberg–adjusted *P* value
PHQ-9: 9-item Patient Health Questionnaire
RCI: reliable change index
RT: response time
SDRT: SD of response time, response time variability
SMR: sensorimotor rhythm
TAR: theta/alpha ratio
TBR: theta/beta ratio
TP: temporal pole
z: zero

Edited by A Mavragani; submitted 13.12.21; peer-reviewed by DLG Rodrigues; comments to author 20.05.22; revised version received 31.05.22; accepted 09.06.22; published 08.07.22.

Please cite as:
Whitehead JC, Neeman R, Doniger GM
Preliminary Real-World Evidence Supporting the Efficacy of a Remote Neurofeedback System in Improving Mental Health: Retrospective Single-Group Pretest-Posttest Study
JMIR Form Res 2022;6(7):e35636
URL: https://formative.jmir.org/2022/7/e35636
doi:10.2196/35636
PMID:35802411

Original Paper

Novel Implementation Strategy to Electronically Screen and Signpost Patients to Health Behavior Apps: Mixed Methods Implementation Study (OptiMine Study)

Zarnie Khadjesari[1], PhD; Tracey J Brown[1], MSci; Alex T Ramsey[2], PhD; Henry Goodfellow[3], PhD; Sherine El-Toukhy[4], PhD; Lorien C Abroms[5], PhD; Helena Jopling[6], MBBS, MPhil; Arden Dierker Viik[6], MS; Michael S Amato[7], PhD

[1]Behavioural and Implementation Science Research Group, School of Health Sciences, University of East Anglia, Norwich, United Kingdom

[2]Department of Psychiatry, Washington University School of Medicine, St Louis, MO, United States

[3]Department of Primary Care and Population Health, University College London, London, United Kingdom

[4]Division of Intramural Research, The National Institute on Minority Health and Health Disparities, The National Institutes of Health, Bethesda, MD, United States

[5]Department of Prevention and Community Health, Milken Institute School of Public Health, The George Washington University, Washington DC, DC, United States

[6]Department of Public Health, West Suffolk NHS Foundation Trust, Bury St Edmunds, United Kingdom

[7]Truth Initiative, Washington DC, DC, United States

Corresponding Author:
Zarnie Khadjesari, PhD
Behavioural and Implementation Science Research Group
School of Health Sciences
University of East Anglia
Norwich Research Park
Norwich, NR4 7TJ
United Kingdom
Phone: 44 01603457648
Email: Z.Khadjesari@uea.ac.uk

Abstract

Background: Behavior change apps have the potential to provide individual support on a population scale at low cost, but they face numerous barriers to implementation. Electronic health records (EHRs) in acute care hospitals provide a valuable resource for identifying patients at risk, who may benefit from behavior change apps. A novel, emerging implementation strategy is to use digital technologies not only for providing support to help-seeking individuals but also for signposting patients at risk to support services (also called *proactive referral* in the United States).

Objective: The OptiMine study aimed to increase the reach of behavior change apps by implementing electronic signposting for smoking cessation and alcohol reduction in a large, at-risk population that was identified through an acute care hospital EHR.

Methods: This 3-phase, mixed methods implementation study assessed the acceptability, feasibility, and reach of electronic signposting to behavior change apps by using a hospital's EHR system to identify patients who are at risk. Phase 1 explored the acceptability of the implementation strategy among the patients and staff through focus groups. Phase 2 investigated the feasibility of using the hospital EHR to identify patients with target risk behaviors and contact them via SMS text message, email, or patient portal. Phase 3 assessed the impact of SMS text messages sent to patients who were identified as smokers or risky drinkers, which signposted them to behavior change apps. The primary outcome was the proportion of participants who clicked on the embedded link in the SMS text message to access information about the apps. The acceptability of the SMS text messages among the patients who had received them was also explored in a web-based survey.

Results: Our electronic signposting strategy—using SMS text messages to promote health behavior change apps to patients at risk—was found to be acceptable and feasible and had good reach. The hospital sent 1526 SMS text messages, signposting patients to either the National Health Service Smokefree or Drink Free Days apps. A total of 13.56% (207/1526) of the patients clicked on the embedded link to the apps, which exceeded our 5% a priori success criterion. Patients and staff contributed to the SMS text message content and delivery approach, which were perceived as acceptable before and after the delivery of the SMS text

messages. The feasibility of the SMS text message format was determined and the target population was identified by mining the EHR.

Conclusions: The OptiMine study demonstrated the proof of concept for this novel implementation strategy, which used SMS text messages to signpost at-risk individuals to behavior change apps at scale. The level of reach exceeded our a priori success criterion in a non–help-seeking population of patients receiving unsolicited SMS text messages, disconnected from hospital visits.

International Registered Report Identifier (IRRID): RR2-10.2196/23669

(JMIR Form Res 2022;6(7):e34271) doi:10.2196/34271

KEYWORDS

electronic health record; EHR; alcohol reduction; electronic messages; proactive messages; proactive outreach; smoking cessation; tobacco use; alcohol use; alcohol; smoking; mobile health; mHealth; mobile app

Introduction

More than one-third of cancers are preventable by adopting a healthier lifestyle, such as stopping smoking and reducing risky drinking; tackling these health behaviors would lead to significant health benefits, minimize comorbidity, and reduce the burden on the National Health Service (NHS) [1-3]. The traditional approach to health promotion in the hospital environment is centered on health professionals providing verbal advice or information in leaflets, with or without the offer of referral to a lifestyle service. This approach relies on the clinician's knowledge, skills, and confidence; time; and the availability of health promotion literature, of which all are significant barriers to delivery [4,5]. Attempts have been made to boost health promotion activities in acute care hospitals in England. For example, (1) financial incentives have been provided for hospitals that undertake screening and brief advice for smoking and risky drinking (2018-2020) [6]; (2) brief advice on health behaviors through opportunistic day-to-day interactions between patients and health care professionals is promoted via Making Every Contact Count, an NHS initiative [7]; and (3) public health specialists are increasingly being embedded into the acute care context in recognition of the links between health behaviors and chronic illnesses, such as cancer, and the pressure this puts on the NHS. However, these interventions are typically resource-intensive or have been discontinued owing to resource constraints.

Digital interventions, such as behavior change websites or apps, have been developed to bridge the evidence-to-practice gap. They can be used as an alternative or adjunct to face-to-face delivery and have a substantial evidence base as interventions [8-12]. Digital interventions can provide effective individual support on a population scale at low cost [13]. They overcome barriers to delivering face-to-face interventions, such as removing the stigma associated with seeking help and alleviating pressure on busy health professionals to deliver brief interventions. Digital interventions for addictive behaviors are more commonly offered in higher education, primary care, and community settings (although far from routine) but rarely in the context of acute health services [8,9,14-16]. Furthermore, the implementation of effective digital interventions is often not considered and relies on the help-seeking behavior of motivated individuals. Another underused digital resource in the acute care setting is the electronic health record (EHR) for screening patients at risk. Although screening is commonplace for promoting medication adherence, vaccine uptake, or cancer screening, systematic screening for health behaviors (such as risky alcohol consumption and tobacco smoking) for signposting to support services is in its infancy.

We developed a novel implementation strategy that uses the combined potential of three existing technologies to bridge the evidence-to-practice gap: (1) acute care hospital EHR for identifying at-risk individuals at scale, (2) health promotion apps and websites that provide behavior change support at an individual level, and (3) SMS text messages that are commonly used by health services to communicate with patients. Using electronic messages (such as SMS text message or email) to signpost (referred to as *proactive outreach* in the United States [17-19]) patients who are at risk, identified by the EHR, to health promotion apps is cheap, efficient, quick to implement using existing infrastructure, and scalable to other health behaviors and health care settings and has the potential to achieve behavior change at the population level. Although SMS text messaging has been used successfully as a treatment tool for smoking cessation [11], using it as a primary channel for outreach and enrollment to support services or treatment represents a novel application with preliminary success. Krebs et al [18] used outreach SMS text messages in a large New York health system to connect patients to quitline counseling. Furthermore, Abroms et al [20] used SMS text messaging to connect patients in the emergency department to a smoking cessation SMS text messaging program or quitline counseling.

So far, no previous studies have evaluated the ability of electronic messages to signpost patients who are not seeking help to behavior change apps, disconnected from hospital visits. The aim of this 3-phase, mixed methods implementation study (the OptiMine study) was to explore the acceptability (phase 1), feasibility (phase 2), and reach (phase 3) of electronic messages to signpost patients who smoke tobacco and drink alcohol at risky levels to behavior change apps. The primary outcome of phase 1 was a qualitative synthesis of patient-perceived attributes of signposting. The primary outcome of phase 2 was the estimated size and characteristics of the population that may be reached by the intervention. Findings from these 2 phases informed the design of phase 3, the primary outcome of which was the proportion of contacted patients who followed the signpost to access more information about the promoted behavior change apps (ie, the *click rate*). The approach and methods have been reported in our published protocol [21].

Methods

Ethics Approval

The NHS Research Ethics Committee and the Health Research Authority approved this study in June 2019. EHR data were accessed and analyzed by staff within the West Suffolk NHS Foundation Trust (WSFT) information service team. Patients at risk were identified by the WSFT information service team and received electronic messages directly from the hospital. All analyses were conducted by the WSFT information service and public health teams. The study team members had access only to anonymized, aggregated data. Authorization for using EHR data at the WSFT was provided by the Information Governance Team via Data Protection Impact Assessment (approval date: September 20, 2019) [22].

Theoretical Frameworks

We used the taxonomy of implementation outcomes by Proctor et al [23] to design this implementation study, focusing on acceptability, feasibility, and reach. Acceptability and feasibility are important precursors for effective uptake and reach of an intervention and were explored using qualitative focus groups (acceptability before implementation), a web-based survey (acceptability after implementation), and routinely collected data from the EHR (feasibility). We assessed *reach* as our measure of implementation success, operationalized as the proportion of patients who engaged with our SMS text message by clicking on an embedded link to access free apps to support health behavior change. The topic guides and survey exploring the acceptability of delivering electronic messages to patients were informed by the Perceived Attributes of eHealth Innovations [24]—an extension of the Diffusion of Innovations Theory [25], which has been applied through a validated questionnaire to test the acceptability of a digital innovation [26]. Further information on the model and its application in this study is provided in our published protocol [21].

Setting

The research was conducted between April 2019 and July 2020 within the West Suffolk Hospital, an acute NHS provider renowned for its world-leading delivery of care using digital technologies, that is, an acute Global Digital Exemplar Trust [27]. eCare (trade name: Cerner Millenium), the EHR system at the WSFT, was launched in 2016 and includes records for all outpatients and inpatients registered with the hospital. The EHR contains contact information, demographics, and health data such as chronic disease status needed for this study. Data on smoking and alcohol consumption status were collected as part of the lifestyle screening survey or the Activities of Daily Living assessment, routinely provided to patients on admission to the hospital. The hospital is situated in the rural region of West Suffolk in the East of England, where population characteristics are similar to those of the general population in England, with slightly higher proportions of individuals aged >65 years and White British residents [28].

Implementation Strategy and Behavior Change Apps

We electronically signposted patients to the NHS Smokefree and Drink Free Days apps. Public Health England has developed the Smokefree and the Drink Free Days apps as part of their One You campaign for supporting healthy lifestyles. These apps are freely available on the web [29] and are heavily promoted in the United Kingdom mass media. The apps are theoretically informed and use evidence-based behavior change techniques, such as goal setting and self-monitoring. This mixed methods implementation study explored acceptability (phase 1) and feasibility (phase 2) as important precursors for the reach of our implementation strategy (phase 3).

Phase 1: Acceptability of Electronic Signposting (Before Implementation)

Acceptability of the implementation strategy was explored before its delivery via focus groups. Patients and staff were invited to participate in face-to-face focus groups to explore their perspectives. Patients were eligible if they were smoking or drinking alcohol regularly. Eligible staff were those (1) in senior IT management roles, (2) responsible for administering lifestyle screening, or (3) involved in EHR data management and hospital communications. Patients were identified via volunteer coordinators, a news story on the hospital website, and a recruitment stall at the main hospital entrance. Staff were recruited directly through email invitation and a weekly staff newsletter. All participants were provided with a participant information sheet and a consent form. Patient focus groups were conducted on-site at the education center at WSFT, whereas staff focus groups were conducted in meeting rooms. Focus groups were audio-recorded and transcribed verbatim by a professional transcription company, removing any identifiable data. Transcripts were coded using NVivo (QSR International). Framework analysis was used to synthesize the findings of the focus groups, based on 3 domains of the Perceived Attributes Theory: compatibility (ie, degree to which electronic messaging was consistent with patient preferences), complexity (ie, degree to which electronic messaging was difficult to understand or act on), and relative advantage (ie, degree to which electronic messaging was superior to alternative or more traditional approaches) [24]. Subthemes were focused on the pragmatic development, refinement, and delivery of the implementation strategy.

Phase 2: Feasibility of Electronic Signposting

Feasibility of using the EHR to identify patients at risk who have mobile phone numbers, email addresses, and patient portal access was explored through data mining. Patients with alcohol consumption or smoking status that had been recorded or updated within the past 13 months were included. The Activities of Daily Living and customized lifestyle screening assessments are used by the WSFT to record alcohol consumption and smoking status on admission to the hospital. The following data were extracted and aggregated from the EHR by a hospital-based information analyst: smoking status (yes or no), alcohol consumption status: Alcohol Use Disorders Identification Test–Consumption (AUDIT-C) score [30] (low risk: 0-4, at risk: 5-9, and dependent: 10-12), sex (woman or man), mobile phone number (yes or no), email address (yes or no), and patient portal access (yes or no). Frequencies and percentages of patients with valid data in each of the fields mentioned

previously were reported to the study team, who had no direct access to the underlying data.

Phase 3: Reach of Electronic Signposting

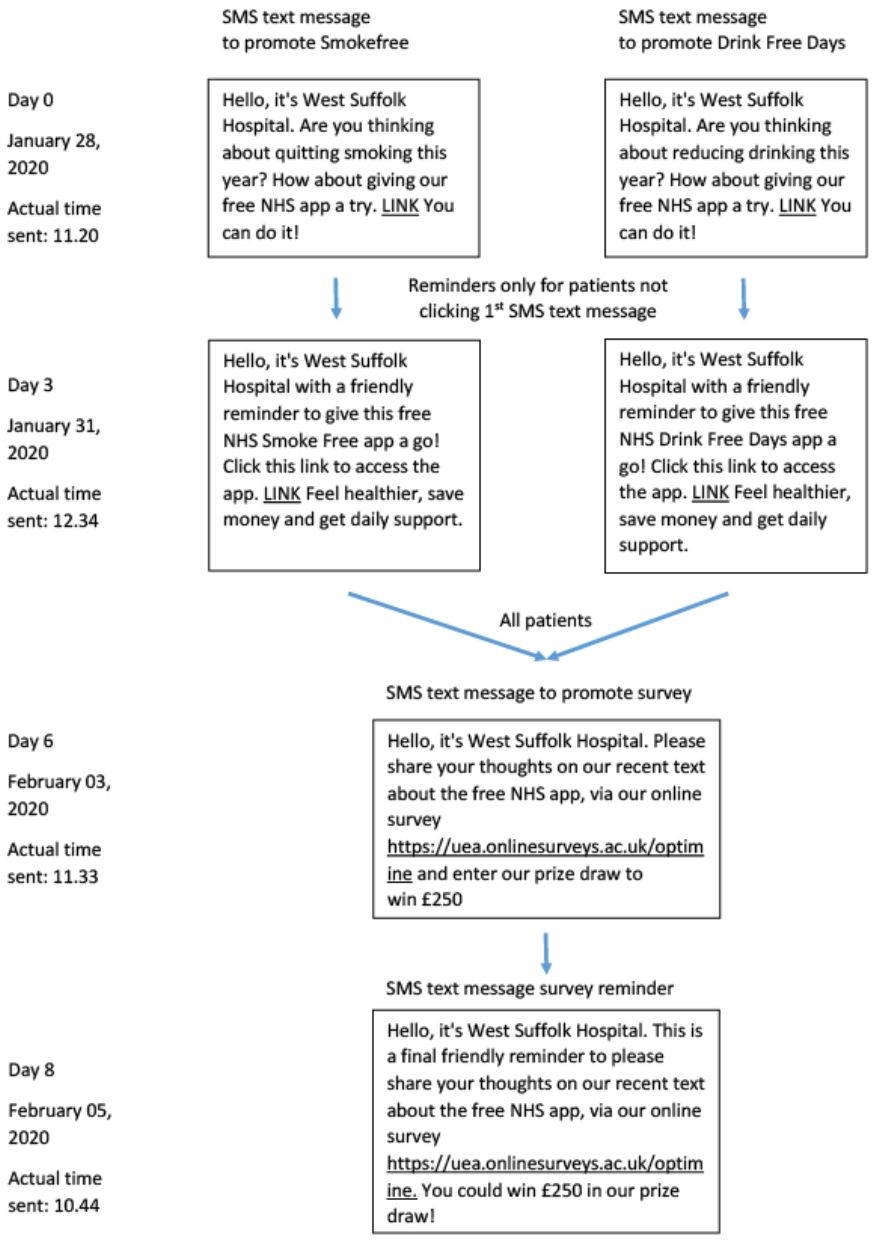
Participants

Eligible patients were adults (aged ≥18 years) recorded as smoking or drinking alcohol at risky levels within the past 13 months, with a valid mobile number. Patients were excluded if they were pregnant, were registered on the end-of-life pathway, or had opted out of communications from the hospital. As specified in the protocol, a minimum sample size of 383 per risk profile group would allow calculation of 95% CIs within a margin of –5% to +5% points, assuming a population proportion of 50% and a population size of 100,000 [21].

Procedure

SMS text messages were selected as the most acceptable and feasible electronic format based on the findings of phase 1 and phase 2. Although the protocol sought to compare 3 risk profiles (exclusive smokers, exclusive risky drinkers, and both), to reach the a priori specified minimum sample size, it was necessary to collapse to two risk profiles: (1) exclusive smokers and (2) risky drinkers regardless of smoking status. The hospital's SMS text messaging system was used to send an initial message to all the participants, signposting to either the Smokefree or Drink Free Days apps based on risk profile, with a second reminder message sent 3 days later to any participant who had not clicked the link yet. Unique link URLs were used to identify the participants who clicked the link. Figure 1 illustrates the content and delivery schedule of the messages.

Figure 1. Flowchart of SMS text message content and delivery.



Data Collection

To determine the characteristics of patients to whom SMS text messages were sent, data were retrieved from eCare, and analyses were performed by a public health manager with the support of coauthors MSA and ZK in July 2020. The following variables were extracted from the system and categorized as follows:

1. Name (for data linkage purposes only)
2. Hospital and NHS number (for data linkage purposes only)
3. Date of birth (for data linkage purposes only)
4. Age—categorized as 18-25, 26-35, 36-45, 46-55, 56-65, and >66 years
5. Sex—woman or man
6. Ethnicity—categories were merged to increase patient numbers in less-populated groups: any White background, any other background, and not known or not stated; refer to the study protocol for individual categories [21]
7. Postcode—used to calculate index of multiple deprivation decile—grouped into quintiles
8. Date of smoking or alcohol consumption screening at the hospital—categorized as days since screening: 1-4, 4-7, 7-10, 10-13, and >13 months
9. SMS text message sent, targeting smoking cessation or alcohol reduction
10. Embedded link to apps clicked or not clicked

Indices of Multiple Deprivation

Postcodes were entered into the government multiple deprivation lookup to obtain the deciles of multiple deprivation for each patient [31], where decile 1 represents the most deprived 10% of the population and decile 10 represents the least deprived 10% of the population.

Recency of Screening

The number of days since patients were screened for smoking or alcohol consumption was calculated by subtracting the date of smoking or alcohol screening from the date on which the smoking or alcohol SMS text message was sent. This was rounded to a whole number.

Health Data

The long-term health conditions stored as structured data in eCare were audited against the whole medical record to determine its suitability for use. Auditors used inpatient, emergency department, and general practice notes to collate patients' medical history. More than half of the manually audited records found different health data than those retrieved from the patient records. Thus, the health data were considered as too incomplete and inaccurate to be included in the study. Refer to the study protocol for the list of long-term health conditions originally intended for extraction from eCare [21].

Missing Data in AUDIT-C Fields

A large proportion of the records (1424/1975, 72.1%) that documented alcohol status had missing data in the AUDIT-C fields. Scores were calculated from the available fields, imputing a value of zero for missing fields, based on information from hospital staff that fields were most likely skipped because they were not applicable (eg, erroneously left blank instead of selecting *0*). Patients with missing data and calculated AUDIT-C scores of 7 to 8 were excluded from the study because it was not possible to be confident that these patients were not dependent drinkers, for whom the intervention would be clinically inappropriate.

Statistical Modeling

We determined a priori that 5% reach, as evidenced by clicking on the embedded links to the apps, would constitute a clinically meaningful level of reach, given the low-burden and scalable nature of the interventions [21]. This success criterion was based on the rationale that reaching even this modest proportion of non–help-seeking individuals with high-risk drinking and smoking behaviors via unsolicited SMS text messages at a time disassociated with their last hospital visit could have great impact at a population health level. The number of patients who clicked the embedded link within their respective SMS text message was reported along with their baseline characteristics. The denominator for reach rate was the total number of patients with that characteristic who received an SMS text message. R software (R Foundation for Statistical Computing) was used to model the logistic regression, where we used the glm function with a Poisson distribution. To investigate potential relationships of common demographic covariates and assess any effect of screening recency, the model included all patient-level variables that were available: sex, age group, ethnicity group, index of multiple deprivation quintile, and days since screening group. Relative risks (RRs) with 95% CIs for each characteristic were derived from the model in R, using exponentiated values of the model coefficients.

Acceptability of Electronic Signposting (After Implementation)

Acceptability of the implementation strategy was also explored after its delivery via a web-based survey. Eligible patients were those to whom a signpost SMS text message was sent. Then, 6 days after the initial signpost SMS text message was sent, 2 additional SMS text messages were sent to the participants (an initial message and a reminder), inviting them to participate in a web-based survey about their views on receiving the signpost SMS text messages. Participant information sheets and consent forms were incorporated into the web-based survey. We used JISC web-based surveys, a free web-based platform designed for academic research and public sector organizations. Participants were given 15 days to complete the survey.

Results

Phase 1: Acceptability of Electronic Signposting (Before Implementation)

A total of 10 patients participated in 2 focus groups (group 1: n=3, 30% of the participants; group 2: n=6, 60% of the participants) and 1 individual interview (n=1, 10% of the participants; owing to low turnout for the focus group). A total of 14 staff members participated in 3 focus groups (group 1: n=5, 36% senior managers; group 2: n=1, 7% nurse and n=1, 7% pharmacy technician; group 3: n=7, 50% members of IT staff and communications officers). Patients' ages were collected as categories and ranged from 46 to >66 years; 60% (6/10) were

women. Most patients (7/10, 70%) were members of the hospital's volunteer group and patient portal user group, with 20% (2/10) of them being members of the public and 10% (1/10) being members of staff identified via the recruitment stall. Findings from this phase suggested that most patients found SMS text messaging as the most acceptable form of electronic message for receiving communications. A more detailed summary of the key findings under each theme and subtheme is presented in Multimedia Appendix 1 (qualitative focus group findings), along with illustrative quotes and the approach that informed the development and delivery of electronic messages.

Phase 2: Feasibility of Electronic Signposting

The dynamic nature of an acute care hospital EHR database was found to be an implementation challenge. The mining of data from eCare needed to be performed on different days, owing to the size of the data queries, and with different search strategies depending on the fields required. Furthermore, the results changed depending on the date on which the queries were run (both the denominators and numerators). Queries that were intended to generate data sets that would include the whole EHR population could only be achieved by limiting the number of variables in the data set; large queries often took several hours to run and were at risk of timing out before they had completed.

The total number of adult patients in the eCare system on October 9, 2019, was 228,982, of which 1092 (0.48%) adults were recorded as smokers (smoking and drinking status data extracted on January 10, 2020). Among the 0.74% (1702/228,982) of the patients who were recorded as drinking alcohol, 23.62% (402/1702) patients were recorded as at-risk drinkers. An additional 0.55% (1249/228,982) and 0.02% (51/228,982) of the patients were reported as drinking at low risk or dependent levels, respectively. Most patients (226,784/228,982, 99.04%) had missing data in the EHR regarding their tobacco or alcohol use status. Smoking and drinking status are only recorded in a way that is retrievable when a patient has an inpatient admission. The proportion of the catchment population that is admitted each year is typically 13% [32]. The proportions of admitted patients who were screened for smoking and alcohol use in the financial year 2018 to 2019 were reported by WSFT information service staff to be 64% and 68%, respectively.

Of the 228,982 patients, 146,171 (63.84%) had mobile phone numbers, of which 707 (0.48%) were recorded as smokers and 271 (0.19%) were recorded as at-risk drinkers. Of the 228,982 patients, the total number of patients with email addresses was 32,375 (14.14%), of which 137 (0.42%) were recorded as smokers and 115 (0.36%) were recorded as at-risk drinkers. The database of patient portal users was independent of the EHR, and could not be linked to the EHR by the WSFT information service staff. The numbers of patient portal users recorded as smoking and drinking could not be determined. Therefore, the consideration of the patient portal as a form of message delivery was discontinued in this study. Combined with the findings from phase 1, these feasibility findings from phase 2 reinforced the decision to use SMS text messages as the channel for the electronic messages.

Phase 3: Reach of Electronic Signposting

Baseline Characteristics and Risk Profile Groups

On the basis of the findings in phase 2, the participant sample was mined from eCare on January 10, 2020. The sampling frame used the most recent admissions to the hospital with validated data to identify patients whose records were most likely to be up to date and accurate. A 13-month time frame (October 1, 2018, to November 30, 2019) was used to obtain a sufficiently large population of patients to meet the minimum sample size required, which was balanced against recency of the data. A total of 6521 people admitted during the time frame were aged ≥18 years, not on the end-of-life pathway, and not pregnant and had a mobile phone number recorded and either their smoking or alcohol status recorded. Owing to the relatively small number of patients who could be identified as exclusively consuming alcohol at risky levels but who were nonsmokers, all participants with risky drinking were combined into a single risk profile group, regardless of their smoking status. The resulting 2 risk profile groups (smoking only and risky drinking with or without smoking) represent a deviation in analysis from the protocol, which aimed to create 3 risk profile groups (smoking only, risky drinking only, and smoking and risky drinking). This deviation was necessary to meet our a priori minimum sample size of 383 participants per group, which was also specified in the protocol. The decision to collapse the risky drinking groups was based exclusively on the baseline data, before the analyses of outcomes. Ultimately, of the 1526 individuals, the selected sample included 1103 (72.28%) individuals who were recorded as smokers only, 276 (18.09%) individuals who were recorded as risky drinkers only, and 147 (9.63%) individuals who were recorded as both smokers and risky drinkers. Table 1 shows the baseline characteristics of the participants.

Table 1. Baseline characteristics.

Characteristics	Smoker only[a] (n=1103)	Risky drinker with or without also being a smoker[b] (n=423)
Sex, n (%)		
Women	553 (50.14)	129 (30.5)
Men	550 (49.86)	294 (69.5)
Not recorded	N/A[c]	N/A
Age (years), mean (SD)	47.78 (17.68)	56.43 (16.27)
18-25, n (%)	124 (11.24)	17 (4)
26-35, n (%)	200 (18.13)	29 (6.9)
36-45, n (%)	192 (17.41)	60 (14.2)
46-55, n (%)	205 (18.59)	90 (21.3)
56-65, n (%)	176 (15.96)	98 (23.2)
66-75, n (%)	125 (11.33)	78 (18.4)
>75, n (%)	81 (7.34)	51 (12.1)
Not recorded	N/A	N/A
Ethnicity, n (%)		
White	1021 (92.57)	395 (93.4)
Mixed	2 (0.18)	1 (0.2)
Asian or Asian British	2 (0.18)	1 (0.2)
Black or Black British	5 (0.45)	0 (0)
Other ethnic groups	21 (1.90)	8 (1.9)
Not recorded	52 (4.71)	18 (4.3)
Index of multiple deprivation (quintile), n (%)		
1	77 (6.98)	17 (4)
2	279 (25.29)	92 (21.7)
3	371 (33.64)	147 (34.8)
4	251 (22.76)	111 (26.2)
5	121 (10.97)	49 (11.6)
Not recorded	4 (0.36)	7 (1.7)
Recency of screening data (months), n (%)		
0-1	N/A	N/A
1-3	119 (10.79)	54 (12.8)
3-6	249 (22.57)	127 (30)
6-12	485 (43.97)	209 (49.4)
>12	250 (22.67)	33 (7.8)
Not recorded	N/A	N/A

[a]Group includes all participants recorded as smoking and not drinking alcohol at risky levels.

[b]Group includes all participants recorded as drinking alcohol at risky levels, regardless of smoking status, owing to sample size considerations. The risky drinker status is defined as an Alcohol Use Disorders Identification Test–Consumption score of 5 to 10 if 3 of its items were completed or a score of 5 to 6 if only 2 items were completed.

[c]N/A: not applicable.

Reach

In January 2020, an SMS text message was sent to 1526 patients, signposting them to either the NHS Smokefree (n=1103, 72.28%) or the Drink Free Days apps (n=423, 27.72%). A total of 13.56% (207/1526) of the participants clicked on the embedded link to the apps (smokers: 26/207, 12.56% and risky drinkers: 34/207, 16.43%), which exceeded our 5% a priori

success criterion in both groups. Figure 2 shows a CONSORT (Consolidated Standards of Reporting Trials) diagram. Characteristics of patients who clicked the embedded link within the SMS text message versus the characteristics of those who did not click the embedded link are presented in Table 2. The only significant differences in baseline characteristics were observed among smokers. Among smokers, the lowest click rate was among participants aged ≥66 years (6.8%). Compared with this age group as reference, smoking participants aged 36 to 45 years (18.2%; RR=2.6; 95% CI 1.4-4.6) and those aged 56 to 65 years (13.1%; RR=1.98; 95% CI 1.05-3.73) were significantly more likely to click. Male smokers were significantly less likely to click than female smokers (10.2% vs 14.8%; RR=0.71; 95% CI 0.51-0.98). Although no significant differences by age were observed among risky drinkers, a numerically different distribution emerged, such that risky drinkers aged 56 to 65 years were the most likely to click the link (20.4%).

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) diagram. *Alcohol Use Disorders Identification Test–Consumption (AUDIT-C) score of 5-10 if 3 fields populated or score of 5-6 if 2 fields populated; **AUDIT-C score<5; ***both=patients who are both smokers and risky drinkers.

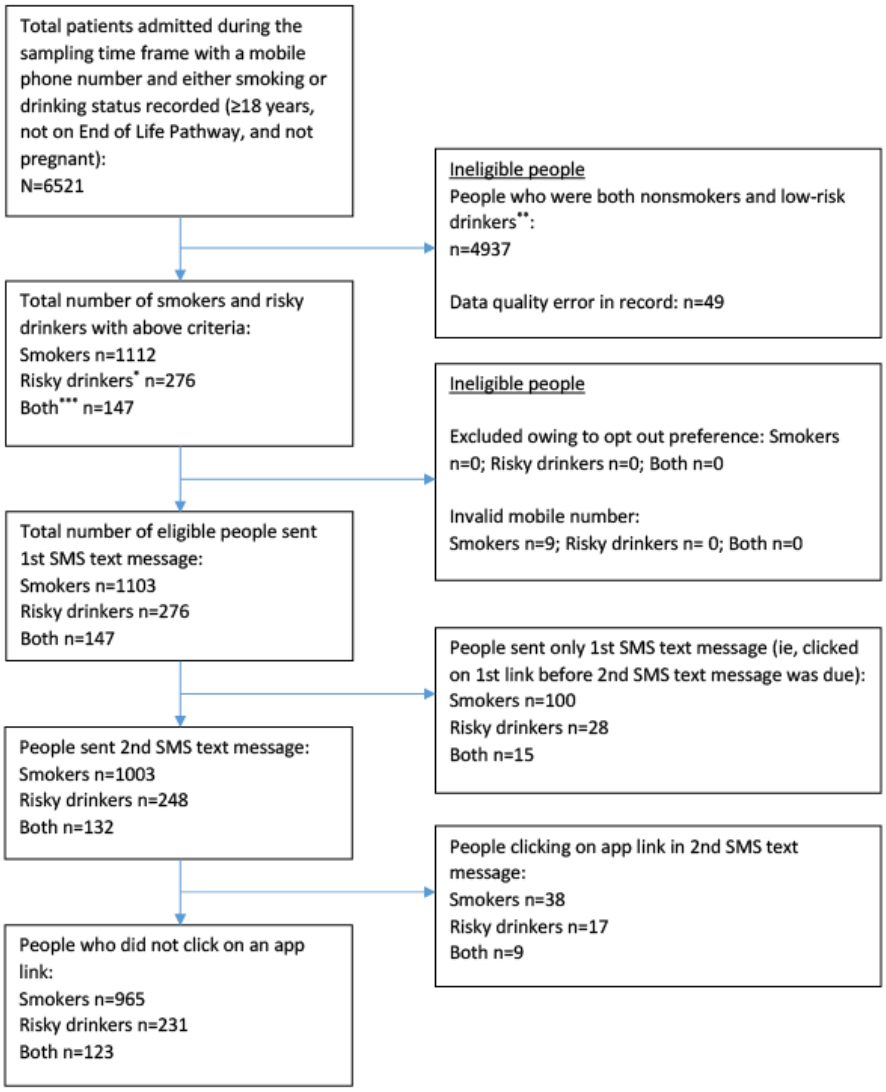


Table 2. Comparison of characteristics of patients who clicked versus those who did not click on links to apps within the SMS text message.

Characteristics	Smokers—patients who clicked versus those who did not click[a] (n=1103)			Risky drinkers—patients who clicked versus those who did not click[b] (n=423)		
	Patients who clicked, n (%)	RR[c] (95% CI)	*P* value	Patients who clicked, n (%)	RR (95% CI)	*P* value
Overall click rate	138 (12.51)	—[d]	N/A[e]	69 (16.3)	N/A	N/A
Sex						
Women	82 (14.83)	Reference	—	21 (16.3)	Reference	—
Men	56 (10.18)	0.71 (0.51-0.98)	.04	48 (16.3)	0.92 (0.56-1.51)	.74
Age (years)						
18-25	14 (11.29)	1.55 (0.76-3.13)	.23	1 (5.9)	0.38 (0.06-2.48)	.31
26-35	26 (13)	1.86 (1-3.44)	.05	1 (3.5)	0.22 (0.03-1.44)	.11
36-45	35 (18.23)	2.56 (1.42-4.64)	.002	10 (16.7)	1.18 (0.58-2.43)	.65
46-55	26 (12.68)	1.83 (0.99-3.39)	.06	17 (18.9)	1.24 (0.67-2.31)	.50
56-65	23 (13.07)	1.98 (1.05-3.73)	.03	20 (20.4)	1.34 (0.74-2.42)	.34
>66	14 (6.80)	Reference	—	20 (15.5)	Reference	—
Ethnicity						
White	120 (11.75)	Reference	—	65 (16.5)	Reference	—
People of color	6 (20)	1.63 (0.74-3.57)	.22	2 (20)	1.53 (0.39-6.06)	.54
Not known or not stated	12 (23.08)	2.12 (1.20-3.75)	.01	2 (11.1)	0.64 (0.17-2.44)	.52
Postcode (for index of multiple deprivation; quintile)						
1	10 (12.99)	0.96 (0.45-2.04)	.92	2 (11.7)	0.79 (0.18-3.47)	.76
2	37 (13.26)	0.96 (0.55-1.68)	.88	15 (16.3)	1.02 (0.45-2.30)	.96
3	49 (13.21)	0.96 (0.56-1.65)	.90	23 (15.7)	1.01 (0.48-2.16)	.97
4	26 (10.36)	0.76 (0.42-1.37)	.36	21 (18.9)	1.25 (0.58-2.69)	.57
5	16 (13.22)	Reference	—	8 (16.3)	Reference	—
Recency of screening data (months)						
1-4	28 (13.93)	1.11 (0.66-1.88)	.70	16 (16)	1.2 (0.17-8.33)	.85
4-7	29 (11.20)	0.88 (0.52-1.47)	.61	16 (12.8)	0.99 (0.14-6.91)	.99
7-10	34 (14.72)	1.18 (0.71-1.96)	.52	17 (15.5)	1.21 (0.17-8.44)	.85
10-13	24 (10.08)	0.80 (0.46-1.38)	.42	19 (23.2)	1.82 (0.26-12.72)	.54
>13	23 (13.22)	Reference	—	1 (16.7)	Reference	—

[a]These messages were sent to patients recorded as smoking and not recorded as drinking alcohol at risky levels.

[b]These messages were sent to all patients recorded as drinking alcohol at risky levels, including those who were also recorded as smoking. These data were merged owing to the low number of patients recorded as drinking at risky levels.

[c]RR: relative risk.

[d]Not available.

[e]N/A: not applicable.

Acceptability (After Implementation)

The survey was completed by 3.67% (56/1526) of participants. Among the 56 survey responders, 18 (32%) participants reported that they had clicked on the link within the SMS text message and 9 (16%) participants reported that they downloaded the app. Approximately two-thirds (36/56, 64%) of the participants found the messages to be at least slightly *helpful*. The message was found to be *not at all difficult* by almost all patients (55/56, 98%). Most participants (51/56, 91%) were happy with the wording of the SMS text message. Almost half of the participants (26/56, 46%) reported that they would like the hospital to deliver this service; however, an opt-out option was also popular (21/56, 38%). Most of the respondents (42/56, 75%) submitted free-text comments. Common positive themes were that the SMS text messages were supportive (7/56, 13%), easy to understand (6/56, 11%), and brief (6/56, 11%). Common

negative themes were that the messages were irrelevant (9/56, 16%), not supportive (4/56, 7%), and unwanted or unexpected (3/56, 5%).

Discussion

Principal Findings

Our electronic signposting strategy that used SMS text messages to promote health behavior change apps to patients at risk via an acute care hospital EHR was found to be acceptable and feasible and to achieve clinically meaningful reach. The hospital sent 1526 SMS text messages signposting patients to either the NHS Smokefree or Drink Free Days apps, with 13.56% (207/1526) of the patients clicking on the embedded link to the apps. The level of reach exceeded our 5% a priori success criterion in a non–help-seeking population of patients receiving unsolicited SMS text messages, disconnected from hospital visits. The strategy was found to be acceptable both before and after implementation. The OptiMine study demonstrated the proof of concept for this novel implementation strategy, specifically signposting digital health promotion interventions at scale to at-risk individuals who are not seeking help.

A 13.56% (207/1526) click rate is promising, considering the comparable outcomes reported in the literature. For example, in a study to promote colorectal cancer screening, patients—identified from EHRs as being overdue for their colorectal screening—received an electronic message from their primary care physicians with information about their overdue status, methods to arrange a screening appointment, and a web-based risk assessment tool for colorectal cancer [33]. The intervention led to 3% of the patients requesting colorectal screening. Our 13.56% (207/1526) click rate exceeds this result. However, it is important to account for the action requested from patients, which will trigger varying levels of response, such as the 3% response rate for colorectal screenings versus the 13.56% (207/1526) click rate on our embedded link to behavioral interventions. Future research should build on this study and aim to define appropriate thresholds for e-referral interventions that correspond to different behaviors that require varying degrees of involvement from patients (eg, time and effort invested, duration of behavior, and difficulty of behavior). Furthermore, it is important to note that our SMS text messages were sent from the patients' hospital, which would be expected to foster trust in the messages and their source and increase uptake.

Using SMS text messages to signpost patients to addictive behavior apps is a scalable implementation strategy, which can provide individualized support to patients at risk. As such, cost-effectiveness increases as the strategy is scaled up. Our low-cost and low-burden strategy has the potential to reach large number of patients at high risk in a novel form (proactive referral to existing tools). Health promotion is typically delivered in primary care and community settings, but it is equally important in acute care settings, as health behaviors such as smoking and risky drinking can cause long-term health conditions and exacerbate existing health conditions. Following the success of this study, the WSFT plans to routinely deliver health promotion advice using electronic signposting.

Strengths and Limitations

The mixed methods design helped to explore multiple implementation outcomes that assess both proximal outcomes and indicators of success [23]. This was a novel and pragmatic study, facilitated by NHS staff in a busy hospital setting. Major strengths of this study included the real-world application of the strategy with patients at the hospital and readiness of the implementation context. Senior-level leadership and buy-in from a public health consultant within the digital health team were instrumental in the successful implementation of this project. Global Digital Exemplar Trusts have EHRs and in-house expertise to support the setup and delivery of SMS text message signposting. We identified the following stakeholders as pivotal to the successful setup, implementation, and evaluation of the strategy: patient representatives, information analysts (assess infrastructure and access to data), deputy chief information officer (design and oversight of message implementation), IT integration developer (implement SMS text message signposting and send messages), public health manager (in-house data analysis), communications team, volunteer coordinator, and administrative support. Although the readiness of the implementation context is considered a strength of our study, it could be seen as a limitation to broad scalability in less-ready, low-resourced contexts, which may serve more disadvantaged or underserved populations.

There were challenges in using the EHR as a research database. The data set was extracted directly from the EHR at different time points owing to the size of the EHR, time to download the data, and workload of hospital staff. Queries for different parts of the data were run on different days over a 4-month period from October 2019. Each day, the total number of EHR records changes owing to new records being created, existing records being edited, and people dying. Smoking and alcohol consumption status were not retrievable from hospital day patients, which meant that there were large amounts of missing data. Furthermore, there were difficulties in retrieving accurate health condition data, and matching to eligible patients was not possible. Health condition data rely on the population of the appropriate fields in a patient record; if the condition is listed in the wrong place or not recorded at all, it cannot be electronically retrieved. This occurs with some frequency across the EHR, and this limitation needs to be addressed before using electronically retrieved health data for similar studies in the future. As such, the full capacity of patient reach was not used owing to missing data, and the impact of long-term health conditions on the likelihood of clicking the embedded link is unknown.

Our primary outcome—whether each participant did or did not click the embedded link—directly measured the willingness to access information about behavior change apps among a non–help-seeking population. However, we did not measure the proportion of participants who actually downloaded and used the app. Although technical limitations prevented us from measuring app use, uptake should be assessed through future studies to support broader dissemination of this approach. In addition, response rates to the web-based questionnaire were low and may have been subject to response bias. This may have been owing to message fatigue, where patients had previously

received up to 2 SMS text messages signposting to the apps, followed by 2 messages inviting participation in the survey. Finally, the SMS text messages did not include instructions for opting out of future messages, which is a requirement in the United States. Future studies should examine click rates in the presence of explicit opt-out instructions.

Implications for Further Research and Practice

The WSFT views this approach as the future for routine delivery of digital interventions for health promotion within their acute care context and as an important adjunct to the opportunistic behavior change interventions that are made by health care professionals. Further refinement and evaluation are needed to optimize the SMS text message content and delivery approach. Additional studies with other populations are needed to understand the best strategy for implementation at other hospitals and institutions. The COVID-19 pandemic has highlighted the urgent need for scalable health promotion support, which can be delivered via digital technology. There is also an international call for greater access to high-quality, safe, and effective addictive behavior apps [34]. Regarding refinement and further evaluation of SMS text message signposting to apps, we propose the following ideas for further development:

1. Tailor message content to groups of people who are less likely to engage, such as older people, low socioeconomic status groups, and ethnic minority groups.
2. A timely trigger for the SMS text message, possibly in the context of a face-to-face consultation, may be more effective than sending all the messages at the same time. The SMS text messages were sent in January 2020 to optimize the click rate by taking advantage of the seasonal high demand for support, and therefore, other times of the year may be less or more likely to have high rates of engagement.
3. Use the strategy to target other health behaviors.
4. Explore other implementation outcomes, such as implementation cost and cost-effectiveness, compared with other methods for promoting lifestyle change; sustainability; and fidelity, including engagement with the app.
5. Investigate system interventions for addressing barriers identified regarding the reliability of EHR data.

Conclusions

The OptiMine study demonstrated the proof of concept for this novel implementation strategy, which used SMS text messages to signpost at-risk individuals to behavior change apps at scale. The level of reach exceeded our a priori success criterion for a non–help-seeking population of patients receiving unsolicited SMS text messages, disconnected from hospital visits. These findings suggest that electronic signposting to support is an effective method for health systems to proactively engage meaningful proportions of their at-risk populations.

Acknowledgments

The authors are grateful to the patients and staff at the West Suffolk Foundation Trust, who participated in this study, and the team that facilitated the study on-site, namely, Adam Jull, Cassia Nice, Josh Wigley, Liam McLaughlin, Simon Reeve, Anna Hollis, Penny Molkenthin, and Paul Oats. This study was funded by a Cancer Research UK Cancer Policy Research Centre innovation grant (C69129/A28665). The authors acknowledge the support from the National Institute for Health Research (NIHR) Clinical Research Network. The effort of SE was supported by the Division of Intramural Research of the National Institute on Minority Health and Health Disparities, National Institutes of Health. ATR was supported by the National Institute on Drug Abuse grant (K12DA041449). This is a summary of the study supported by the NIHR Applied Research Collaboration East of England. The views expressed are those of the authors and not necessarily those of the National Health Service, NIHR, UK Department of Health and Social Care, US National Institutes of Health, US Department of Health and Human Services, or West Suffolk Foundation Trust.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Qualitative focus group findings.
[DOCX File , 24 KB - formative_v6i7e34271_app1.docx]

References

1. Statistics on Alcohol, England 2020. NHS Digital. URL: https://digital.nhs.uk/data-and-information/publications/statistical/statistics-on-alcohol/2020 [accessed 2022-06-01]
2. Statistics on Smoking, England 2020. NHS Digital. URL: https://digital.nhs.uk/data-and-information/publications/statistical/statistics-on-smoking/statistics-on-smoking-england-2020 [accessed 2022-06-23]
3. Newton JN, Briggs AD, Murray CJ, Dicker D, Foreman KJ, Wang H, et al. Changes in health in England, with analysis by English regions and areas of deprivation, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015 Dec 05;386(10010):2257-2274 [FREE Full text] [doi: 10.1016/S0140-6736(15)00195-6] [Medline: 26382241]

4. Keyworth C, Epton T, Goldthorpe J, Calam R, Armitage CJ. 'It's difficult, I think it's complicated': health care professionals' barriers and enablers to providing opportunistic behaviour change interventions during routine medical consultations. Br J Health Psychol 2019 Sep 12;24(3):571-592 [FREE Full text] [doi: 10.1111/bjhp.12368] [Medline: 30977291]
5. Keyworth C, Epton T, Goldthorpe J, Calam R, Armitage CJ. Are healthcare professionals delivering opportunistic behaviour change interventions? A multi-professional survey of engagement with public health policy. Implement Sci 2018 Sep 21;13(1):122 [FREE Full text] [doi: 10.1186/s13012-018-0814-x] [Medline: 30241557]
6. Health matters: tobacco and alcohol CQUIN. NHS England. URL: https://www.gov.uk/government/publications/health-matters-preventing-ill-health-from-alcohol-and-tobacco/health-matters-preventing-ill-health-from-alcohol-and -tobacco-use [accessed 2022-06-23]
7. Making Every Contact Count (MECC): consensus statement. NHS. URL: https://www.england.nhs.uk/publication/making-every-contact-count-mecc-consensus-statement/ [accessed 2022-06-23]
8. Sundström C, Blankers M, Khadjesari Z. Computer-based interventions for problematic alcohol use: a review of systematic reviews. Int J Behav Med 2017 Oct;24(5):646-658 [FREE Full text] [doi: 10.1007/s12529-016-9601-8] [Medline: 27757844]
9. Riper H, Hoogendoorn A, Cuijpers P, Karyotaki E, Boumparis N, Mira A, et al. Effectiveness and treatment moderators of internet interventions for adult problem drinking: an individual patient data meta-analysis of 19 randomised controlled trials. PLoS Med 2018 Dec 18;15(12):e1002714 [FREE Full text] [doi: 10.1371/journal.pmed.1002714] [Medline: 30562347]
10. Colbert S, Thornton L, Richmond R. Smartphone apps for managing alcohol consumption: a literature review. Addict Sci Clin Pract 2020 May 07;15(1):17 [FREE Full text] [doi: 10.1186/s13722-020-00190-x] [Medline: 32381062]
11. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y, Dobson R. Mobile phone text messaging and app-based interventions for smoking cessation. Cochrane Database Syst Rev 2019 Oct 22;10:CD006611 [FREE Full text] [doi: 10.1002/14651858.CD006611.pub5] [Medline: 31638271]
12. Behaviour change: digital and mobile health interventions. National Institute for Health and Care Excellence. URL: https://www.nice.org.uk/guidance/ng183/resources/behaviour-change-digital-and-mobile-health-interventions-pdf-66142020002245 [accessed 2022-06-23]
13. Abrams DB, Graham AL, Levy DT, Mabry PL, Orleans CT. Boosting population quits through evidence-based cessation treatment and policy. Am J Prev Med 2010 Mar;38(3 Suppl):S351-S363 [FREE Full text] [doi: 10.1016/j.amepre.2009.12.011] [Medline: 20176308]
14. Ramsey AT, Satterfield JM, Gerke DR, Proctor EK. Technology-based alcohol interventions in primary care: systematic review. J Med Internet Res 2019 Apr 08;21(4):e10859 [FREE Full text] [doi: 10.2196/10859] [Medline: 30958270]
15. Kaner EF, Beyer FR, Garnett C, Crane D, Brown J, Muirhead C, et al. Personalised digital interventions for reducing hazardous and harmful alcohol consumption in community-dwelling populations. Cochrane Database Syst Rev 2017 Sep 25;9:CD011479 [FREE Full text] [doi: 10.1002/14651858.CD011479.pub2] [Medline: 28944453]
16. Nolan MB, Warner DO. Perioperative tobacco use treatments: putting them into practice. BMJ 2017 Sep 06;358:j3340. [doi: 10.1136/bmj.j3340] [Medline: 28877905]
17. Kruse GR, Park E, Haberer JE, Abroms L, Shahid NN, Howard SE, et al. Proactive text messaging (GetReady2Quit) and nicotine replacement therapy to promote smoking cessation among smokers in primary care: a pilot randomized trial protocol. Contemp Clin Trials 2019 May;80:48-54 [FREE Full text] [doi: 10.1016/j.cct.2019.03.006] [Medline: 30923022]
18. Krebs P, Sherman S, Wilson H, El-Shahawy O, Abroms L, Zhao X, et al. Text2Connect: a health system approach to engage tobacco users in quitline cessation services via text messaging. Transl Behav Med 2020 Feb 03;10(1):292-301. [doi: 10.1093/tbm/ibz033] [Medline: 32011721]
19. Japuntich SJ, Hammett PJ, Rogers ES, Fu S, Burgess DJ, El Shahawy O, et al. Effectiveness of proactive tobacco cessation treatment outreach among smokers with serious mental illness. Nicotine Tob Res 2020 Aug 24;22(9):1433-1438. [doi: 10.1093/ntr/ntaa013] [Medline: 31957794]
20. Abroms LC, Wu K, Krishnan N, Long M, Belay S, Sherman S, et al. A pilot randomized controlled trial of text messaging to increase tobacco treatment reach in the emergency department. Nicotine Tob Res 2021 Aug 18;23(9):1597-1601. [doi: 10.1093/ntr/ntab036] [Medline: 33684207]
21. Amato MS, El-Toukhy S, Abroms LC, Goodfellow H, Ramsey AT, Brown T, et al. Mining electronic health records to promote the reach of digital interventions for cancer prevention through proactive electronic outreach: protocol for the mixed methods OptiMine study. JMIR Res Protoc 2020 Dec 31;9(12):e23669 [FREE Full text] [doi: 10.2196/23669] [Medline: 33382041]
22. Data Protection Impact Assessment (DPIA). NHS West Suffolk. URL: https://www.wsh.nhs.uk/CMS-Documents/DPIA/DPIA-OptiMine.pdf [accessed 2022-06-23]
23. 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]
24. Atkinson NL. Developing a questionnaire to measure perceived attributes of eHealth innovations. Am J Health Behav 2007;31(6):612-621. [doi: 10.5555/ajhb.2007.31.6.612] [Medline: 17691874]
25. Rogers EM. Diffusion of Innovations. 5th Edition. New York, NY: Free Press; 2003.

26. Khadjesari Z, Boufkhed S, Vitoratou S, Schatte L, Ziemann A, Daskalopoulou C, et al. Implementation outcome instruments for use in physical healthcare settings: a systematic review. Implement Sci 2020 Aug 18;15(1):66 [FREE Full text] [doi: 10.1186/s13012-020-01027-6] [Medline: 32811517]
27. Global digital exemplars. NHS. URL: https://www.england.nhs.uk/digitaltechnology/connecteddigitalsystems/exemplars/ [accessed 2022-06-23]
28. Welcome to the Suffolk Observatory. Suffolk Observatory. URL: https://www.suffolkobservatory.info/ [accessed 2022-06-23]
29. Better health-Kickstart your health. NHS. URL: https://www.nhs.uk/better-health/ [accessed 2022-06-24]
30. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [doi: 10.1001/archinte.158.16.1789] [Medline: 9738608]
31. English indices of deprivation. Ministry of Housing, Communities & Local Government. URL: http://imd-by-postcode.opendatacommunities.org/imd/2019 [accessed 2022-06-23]
32. NHS Acute (Hospital) Trust Catchment Populations. Public Health England. URL: https://tinyurl.com/3a2u85ft [accessed 2022-06-23]
33. Sequist TD, Zaslavsky AM, Colditz GA, Ayanian JZ. Electronic patient messages to promote colorectal cancer screening: a randomized controlled trial. Arch Intern Med 2011 Apr 11;171(7):636-641 [FREE Full text] [doi: 10.1001/archinternmed.2010.467] [Medline: 21149743]
34. Khadjesari Z, Brown T, Naughton F. Regulation and accreditation of addictive behaviour applications-navigating the landscape. Addiction 2021 Dec;116(12):3276-3283 [FREE Full text] [doi: 10.1111/add.15484] [Medline: 33739480]

Abbreviations

AUDIT-C: Alcohol Use Disorders Identification Test–Consumption
CONSORT: Consolidated Standards of Reporting Trials
EHR: electronic health record
NHS: National Health Service
NIHR: National Institute for Health Research
RR: relative risk
WSFT: West Suffolk NHS Foundation Trust

Edited by A Mavragani; submitted 14.10.21; peer-reviewed by M Bird, W Cha; comments to author 08.01.22; revised version received 12.03.22; accepted 31.03.22; published 11.07.22.

Please cite as:
Khadjesari Z, Brown TJ, Ramsey AT, Goodfellow H, El-Toukhy S, Abroms LC, Jopling H, Dierker Viik A, Amato MS
Novel Implementation Strategy to Electronically Screen and Signpost Patients to Health Behavior Apps: Mixed Methods Implementation Study (OptiMine Study)
JMIR Form Res 2022;6(7):e34271
URL: https://formative.jmir.org/2022/7/e34271
doi:10.2196/34271
PMID:35816374

Original Paper

A Web Platform for Standardized Data Acquisition, Processing, and Export in the Child Psychopathology Clinical Routine (MedicalBIT): Design and Implementation Study

Paola Colombo[1], MSc, PhD; Silvia Busti Ceccarelli[1], MSc; Stefano Pacchiarini[2], ING; Stefano Cribellati[2], ING; Massimo Molteni[1], MD

[1]Child Psychopathology Unit, Istituto di Ricovero e Cura a Carattere Scientifico Eugenio Medea, Bosisio Parini, Italy
[2]SEGE Srl, Milan, Italy

Corresponding Author:
Silvia Busti Ceccarelli, MSc
Child Psychopathology Unit
Istituto di Ricovero e Cura a Carattere Scientifico Eugenio Medea
Via Don Luigi Monza 20
Bosisio Parini, 23842
Italy
Phone: 39 +39 031 877623
Email: silvia.busti@lanostrafamiglia.it

Abstract

Background: The rapid extent of digital innovation for the collection of data has transformed the way in which health professionals collect, share, and analyze health information for better clinical decision-making and health care. In the last decade, there has been an increased interest in telemedicine by mental health agencies; the gap between the need for care and both diagnosis and treatment is wide, and digital technology could play an important role in filling this gap. However, there are limited data on the effectiveness of the clinical process and cost-effectiveness of most telemedicine applications.

Objective: This study examined the implementation of the first Italian online, web-based, comprehensive screening tool and described the screening and diagnostic process through the interactive web platform in a child psychopathology clinic. This is a feasibility study that aims to present the design and implementation of the best practices to improve patient experiences and clinical outcomes. Moreover, the paper evaluates the platform with qualitative and quantitative measures.

Methods: We planned, designed, and implemented a web-based system to collect, store, and manage clinical data. The platform was developed by a multidisciplinary team composed of researchers, clinicians, and informatics professionals through different steps. First, we defined the clinical information to be collected. A number of measures were chosen, tapping several clinical risk areas such as neurodevelopmental disorders and emotional and behavioral problems. The web application architecture and process were then designed. The three phases of process design are described in detail: design of the input interface, processing design, and design of the output interface. Finally, the system has been implemented and evaluated. Based on indicators recommended by the National Quality Forum and the Italian National Guidelines, we evaluated the quality of the system and used quantitative measures that were replicable and comparable over time.

Results: We present the implemented architecture and features of Medea Information and Clinical Assessment On-Line (MedicalBIT), and we provide performance measures for the data collected between October 2018 and June 2021. The measured concepts pertain to four domains: access to care, financial impact/cost, experience, and effectiveness.

Conclusions: In this study, we present the successful implementation of an innovative digital tool. The findings of this study show that the implemented web-based platform appears to be an efficient, cost-effective, and feasible way to improve digital care in the field of child psychiatry.

(JMIR Form Res 2022;6(7):e36757) doi:10.2196/36757

KEYWORDS

digital health; big data; developmental psychopathology; neurodevelopmental disorders; digital data; digital innovation; mental health; screening tool; children; psychopathology; web platform; digital intervention; clinical outcome

Introduction

Digital data collection is an emerging trend in various fields, including the medical and psychological ones. Digital innovation is changing the way health information is collected, shared, and analyzed for better clinical decision-making and health care. Digital innovation is also rapidly expanding in the medical field, significantly improving the quality of health care, reducing health care costs, and enhancing research processes [1]. The rapid evolution of technology has recently promoted web platforms for big data collection in the health care field, and related publications are exponentially increasing [2].

"Big data" refers to a voluminous collection of information taking place quickly without affecting quality, and there are several well-known definitions describing this [3-5]. Pastorino and colleagues [2] stress the importance of defining data as both smart and big because big data presents a substantial potential when it is meaningful. This implies using data for improving health conditions by searching for increasingly clearer and more accurate links between causes, diseases, therapies, and outcomes. In the Study on Big Data in Public Health, Telemedicine, and Healthcare [6], the European Commission identified four macroareas in health care for big data use: (1) early signs for detection, diagnosis, and intervention; (2) identification of risk factors for diseases to improve prevention; (3) enhancing pharmacovigilance and patient safety by communication of real-time information; and (4) improvement in outcome prediction.

All this is possible since a large amount of data is an invaluable resource for epidemiological studies, analyses of general population needs, treatment evaluation, and experimental designs on the target population. If research develops in this direction, it will enable precision medicine that will contribute to optimization of resources: the right care for the right patient at the right time. Therefore, smart use of big data can provide a possible answer to the need for care: socioeconomic and clinical sustainability. In line with this, health technology assessments aim to inform on safe, effective, patient-centered policy-making as well as determine the greatest value for it.

Against this background, telehealth [7] should be seen not only as a complement or alternative to traditional clinical practice but also as an ideal channel to collect and analyze big data. Besides rapid technological progress, COVID-19 spurred an exponential growth in telehealth, and there will be no return to the prepandemic situation according to the American Telemedicine Association [8]. There are several fields of telehealth, just like there are medical specialties. For example, investments in information technology are promoting the development of telepsychiatry. According to Allen [9] and Torous and colleagues [10] one of the most important contributions of artificial intelligence for psychiatry concerns apps, as shown by their innovative project Learn, Assess, Manage, and Prevent (MindLAMP). MindLAMP is a digital platform used both for clinic and research purposes: it receives patient input and aggregates data, it guides reflection, and it helps orient toward treatments. Through the app, patients can input data in an active way (survey, symptom registration, etc) or in a passive way (information collected in the background even if the user is not using the app), and they can receive some advice and mindfulness resources [11]. It is evident that, in the last decade, there has been an increased interest in telemedicine by mental health agencies; the gap between the need for care and possible answers both for diagnosis and treatment is wide, and digital technology could play a fundamental role [12]. In the child and adolescence mental health services, the parents and caregivers of young patients must be involved in the clinical diagnostic process, both as powerful sources of information to be used by clinicians and to obtain a clear understanding about their child's difficulties. This engagement should be conducted in a comfortable environment for the participants.

Telemedicine seems to be efficient both for the assessment and treatment. For example, among services for the diagnosis of autism spectrum disorder (ASD), the use of telemedicine has shown some promising results in terms of observation [13-15]. As for remote treatment, several recent studies on telemedicine and developmental disorders such as autistic spectrum disorder, specific learning disorder, specific language impairment, dyspraxia, and acquired brain injuries show encouraging results [16-20].

A review of the literature by the National Quality Forum—a no-profit organization for health care development—found a positive effect of telehealth on the quality of processes, outcomes, and costs. It also identified the medical areas where telehealth spread more extensively, with a correspondingly greater increase in scientific publications, namely, dermatology, mental health, rehabilitation, medical management, and chronic diseases. However, despite this topic being current and debated, telemedicine is relatively recent, and there is a lack of studies assessing its quality [21].

In our study, we examined the implementation of the first Italian online, web-based, comprehensive screening tool: Medea Information and Clinical Assessment On-Line (MedicalBIT) [22]. We carried out a feasibility study aimed to present the design and the implementation of the best practices to improve users' experiences and discuss the evaluation of the platform with qualitative and quantitative measures. Additionally, we intend to describe a web-based screening and diagnostic process in a child psychopathology clinic.

Methods

Overview

This section describes the key elements of the platform implementation.

We planned, designed, and implemented a web-based system to collect, store, and manage clinical data through different steps. The whole eHealth system is composed of users (patients and clinicians), a service provider (the Association La Nostra Famiglia-IRCCS Eugenio Medea, a no-profit organization providing care and rehabilitation to children with disabilities), and a service developer (SE-GE Consulting Company).

The main purpose of the MedicalBIT platform is to support diagnostic flow by systematic data collection.

Patients can answer questionnaires comfortably from home through a user-friendly and easily accessible interface, saving time traveling to and staying in a clinic. Clinicians have access to a real-time description of individual patients' symptoms and a graphical output of possible alerts through clinically based algorithms. Clinicians then assign patients to different diagnostic paths for principal neurodevelopmental disorders (Attention-Deficit/Hyperactivity Disorder, Autism Spectrum Disorder, Specific Learning Disorders, Language Disorders) and behavioral and emotional disorders based on this output. Taking into account the international gold standard, a clinical assessment is performed for each path. The final diagnostic output may differ from the initial path assignment based on the assessment results.

Finally, the data collected can be easily exported in Excel (Microsoft Corporation) format for research analyses: researchers can use collected data for further analysis to develop predictive models.

The platform was developed by a multidisciplinary team composed of researchers, clinicians, and informatics professionals. Figure 1 provides an overview of the workflow, and Figure 2 displays some picture frames of the web platform home page. A description of each step is reported below:

1. Defining clinical information to be collected
2. Web application architecture and process design
3. System implementation and evaluation

Figure 1. The diagnostic flow on Medea Information and Clinical Assessment On-Line (MedicalBIT).

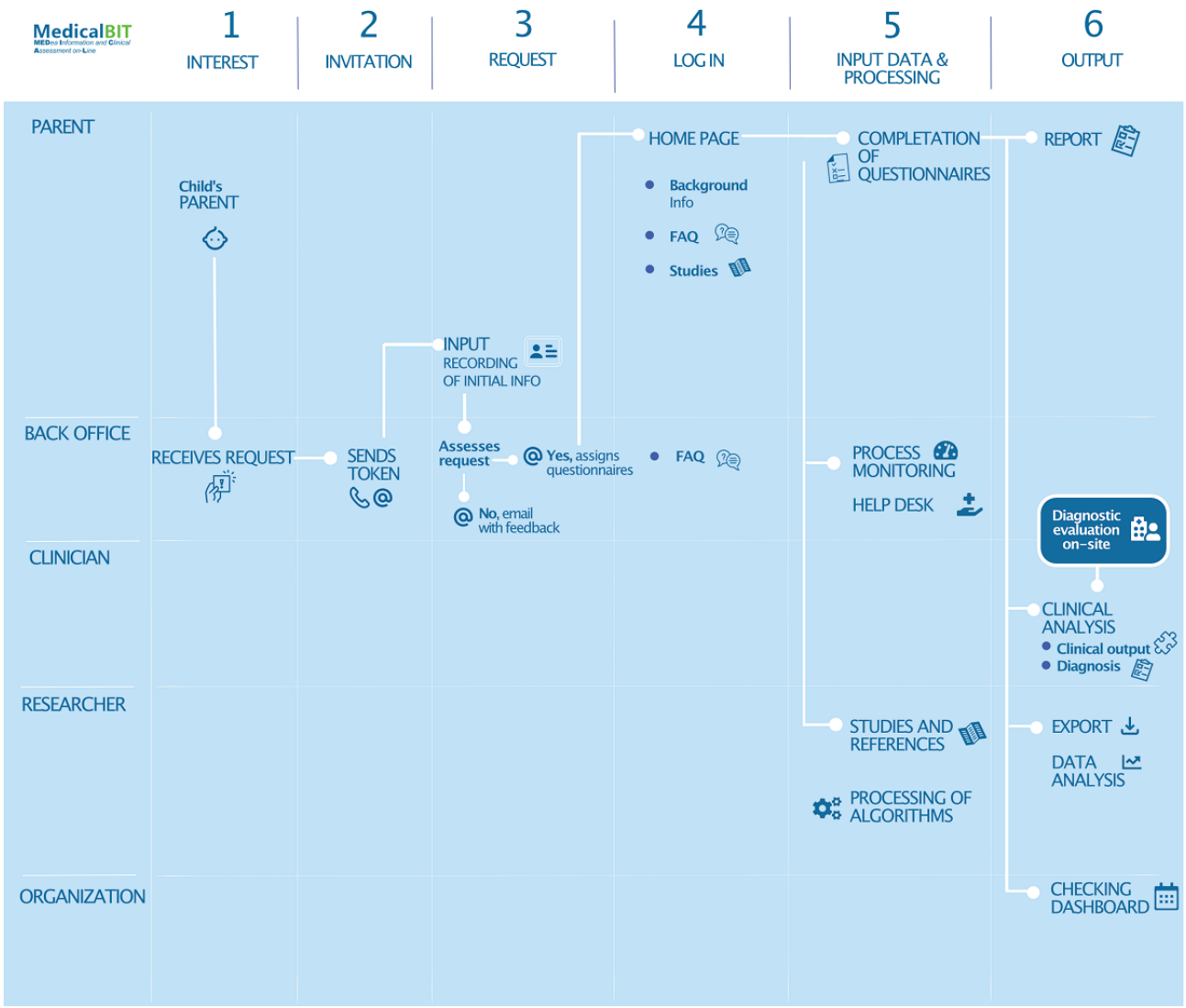
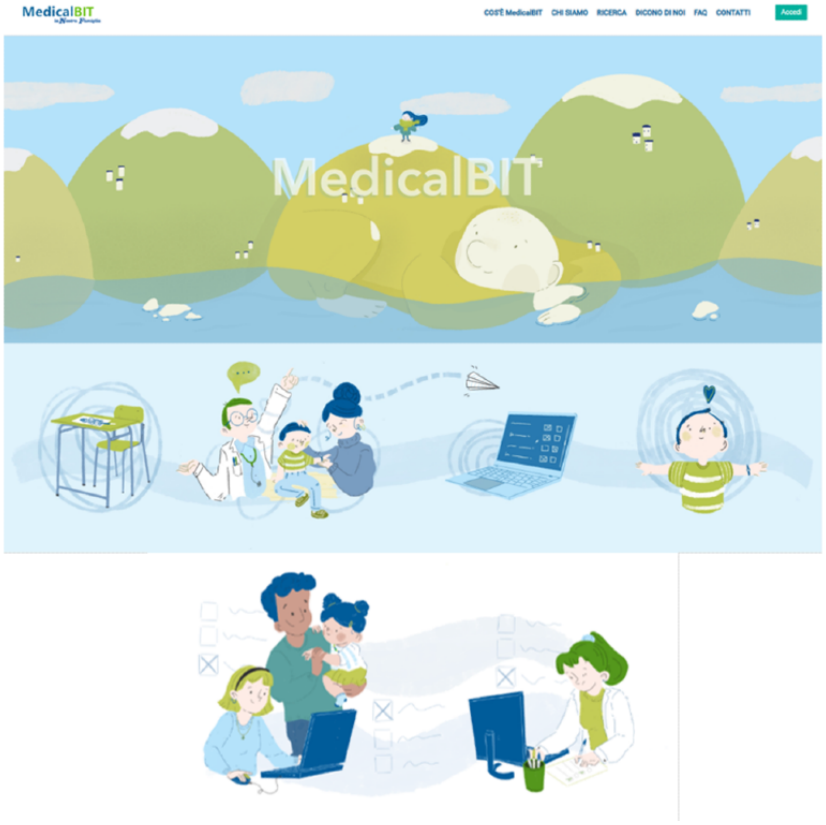


Figure 2. The main menu and some pictures of the Medea Information and Clinical Assessment On-Line (MedicalBIT) home page.



Defining Clinical Information to Be Collected

A number of measures were chosen by tapping several clinical risk areas, such as neurodevelopmental disorders and emotional and behavioral problems. Measures were selected for their feasibility in routine clinical practice (ie, brevity, free availability, validation in children and young people, and translation) and psychometric performance (ie, validity, reliability, and sensitivity to change).

The following questionnaires were selected:

- Risk factors: A questionnaire to explore biological and environmental risk factors such as family composition; presence/absence of psychiatric diseases in parents or close relatives; prenatal, perinatal, and postnatal factors; and developmental milestones.
- ASD: The Modified Checklist for Autism in Toddlers [23] and the Autism Spectrum Quotient: Children's Version [24] were used for the detection of ASD. The first is one of the most widely used ASD toddler screening instruments; it is easily accessible and low-cost. The second is a brief, parent-reported, 50-item questionnaire to quantify autistic traits in children aged 4-11 years.
- Emotional and behavioral problems: The Strengths and Difficulties Questionnaire [25] is a brief instrument widely used to assess main areas of developmental psychopathology and personal strengths. It consists of 25 items and is available in three forms depending on responders: parents, teachers, and adolescents (self-report).
- Other neurodevelopmental disorders: Ad hoc screening tools for specific language impairment and specific learning disorder were implemented by our institute's research group working on specific learning, language, and communication disorders.

Overall, these tools cover ages ranging from 18 months to 15 years.

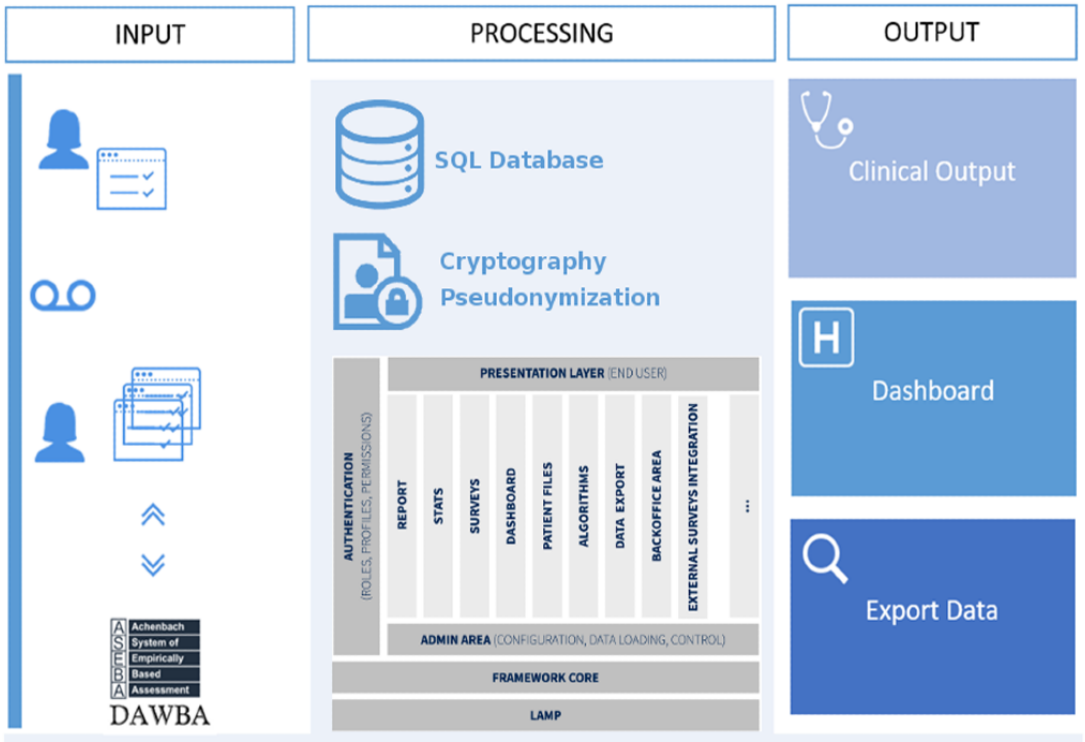
After this initial screening step, other standardized assessment tools can be administered too, such as the Child Behavior Checklist (CBCL) [26,27] and the Development and Well-Being Assessment (DAWBA) [28]. Through an interoperability system, links and credentials can be sent to caregivers to complete questionnaires (CBCL thorough ASEBA-Web) and the interview (DAWBA thorough Dawba.net), and results can then be imported by uploading appropriately encrypted files.

Web Application Architecture and Process Design

Figure 3 shows a high-level diagram of the platform development process. The three phases of process design are described in detail below:

- Design of the input interface
- Processing design
- Design of the output interface

Figure 3. The developed web app architecture design.



Design of the Input Interface

An easy-to-use interface was developed to support the interaction between front-end users and the web-based application; the web interface is simple to learn and easy to use, and it displays information in a consistent and progressive manner and maximizes functionalities. To be displayed properly on most mobile devices, the platform implements a responsive web design.

The system also relies on the active engagement of caregivers who send their data to the informatics system using a PC and mobile device. Patients scheduled for an appointment at the Child Psychopathology Unit (IRCCS Medea) get an email with a token for the first web app. They are then guided through a series of steps: completion of a brief form with primary clinical information, verifying of this information, and filling out of the registration form. They then receive a set of questionnaires according to their age and main characteristics, along with the link to the MedicalBIT assessment for completion.

Users are not asked to complete all forms in one session, they can resume from where they last left. After completing all the selected questionnaires, patients are scheduled for a doctor's visit.

To support the user's interaction with the platform, a user-friendly back-end interface was implemented. Through a control panel, staff can handle main functions such as administration rights, patients' lists, and support for data entry.

Processing Design: Technical Implementation of Data Storage and Automated Scoring Algorithm

At the end of each survey, the system immediately processes the caregiver's answers using artificial intelligence algorithms. Starting from scores provided for every possible answer and relying on appropriate threshold values, alerts, and key performance indicators are generated, which are useful to health professionals not only to monitor single patients but also for statistics and research purposes.

The whole process is General Data Protection Regulation compliant. The web portal is protected with a digital certificate issued by a certification authority, implements HTTPS (the secure browsing protocol for the World Wide Web), and is hosted on a virtual server in the GARR cloud environment. GARR is the Italian ultrabroadband network for education and research.

The infrastructure provides high reliability, service continuity, and data protection. Backup takes place daily, and we rely on disaster recovery as a service in an alternative environment to ensure availability in case of downtime. We used open source solutions for easier integration with company applications and to avoid vendor lock-in. Our platform is compliant with Italian guidelines [29]: "To ensure an effective support to all company processes, it is necessary to guarantee data univocity and integrity, real-time updating, historicization and audit trail, ergonomics, standardization, integration, stability, availability, security, privacy, innovation, evolvability."

All data are unique; nothing is duplicated. Information can be accessed only by authorized users and is protected from unauthorized changes; furthermore, confidential data is safeguarded.

There are two different access levels:

- Front-end user
  - Caregivers filling out the forms can only view and download the information concerning their child. They have no access to scores generated by the automatic

scoring algorithms. Multiple patients can be associated to each caregiver (eg, in the case of siblings).

- Back-end users
  - Health care professionals viewing the answers and any alerts enter the diagnoses at the end of the diagnostic process.
  - Back-office users with full operation rights
  - The system administrator has access to additional features such as dashboard and data export (see next paragraph).

The integrity of data and documents is guaranteed by the use of appropriate logs of activity and changes. These logs are also useful from a legal point of view. All information relating to a particular patient can also be deleted by users with administrator rights. Documents saved on the server are protected by encryption; to comply with the privacy legislation, the information stored in the database is pseudo-anonymized.

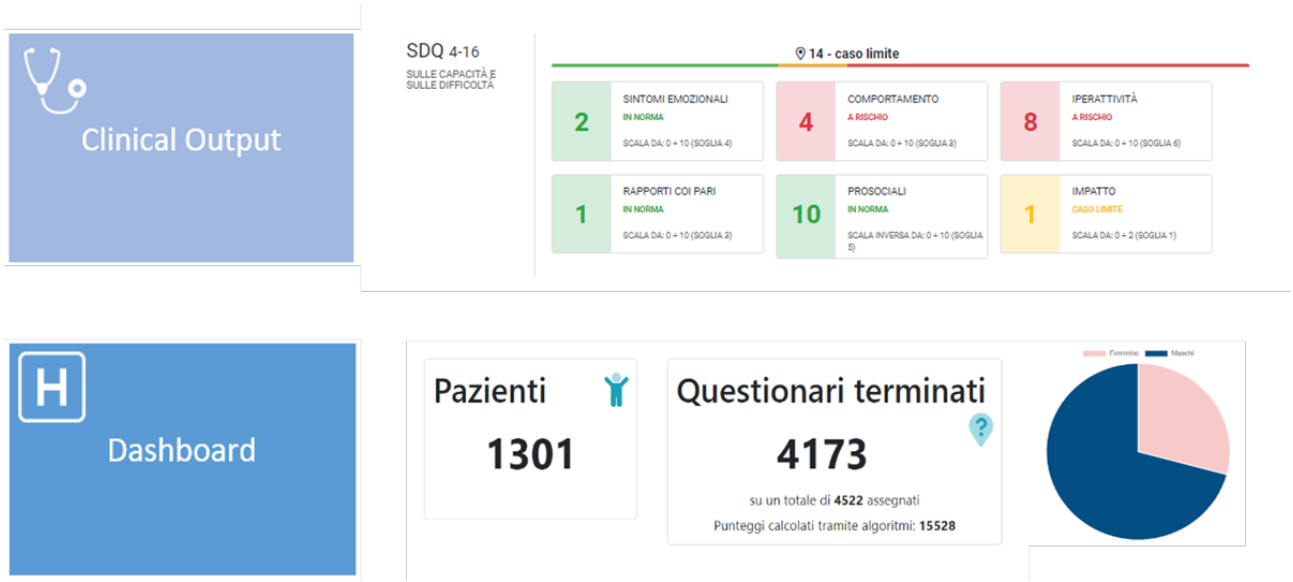
Design of the Output Interface

A real-time graphing, panel, and table interface is available only to back-end users and contains all processed data in real time. This data is also displayed in interactive dashboards available at different access levels. Two different types of reporting are created to display the collected data (Figure 4):

1. Clinical output: an individual-level output, reporting patient outcome scores
2. Dashboard: group-level data, reporting summarized data in a dashboard

All data stored in the database system can be exported in Excel format for further processing.

Figure 4. Different output levels.



System Implementation and Evaluation

This digital system has been on since October 2018 and has been used by 1301 patients (front-end users) and back-end users, including 17 clinicians, 2 back-office operators, and administration operators.

Back-end users are employed by the provider service: back-office operators are office secretaries and administration operators are psychologists and researchers who designed the service in collaboration with the data scientist and consulting company SEGE srl.

Based on indicators recommended by the National Quality Forum [21] and the Italian National Guidelines [29], we evaluated the quality of the system and used quantitative measures that were replicable and comparable over time. As shown in Textbox 1, the selected measures pertain to four areas: access to care, financial impact/cost, experience, and effectiveness.

Textbox 1. Telehealth measurement framework domains and subdomains.

Access to care

- Access for parent/caregiver
- Access to information

Effectiveness

- System effectiveness
- Operational effectiveness
- Technical effectiveness

Costs

- The financial impact to family/caregiver
- The financial impact to care team

Experience

- Patient, family, or caregiver experience

Results

According to the Italian National Guidelines [29] and the National Quality Forum [21], we provide performance measures available for data collected between October 2018 and June 2021 (Table 1 and Figure 5). This process is a work in progress, and in the future, we aim to capture more indicators about different domains and subdomains.

Table 1. Platform performance measures.

Domains and subdomains	Qualitative measures	Quantitative measures
Access to care		
• Access for parent/caregiver • Access to information	• The web system is responsive to different devices • Family's patient can access the system using dedicated access points in the institute • Clear instructions in the home page • Responsive technical assistance	• N/A[a]
Effectiveness		
• System effectiveness	• N/A	• Continuity: from October 2018 to June 2021 • Dimension: 1100 new patients accessed the telehealth system and were subsequently sorted to different diagnostic paths. For details about the accesses trend over time, see indicators in Figure 5. Each patient completed on average 4 questionnaires, for a total of 177 items. • Speed: (1) Only a few days elapsed between the initial request from family and the completion of screening questionnaires on the platform. (2) The average length between screening questionnaire completion and completion of the diagnostic process was 106 days, with a decrease over the 3-year period from 2019 to 2021 (see Figure 5e)
• Operational effectiveness	• The system is perfectly integrated within the traditional diagnostic path. • Customized in-person visits are based on web questionnaire output.	• N/A
• Technical effectiveness	• Output dashboard for clinicians and database with collected data for researchers are automatically generated. • The system is integrated with other screening and diagnostic tools implemented on different platforms: the Development and Well-Being Assessment [28] and the Child Behavior Checklist [26,27].	• N/A
Costs		
• Financial impact for family/caregiver • Financial impact for care team	• Decrease in the length/frequency of stay/visit to hospital • Clinician can integrate traditional work setting with smart working	• N/A
Experience		
• Patient, family, or caregiver experience	• N/A	• Caregivers dropout: (1) 5.3% of registered users do not fill in any questionnaire. (2) 99.6% of caregivers that begin a questionnaire, complete it

[a]N/A: not applicable.

Figure 5. (a) Number of accesses of new patients for each semester in the last 3 years: multiple accesses by 1 patient are counted once. (b) Average and SD per month of accesses by new patients in the last 3 years. (c) Dynamic dimension value = (August 2020 to July 2021) / (August 2019 to July 2020) = 560 / 410 = 1.36. This enabled us to compare differences in accesses in the last 2 years. (d) The graph shows the trend in Medea Information and Clinical Assessment On-Line (MedicalBIT) during the different phases of the pandemic. (e) Data show that the average time between the completion of questionnaires and receiving a diagnosis has clearly decreased over the 3-year period: 2019 (mean 150, SD 179.6), 2020 (mean 82, SD 77.3), and 2021 (mean 65, SD 34.9).

Discussion

Principal Results

In this study, we present the design and implementation of the MedicalBIT web platform, which allows patients to complete questionnaires comfortably from home through a user-friendly and easily accessible interface.

This work is based on the European Commission's recommendations in the Study on Big Data in Public Health, Telemedicine, and Healthcare [6] that identified four major areas in health care where big data may be used: (1) early signs for detection, diagnosis, and intervention; (2) identification of risk factors for disease to improve prevention care; (3) enhancing pharmacovigilance and patient safety by communication of real-time information; and (4) improvement in outcomes prediction.

The MedicalBIT platform, as described in this paper, allows for the timely collection of clinical data to support the diagnostic process and subsequent steps (1 and 3), and it enables the development of predictive models to improve preventive care and outcome prediction (2 and 4). A core component of the proposed model is caregivers' perspective of their child's mental health, laying the foundation for evidence-based practice and person-centered care. The MedicalBIT web platform was introduced into the clinical workflow, making patients' and caregivers' care experiences more comfortable. As mentioned, health information technology can save time and reduce costs, thereby improving the health care experience for both patients and clinicians. For this purpose, we also collected customer satisfaction feedback.

Furthermore, performance measures enable us to describe the platform with an objective measurement that allows for comparisons both within the same platform over time and between different web platforms used for similar purposes.

According to reports so far, available quantitative data shows general upward trends for platform use from 2018 to 2021 despite the apparent drop during the first months of the COVID-19 pandemic, which required a complete reorganization of care provision.

This trend is in line with research confirming the value of telepsychiatry [30,31] not only during the pandemic period but also before and, most importantly, after the emergency. As described in the Introduction, use of telemedicine in the psychiatry and developmental psychiatry fields can offer several benefits such as improved efficiency and effectiveness of mental health services: it can reach more people with fewer resources [32].

We hope that these changes will lead to filling the care gap in the field of child psychiatry, which has been appropriately defined as “one of the most difficult and crucial challenges of the next decade” [12].

Our project fits into this still uncertain and experimental scenario, with its pros and cons. Telemedicine in the child psychiatry clinical routine can help both patients and health care providers save time and provide them easy access to the care system. Furthermore, because of its fast and advantageous features, it facilitates the hospital and clinic workflow, as processing and automatic output of clinical data enables the collection of large amounts of data for research purposes. Nevertheless, there are some limitations to consider such as the impact of physical distance on human relationships and ethical and coroner issues [29].

Further studies are clearly necessary to establish evidence-based telepsychiatry-specific standards of care, following guidelines proposed by the American Telemedicine Association [33] and the American Academy of Child and Adolescent Psychiatry [34].

Limitations and Further Development

Based on the current limitations, we are working on a customer satisfaction questionnaire for both families and clinicians to collect feedback to make users feel more engaged in this process and help us as back-office operators to improve services. Moreover, we intend to improve our platform performance measurement tools to obtain more quantitative indicators about the four domains presented in Table 1.

According to Waller and Stotler [35], one of the next steps should be the implementation of a measurement and methodology evaluation of the impact of telehealth on clinical outcomes (in our case, on the diagnostic outcome as an indicator of clinical effectiveness). Our intent is to also implement a platform-based service that could follow and provide assistance to patients after diagnosis (information, psychoeducational materials, rehabilitation tools).

Conclusions

This study describes the successful implementation of an innovative digital tool. According to our results, the web-based platform appears to be a feasible, efficient, and cost-effective method for enhancing digital care in the field of child psychiatry. It also contributes to the collection of big and smart data, in line with European guidelines. This work shows how large data sets may enhance the accuracy and timing of diagnosis, and contribute to more effective interventions based on predictive models.

Acknowledgments

The authors wish to thank clinicians and back-office staff who took part in the web platform co-design and all clinicians using Medea Information and Clinical Assessment On-Line (MedicalBIT). This work has been possible due to the patients’ families, clinicians, researchers, and back-office staff who trust this innovative system.

Authors' Contributions

PC, SBC, and MM conceptualized the study. SBC gathered the resources and curated the data. SBC, SP, and PC prepared the original draft. PC, SBC, SP, and SC reviewed and edited the manuscript. MM supervised the study.

Conflicts of Interest

None declared.

References

1. Monitoring and evaluating digital health interventions: a practical guide to conducting research and assessment. World Health Organization. 2016. URL: https://apps.who.int/iris/bitstream/handle/10665/252183/?sequence=1 [accessed 2022-06-27]
2. Pastorino R, De Vito C, Migliara G, Glocker K, Binenbaum I, Ricciardi W, et al. Benefits and challenges of Big Data in healthcare: an overview of the European initiatives. Eur J Public Health 2019 Oct 01;29(Supplement_3):23-27 [FREE Full text] [doi: 10.1093/eurpub/ckz168] [Medline: 31738444]
3. Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, et al. Big Data: the Next Frontier for Innovation, Competition, and Productivity. Washington, DC: McKinsey Global Institute; 2011.
4. Beyer M, Laney D. The importance of 'Big Data': a definition. Gartner. 2012. URL: https://www.gartner.com/en/documents/2057415#:~:text=%22Big%20data%22%20warrants%20innovative%20processing,to%20business%20goals%20and%20objectives [accessed 2022-06-27]

5. McAfee A, Brynjolfsson E. Big data: the management revolution. Harv Bus Rev 2012 Oct;90(10):60-6, 68, 128. [Medline: 23074865]
6. Study on Big Data in Public Health, Telemedine and Healthcare: final report. European Commission. 2016. URL: https://ec.europa.eu/health/sites/health/files/ehealth/docs/bigdata_report_en.pdf%0A%0A [accessed 2021-09-01]
7. Telehealth: defining 21st century care. American Telemedicine Association. 2020. URL: https://marketing.americantelemed.org/hubfs/Files/Resources/ATA_Telehealth_Taxonomy_9-11-20.pdf [accessed 2022-06-27]
8. The adoption of telehealth. American Telemedicine Association. 2021. URL: https://www.americantelemed.org/wp-content/uploads/2021/05/Adoption-of-Telehealth.pdf [accessed 2022-06-27]
9. Allen S. Artificial intelligence and the future of psychiatry. IEEE Pulse 2020;11(3):2-6. [doi: 10.1109/MPULS.2020.2993657] [Medline: 32559160]
10. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. JMIR Ment Health 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: 10.2196/18848] [Medline: 32213476]
11. Torous J, Wisniewski H, Bird B, Carpenter E, David G, Elejalde E, et al. Creating a digital health smartphone app and digital phenotyping platform for mental health and diverse healthcare needs: an interdisciplinary and collaborative approach. J Technol Behav Sci 2019 Apr 27;4(2):73-85. [doi: 10.1007/s41347-019-00095-w]
12. Fatori D, Polanczyk GV. The role of digital technology in bridging the child mental health care gap. Eur Child Adolesc Psychiatry 2019 Apr;28(4):425-426. [doi: 10.1007/s00787-019-01316-6] [Medline: 30888504]
13. Corona L, Hine J, Nicholson A, Stone C, Swanson A, Wade J, et al. What is the TELE-ASD-PEDS? Vanderbilt Kennedy Center. 2020. URL: https://vkc.vumc.org/vkc/triad/tele-asd-peds [accessed 2022-03-25]
14. Corona L, Wagner L, Wade J, Weitlauf A, Hine J, Nicholson A, et al. Toward novel tools for autism identification: fusing computational and clinical expertise. J Autism Dev Disord 2021 Nov;51(11):4003-4012 [FREE Full text] [doi: 10.1007/s10803-020-04857-x] [Medline: 33417138]
15. Wagner L, Weitlauf AS, Hine J, Corona LL, Berman AF, Nicholson A, et al. Transitioning to telemedicine during COVID-19: impact on perceptions and use of telemedicine procedures for the diagnosis of autism in toddlers. J Autism Dev Disord 2022 May;52(5):2247-2257 [FREE Full text] [doi: 10.1007/s10803-021-05112-7] [Medline: 34085153]
16. Sutherland R, Trembath D, Roberts J. Telehealth and autism: a systematic search and review of the literature. Int J Speech Lang Pathol 2018 Jun;20(3):324-336. [doi: 10.1080/17549507.2018.1465123] [Medline: 29709201]
17. Hodge MA, Sutherland R, Jeng K, Bale G, Batta P, Cambridge A, et al. Literacy assessment via telepractice is comparable to face-to-face assessment in children with reading difficulties living in rural Australia. Telemed J E Health 2019 Apr;25(4):279-287. [doi: 10.1089/tmj.2018.0049] [Medline: 30040538]
18. Wales D, Skinner L, Hayman M. The efficacy of telehealth-delivered speech and language intervention for primary school-age children: a systematic review. Int J Telerehabil 2017;9(1):55-70 [FREE Full text] [doi: 10.5195/ijt.2017.6219] [Medline: 28814995]
19. Miyahara M, Butson R, Cutfield R, Clarkson JE. A pilot study of family-focused tele-intervention for children with developmental coordination disorder: development and lessons learned. Telemed J E Health 2009 Sep;15(7):707-712. [doi: 10.1089/tmj.2009.0022] [Medline: 19694593]
20. Corti C, Oldrati V, Oprandi M, Ferrari E, Poggi G, Borgatti R, et al. Remote technology-based training programs for children with acquired brain injury: a systematic review and a meta-analytic exploration. Behav Neurol 2019;2019:1346987. [doi: 10.1155/2019/1346987] [Medline: 31467613]
21. Creating a framework to support measure development for telehealth. National Quality Forum. 2017. URL: https://www.qualityforum.org/Publications/2017/08/Creating_a_Framework_to_Support_Measure_Development_for_Telehealth.aspx [accessed 2021-09-01]
22. MedicalBIT. URL: https://www.medicalbit.com/it/home/ [accessed 2022-06-27]
23. Robins DL, Casagrande K, Barton M, Chen CA, Dumont-Mathieu T, Fein D. Validation of the modified checklist for Autism in toddlers, revised with follow-up (M-CHAT-R/F). Pediatrics 2014 Jan;133(1):37-45 [FREE Full text] [doi: 10.1542/peds.2013-1813] [Medline: 24366990]
24. Auyeung B, Baron-Cohen S, Wheelwright S, Allison C. The Autism Spectrum Quotient: Children's Version (AQ-Child). J Autism Dev Disord 2008 Aug;38(7):1230-1240. [doi: 10.1007/s10803-007-0504-z] [Medline: 18064550]
25. 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]
26. Achenbach T. Child Behavior Checklist (CBCL), Achenbach System of Empirically Based Assessment. Burlington, VT: ASEBA; 2001.
27. Achenbach T, Rescorla L. Manual for the ASEBA School-Age Forms and Profiles. Burlington, VT: University of Vermont; 2001.
28. Goodman R, Ford T, Richards H, Gatward R, Meltzer H. The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. J Child Psychol Psychiatry 2000 Jul;41(5):645-655. [Medline: 10946756]

29. Linee di indirizzo nazionali sulla telemedicina. Ministero della Salute. 2014. URL: https://www.salute.gov.it/portale/ehealth/dettaglioContenutiEHealth.jsp?lingua=italiano&id=5525&area=eHealth&menu=telemedicina [accessed 2022-06-27]
30. Chang BP, Kessler RC, Pincus HA, Nock MK. Digital approaches for mental health in the age of covid-19. BMJ 2020 Jun 29;369:m2541. [doi: 10.1136/bmj.m2541] [Medline: 32601049]
31. Garofalo M, Vaithilingam S, Ferrando S. Telemedicine for psychiatry and mental health. In: Latifi R, Doarn CR, Merrell RC, editors. Telemedicine, Telehealth and Telepresence: Principles, Strategies, Applications, and New Directions. Cham: Springer; 2021:365-378.
32. Moock J. Support from the internet for individuals with mental disorders: advantages and disadvantages of e-mental health service delivery. Front Public Health 2014;2:65. [doi: 10.3389/fpubh.2014.00065] [Medline: 24967221]
33. Yellowlees P, Shore J, Roberts L, American Telemedicine Association. Practice guidelines for videoconferencing-based telemental health - October 2009. Telemed J E Health 2010 Dec;16(10):1074-1089. [doi: 10.1089/tmj.2010.0148] [Medline: 21186991]
34. American Academy of Child and Adolescent Psychiatry (AACAP) Committee on Telepsychiatry and AACAP Committee on Quality Issues. Clinical update: telepsychiatry with children and adolescents. J Am Acad Child Adolesc Psychiatry 2017 Oct;56(10):875-893. [doi: 10.1016/j.jaac.2017.07.008] [Medline: 28942810]
35. Waller M, Stotler C. Telemedicine: a primer. Curr Allergy Asthma Rep 2018 Aug 25;18(10):54. [doi: 10.1007/s11882-018-0808-4] [Medline: 30145709]

Abbreviations

ASD: autism spectrum disorder
CBCL: Child Behavior Checklist
DAWBA: Development and Well-Being Assessment
MedicalBIT: Medea Information and Clinical Assessment On-Line
MindLAMP: Learn, Assess, Manage, and Prevent

Edited by A Mavragani; submitted 15.02.22; peer-reviewed by B Sindelar, L Cerniglia; comments to author 11.03.22; revised version received 06.05.22; accepted 22.05.22; published 11.07.22.

Please cite as:
Colombo P, Busti Ceccarelli S, Pacchiarini S, Cribellati S, Molteni M
A Web Platform for Standardized Data Acquisition, Processing, and Export in the Child Psychopathology Clinical Routine (MedicalBIT): Design and Implementation Study
JMIR Form Res 2022;6(7):e36757
URL: https://formative.jmir.org/2022/7/e36757
doi:10.2196/36757
PMID:35816380

Original Paper

Developing a Consensus for Adolescent and Young Adult mHealth HIV Prevention Interventions in the United States: A Delphi Method Study

Kayla Knowles[1*], MPH; Nadia Dowshen[1,2,3*], MSHP, MD; Susan Lee[1,2*], MPH; Amanda Tanner[4*], MPH, PhD

[1]PolicyLab, Children's Hospital of Philadelphia, Roberts Center for Pediatric Research, Philadelphia, PA, United States

[2]Craig-Dalsimer Division of Adolescent Medicine, Children's Hospital of Philadelphia, Philadelphia, PA, United States

[3]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[4]Department of Public Health Education, University of North Carolina Greensboro, Greensboro, NC, United States

[*]all authors contributed equally

Corresponding Author:
Kayla Knowles, MPH
PolicyLab
Children's Hospital of Philadelphia
Roberts Center for Pediatric Research
2716 South Street
10th Floor, #10-2222
Philadelphia, PA, 19146
United States
Phone: 1 267 425 1449
Email: knowlesk1@chop.edu

Abstract

Background: Engaging adolescents and young adults (AYAs) who are at elevated risk for HIV acquisition or who are living with HIV in health care has posed a major challenge in HIV prevention and care efforts. Mobile health (mHealth) interventions are a popular and accessible strategy to support AYA engagement despite barriers to care present along the HIV care continuum. Even with progress in the field of mHealth research, expert recommendations for the process of designing, evaluating, and implementing HIV-related mHealth interventions are underdeveloped.

Objective: The aim of this study was to compile expert recommendations on the development, evaluation, and implementation of AYA-focused HIV prevention and care mHealth interventions.

Methods: Experts from adolescent mHealth HIV research networks and investigators of recently funded HIV mHealth projects and programs were identified and invited to complete a series of electronic surveys related to the design, implementation, and evaluation of HIV-related mHealth interventions. A modified Delphi method was used to ask experts to score 35 survey items on a 4-point Likert scale from not important to very important and encouraged experts to leave additional comments in textboxes. Responses were reviewed by the researchers, a team of 4 HIV mHealth intervention experts. The average importance ratings from survey responses were calculated and then categorized as retained (high importance), flagged (mid-level importance), or dropped (no/low importance). Additionally, thematic analysis of expert comments helped modify survey items for the next survey round. An evaluation of the level of agreement among experts on the most important items followed each round until consensus was reached.

Results: Of the 35 invited experts, 23 completed the first survey representing a variety of roles within a research team. Following two rounds of Delphi surveys, experts scored 24 of the 28 (86%) survey items included in round two as important to very important. The final consensus items included 24 recommendations related to the mHealth intervention design process (n=15), evaluation (n=2), and implementation (n=7). The 3 survey items with the highest average scores focused on the design process, specifically, (1) the creation of a diverse team including researchers, app software developers, and youth representation; (2) the importance of AYA-focused content; and (3) the value of an iterative process. Additionally, experts highlighted the importance of establishing the best ways to collect data and the types of data for collection during the evaluation process as well as constructing a plan for participant technology disruption when implementing an mHealth intervention.

Conclusions: The modified Delphi method was a useful tool to convene experts to determine recommendations for AYA-focused HIV prevention and care mHealth interventions. These recommendations can inform future mHealth interventions. To ensure the acceptability, feasibility, and efficacy of these AYA HIV prevention interventions, the focus must be on the specific needs of AYAs by including representation of AYAs in the process, including consistent and relevant content, ensuring appropriate data is collected, and considering technology and health accessibility barriers.

(JMIR Form Res 2022;6(7):e25982) doi:10.2196/25982

KEYWORDS

HIV prevention and care; adolescents; mHealth; mobile apps; Delphi method; health application; HIV care; HIV prevention; health intervention; public health; health care; health accessibility; health technology; digital health; HIV

Introduction

In the United States, HIV is a major cause of morbidity, mortality, and social adversity among adolescents and young adults (AYAs) [1]. In 2019, individuals aged 13-24 years accounted for 21% of newly reported HIV infections [2]. AYAs have poor health outcomes across the HIV care continuum—diagnosis, linkage and retention in care, maintaining an antiretroviral therapy (ART) regimen, and achieving viral suppression [2]. Over 40% of youth living with HIV (YLH) in the United States are unaware of their status compared to 14% of adults [3]. Additionally, for YLH with a known HIV diagnosis, care engagement was greater than the overall population with HIV diagnosis, yet viral suppression was 60% [3], again fairing worse than their adult counterparts. Beyond age, disparities based on race, gender, and sexual orientation have created an environment where communities with the highest incidence of HIV infections have access to the fewest resources [2-4]. Young Black and Latinx gay, bisexual, other men who have sex with men (GBMSM), and transgender women are at the greatest risk of becoming infected with HIV in their lifetime [3]. Experiences of pervasive stigma and systemic racism in and outside health care systems prevent communities most vulnerable to HIV from accessing and engaging in the HIV care continuum [3,5]. High rates of sexual risk behaviors, low rates of HIV testing, and poor engagement in HIV care [2,3] have generated a high demand for culturally competent and accessible AYA-targeted HIV prevention and treatment interventions.

The US Department of Health and Human Services Ending the HIV Epidemic (EHE) initiative outlined strategies to reduce new HIV infections by 90% by 2030 [6]. Efficient implementation of HIV research into practice was an area of focus included in the National HIV/AIDS Strategy to move toward EHE, as well as prioritizing populations most affected by the HIV epidemic [7]. In 2018, more than 90% of US youth regardless of race, income, or parent's education level had access to a smartphone [8], and adolescents expressed a preference for HIV interventions on mobile technologies [9,10]. Hence, mobile health (mHealth) approaches have become increasingly common in engaging AYAs in HIV prevention and care efforts. Focus groups and usability testing studies with YLH have identified privacy, entertainment, social support, and educational content as desired features of mHealth HIV interventions [11-14]. To date, mHealth interventions have used a variety of app features to support engagement, including connecting with other YLH via in-app forums or message boards [15], daily or weekly SMS text message reminders [16-18], and educational games [10,19]. Many mHealth projects, such as weCare, MyPEEPS, and Guy2Guy, prioritize engagement of GBMSM and transgender women with a social support network, inclusive sex education, and culturally relevant content [15,20,21]. mHealth interventions have shown some improvement in outcomes of interest including ART adherence, viral suppression, medical appointment attendance, HIV testing, and perceived social support [10,15,19,20,22]. However, few of these interventions have been tested in large-scale randomized controlled trials, and fewer have been implemented into routine clinical care or community-based settings. Further, there are no existing recommendations for developing, testing, and disseminating mHealth HIV prevention or treatment interventions for AYAs. Accordingly, this study aimed to use a modified Delphi method [23,24] to elicit expert recommendations on the design, evaluation, and implementation of AYA-focused HIV prevention and care mHealth interventions. Study design refers to the road map used to collect and analyze data based on a specific question the researchers hope to answer [25]. An evaluation reflects on how the intervention is implemented and how effective it is over time [26]. An effective evaluation will allow an intervention to adapt over time. Finally, implementation is the adaption and delivery of evidence-based research findings into real-world practice to benefit the community that the intervention was developed to benefit [27].

Methods

Overview

The Delphi method is based on the convergence of expert opinion after multiple rounds of data collection (eg, surveys with Likert scale items) with the final round occurring when agreement emerges among experts [23]. The multiple rounds of data collection and analysis in the Delphi method allow for elimination of survey items that experts agree are unimportant and refinement of the remaining items [23]. The final round leaves a consensus list of recommendations experts agree are crucial to consider. This study used a modified Delphi method [24] to reach consensus on expert recommendations for AYA HIV mHealth interventions. The modified Delphi method used in this study maintained the multiple rounds of data collection, but it omitted the expert-to-expert face-to-face discussion that would occur following data collection due to difficulties in varying schedules and time zones as well as to limit experts from influencing each other. In the absence of the expert

discussion, participants had the option to leave text comments in the survey that were later considered when designing the next survey data collection round.

Study Design

Survey Development

Survey items were generated by 4 mHealth HIV experts, who were investigators for various federally funded projects relevant to this topic. A list of common considerations when developing, evaluating, and implementing HIV mHealth interventions were compiled for the survey. The study team and the survey development group reviewed and edited the initial survey to ensure items were concise and minimized influence on expert participants' scoring.

The consensus survey for the first round included 35 Likert items sorted into five different categories: (1) design: preparation phase; (2) design: core elements; (3) evaluation: preparation and study design; (4) evaluation: data sources; and (5) implementation. Item response options ranged from not important (1) to very important (4), and each item included a "comment" textbox for experts to type in general comments (eg, why they scored an item the way they did) or revised wording suggestions. The survey was programmed in Research Electronic Data Capture (REDCap), a secure website software used for data collection and management in clinical and translational research [28,29], which was used to disseminate the survey to expert participants.

Expert Identification and Selection

The identification of experts who have knowledge of and experience with AYA mHealth HIV prevention and care interventions to participate in this study involved a review of large-scale HIV mHealth intervention networks and recently funded HIV mHealth projects and programs at the start of this study in 2019. Expert participants were identified through two different adolescent mHealth HIV-related networks, the iTech U19 of the National Institutes of Health (NIH)–funded Adolescent Medicine Trials Network [30] and the Health Resources and Services Administration's Special Projects of National Significance Initiative in Use of Social Media to Improve Engagement, Retention, and Health Outcomes along the HIV Care Continuum, 2015-2019 [31]. Additional experts were identified through a search of 2019 grantees of adolescent HIV mHealth interventions in the NIH Research Portfolio Online Reporting Tools Expenditures and Results [32] with keyword Medical Subject Headings (MeSH) [33] terms "adolescent," "HIV," and "mobile health." Expert identification and selection for survey participation aimed to encompass the involvement of a variety of cities and organizations in the United States and to include representation of principal investigators, program directors, project coordinators, protocol chairs, and project staff to capture the range of perspectives and expertise within a study team.

Survey Distribution and Data Collection

The survey was distributed via email with a weblink to participate in the REDCap survey. Experts had a 4-week window to complete the survey. After the round 1 survey window closed, the data were deidentified and analyzed. Survey items were modified based on expert participant survey response scores and accompanying comments resulting in a subsequent round 2 survey. Given the number of invited experts and consideration of their daily responsibilities, instead of having experts review each other's responses in their entirety, experts were informed of the group's overall responses by presenting them only with retained and reworded survey items in the next round for efficiency and privacy. A subsequent round was conducted similarly until a consensus was reached by participants. Only participants from the first round were eligible to participate in the subsequent round. Ultimately, the process was completed in two survey rounds between November 2019 and April 2020. Following successful completion of participation, expert participants were compensated in the form of a US $50 e–gift card for their time and effort in this study.

Data Analysis and Consensus Criteria

Deidentified survey Likert responses were exported from REDCap for analysis in Excel (Microsoft Corporation). Scores for each survey responses were determined based on the mean importance rating from not important (1) to very important (4) and were then categorized as retained, flagged, or dropped. Retained survey responses were those with mean scores between 3 (important) and 4 (very important), and automatically appeared on the next iteration of the survey. Flagged survey responses had a score between 2 (somewhat important) and 3 (important). For flagged survey responses, the SDs were calculated. The survey responses with an SD>1 were dropped, and the other flagged survey responses were revised based on the experts' comments. Finally, survey responses with a score ≤2 (not at all important, 1, to somewhat important, 2) were dropped.

As part of the REDCap survey, expert comments were collected in textboxes, which provided additional information on why they chose a particular survey response, and were reviewed during analysis [34]. Comments were compiled in Word (Microsoft Word) under their survey item, and responses with similar explanations/questions were grouped together for review. Comments were used to inform revisions following round 1 and were especially helpful for reviewing the flagged items, considering where divergence reflected disagreement among experts or unclear wording. Comments were also used to assist in assessing consensus among experts. The revised recommendation list presented to experts were those assessed as "important" (3) to "very important" (4) following the round 2 survey.

Ethical Considerations

This study received an exempt determination from the Children's Hospital of Philadelphia's Institutional Review Board. The beginning of each survey included an overview of the study and consent language where moving to the next page of the survey indicated willingness to participate in the study.

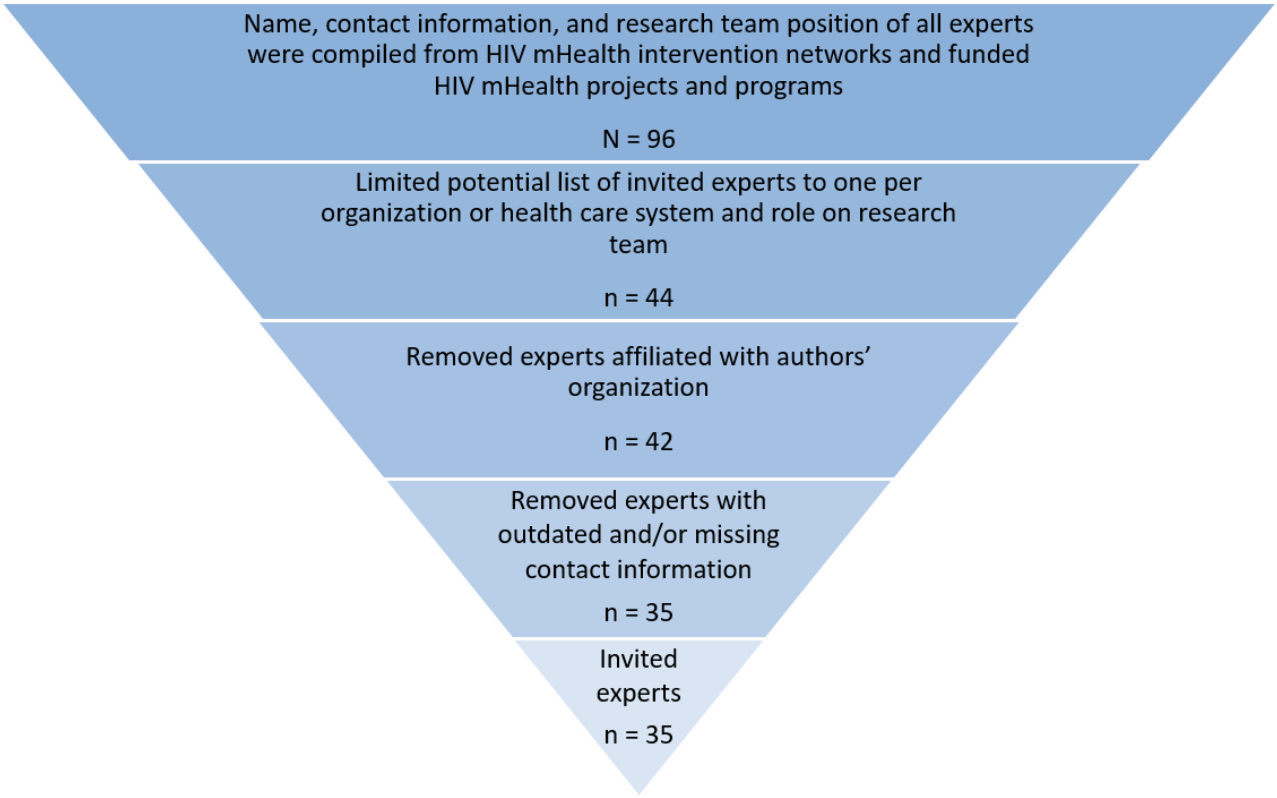
Results

Expert Participation

Survey responses were received from 23 of the 35 (66%) invited adolescent HIV/mHealth experts for round 1 (see Figure 1)

identified using reports from HIV mHealth intervention networks and recently funded HIV mHealth projects and programs. Participating experts represented 20 research organizations across the United States. More than half were principal investigators (n=15) and the remaining experts included various roles on a study team, including project directors (n=3), evaluators (n=2), project coordinators (n=2), and other project staff (n=1). Experts who completed the first survey were reinvited to participate in round 2 with a high response rate of 21 of 23 (91%) experts.

Figure 1. Expert identification and selection.



Round 1

Item Score Analysis

Of the 35 items asked in the round 1 survey, 86% (n=30) fit the criteria to be retained, 14% (n=5) were identified as flagged, and none fit the dropped criteria (not important, 1, to somewhat important, 2) following the analysis of the survey responses. Table 1 displays the expert responses for each section: total retained, flagged by SD, and dropped item. Of the 5 flagged items, 1 had an SD>1 and subsequently was kept for round 2.

Table 1. Consensus survey response scoring results: round 1 (R1; n=35) and round 2 (R2; n=28).

	Design: preparation	Design: core intervention elements	Evaluation: preparation and study design	Evaluation: data source	Evaluation: preparation, study design, and data source	Implementation	Total, n
Total, n							
R1	13	7	4	4	N/A[a]	7	35
R2	11	7	N/A	N/A	3	7	28
Mean (SD)							
R1	3.34 (0.39)	3.60 (0.28)	3.17 (0.28)	3.19 (0.20)	N/A	3.43 (0.20)	N/A
R2	3.41 (0.35)	3.53 (0.31)	N/A	N/A	3.11 (0.32)	3.41 (0.17)	N/A
Retain, n							
R1	10	7	3	3	N/A	7	30
R2	9	6	N/A	N/A	2	7	24
Flag, n[b]							
R1							
SD>1	1	0	0	0	N/A	0	1
SD≤1	2	0	1	1	N/A	0	4
R2							
SD>1	0	0	N/A	N/A	0	0	0
SD≤1	2	1	N/A	N/A	1	0	4
Drop, n							
R1	0	0	0	0	N/A	0	0
R2	0	0	N/A	N/A	0	0	0

[a]N/A: not applicable.

[b]Flagged items with SD≤1 were dropped.

Experts' Supplemental Comments

Themes emerged within the comment section that guided the revision of the round 2 survey. Various questions by experts reflected a misunderstanding about how to rate an item's importance (eg, based on their specific projects or mHealth HIV interventions generally). For round 2, a clarifying statement was included in the instructions to correct this issue. Experts were reminded to focus on overall recommendations that are important even if they do not apply to all intervention types. Additionally, some items (n=6) were revised based on feedback that they were unclear and required further clarification.

Retained Items

After review of the quantitative and qualitative analyses, decisions were made about which items to retain for the next round. A total of 4 flagged items with a SD≤1 were dropped, and 3 others were eliminated due to feedback related to the similarities among items in the evaluation: data source section. Specific data sources such as paradata, electronic medical records (EMRs), self-reports, and qualitative data individually were found to be less important than generally identifying an appropriate data source for evaluation of mHealth HIV interventions. Therefore, 3 data source questions were dropped and 1 was revised into a single simplified data source item. The round 2 survey included 28 survey items reduced from the original 35 survey items in the round 1 survey.

Round 2

Item Score Analysis

Experts were asked to score the importance of the revised 28-item survey. Following the same analytical process as in round 1 (see Table 1), 24 survey items were retained, 4 were flagged, and none fit the dropped criteria (see Table 1). The 4 items that scored in the flagged range (somewhat important, 2, and important, 3) in both round 1 and again in round 2 (see Table 2) were subsequently removed from the recommendation list. Over the two rounds of data collection, experts scored these flagged items consistently as less important compared to the retained items (important, 3, and very important, 4).

Table 2. Rounds 1 and 2 scores for flagged items that were dropped from expert's recommendation list.

Flagged items	Round 1, mean (SD)	Round 2, mean (SD)
Design: preparation		
Weigh the advantages and disadvantages of integration with an electronic medical record	2.83 (0.94)	2.81 (0.81)
Consider the advantages and disadvantages of an intervention that is publicly available vs offered by providers in a clinical setting	3.09 (0.73)	2.86 (1.06)
Design: core intervention elements		
Use a theoretical foundation to guide intervention development	3.00 (0.85)	2.95 (0.74)
Evaluation: preparation, study design, and data source		
Consider a study design other than traditional randomized controlled trials, such as a multiphase optimization strategy, sequential multiple assignment randomized trials, or pragmatic trials to optimize and tailor the intervention	3.13 (0.82)	2.76 (0.70)

Expert Recommendations

Overall, 86% of the items in round 2 were scored by experts as important (3) to very important (4), indicating a consensus was reached, and a third round of survey data collection was not needed. Table 3 provides the experts' final 24 recommendations for developing, evaluating, and implementing HIV prevention and care mHealth interventions.

As shown in Table 3, the majority (15/24) of the final recommendations experts assessed as the most important were related to intervention design. Design encompasses a broad range of considerations from who will be on the design team, cost, content, privacy concerns, and more. For example, expert made the following comments on the importance of design consideration:

> *...The intervention is not for us. It is not something fun to do with hundreds of thousands of dollars and hours. It is a product we're building with our population to make impactful and positive change in the lives and health of our target population.* [On AYA-specific content]
>
> *Having the intervention be stealth during everyday life is important. However, message content, when participants are interacting with the intervention, may need to push the privacy boundaries to make it relevant.* [On importance of privacy]
>
> *When it comes to youth (and perhaps people in general), novelty is important as is continuous development. Nothing is perfect from the beginning. There will always be something that can be tweaked, added, or removed in order to better serve the users of the technology* [On incorporating an iterative process]

Evaluation items made up 2 of the final recommendations. As discussed previously, data sources individually were not viewed as important as tailoring the evaluation to each intervention appropriately. Finally, there were 7 final implementation recommendations.

Table 3. Final expert recommendations and average importance scores (range 1-4: not important to very important).

	Round 1, mean (SD)	Round 2, mean (SD)
Design: preparation recommendations		
1. Have a multidisciplinary team including experts in HIV, adolescent development, software programming, user interface design, and youth	3.83 (0.39)	3.81 (0.40)
2. Include youth throughout the process to maximize engagement	3.78 (0.52)	3.71 (0.56)
3. Weigh challenges and benefits of using existing technology (platforms or apps) vs newly created technology	3.30 (0.70)	3.24 (0.83)
4. Use an iterative process (incorporating feedback and refining) in developing or adapting the intervention	3.65 (0.49)	3.81 (0.51)
5. Design with sustainability in mind	3.52 (0.79)	3.52 (0.68)
6. Consider the balance between cost and functionality	3.56 (0.59)	3.38 (0.59)
7. Develop an intervention that is accessible, for instance it is platform agnostic (ie, can be used on Android, iOS, or Windows) and available to the majority of youth (ie, those with limited cell phone plans and access to Wi-Fi)	3.74 (0.54)	3.67 (0.58)
8. Determine the appropriate level of real human engagement (eg, automated messaging vs live human coach)	3.39 (0.84)	3.38 (0.59)
9. Consider whether you want to design for youth at various points across the continuum or for a specific target audience	3.36 (1.01)	3.33 (0.66)
Design: core intervention elements recommendations		
1. Ensure that intervention content is relevant to the needs of your specific youth population	3.74 (0.54)	3.95 (0.22)
2. Consider the appropriate dose/frequency of the intervention for optimal efficacy	3.61 (0.66)	3.57 (0.60)
3. Determine the most appropriate digital health modalities (eg, SMS text messaging, social media, mobile website, e-coach, app, or telemedicine) or a combination that will allow for the maximal engagement and effectiveness of intervention	3.74 (0.54)	3.52 (0.51)
4. Consider privacy and confidentiality in design (app icon, message content, home screen)	3.78 (0.52)	3.71 (0.46)
5. Ensure content and engagement strategies are developmentally and culturally appropriate	3.78 (0.42)	3.62 (0.50)
6. Maximize engagement strategies with the intervention to address issues with attrition	3.57 (0.51)	3.38 (0.67)
Evaluation: preparation, study design, and data source recommendations		
1. Determine which intervention components are most effective across groups or individuals	3.35 (0.78)	3.19 (0.68)
2. Establish the best ways to collect data and what types of data (eg, Google analytics, paradata, self-report data, or electronic medical record)	3.35 (0.65)	3.38 (0.67)
Implementation recommendations		
1. Consider the cost and logistics of any human component	3.52 (0.67)	3.38 (0.67)
2. Plan for participant technology disruption (lost, stolen, broken phone, or plan cut off) by collecting multiple modes of contact information and making it easy to reupload or log in to a platform	3.78 (0.52)	3.52 (0.51)
3. Consider integration into routine clinical care and community-based services where appropriate	3.17 (0.83)	3.48 (0.75)
4. Seek youth input about strategies to improve engagement with the intervention	3.43 (0.67)	3.62 (0.59)
5. Plan for how intervention informational content will be updated	3.48 (0.67)	3.33 (0.66)
6. Anticipate changes in platforms or operating systems	3.26 (0.69)	3.43 (0.68)
7. Consider strategies to meet milestones of potentially different timelines of research (eg, grant) and technology partners (eg, development and revision)	3.35 (0.83)	3.10 (0.70)

Discussion

Principle Findings and Comparison With Prior Work

The modified Delphi method allowed for the rapid collection of data including the opinions of more than 20 AYA-focused mHealth HIV intervention experts with varied experience across the United States. Following two rounds of data collection, an original list of 35 items were modified into 24 recommendations on the design, evaluation, and implementation of HIV prevention and care mHealth interventions for AYAs. These consensus items provide guidelines for future development, evaluation, and implementation of mHealth interventions as well as for prevention scientists to support the health of AYAs at risk for acquiring or living with HIV. At each stage of AYA-focused mHealth HIV research, these considerations can guide

decision-making to optimize efficacy, acceptability, and accessibility of mHealth technology to work toward ending the HIV epidemic.

The three items that experts scored as most important were related to intervention design: ensuring that intervention content is relevant to the needs of the specific youth population; having a multidisciplinary team including experts in HIV, adolescent development, software programming, user interface design, and youth; and using an iterative process (incorporating feedback and refining) in developing or adapting the intervention. These design recommendations highlight the potential value of youth involvement and using community-engaged research principles (eg, listening to your community stakeholders throughout the research process to build trust and improve health outcomes that outlast the study duration) [35,36]. These processes can ensure a strong foundation for the intervention as well as for intervention engagement and sustainability.

This study also identified that intrinsic elements of the design are critical for accessibility and privacy. The accessibility of mHealth interventions may fluctuate depending on the type of platform (eg, iOS, Android, or web applications), device's operating system version, and adolescent's access to Wi-Fi and data plan. Privacy and confidentiality of HIV-related mHealth interventions has been documented as very important to YLH [13,37,38], and AYA HIV experts in this study agreed. Preventing unwanted disclosures in fear of the stigma associated with HIV may stop adolescents from downloading or using mHealth interventions that do not protect their privacy and confidentiality [37]. Password-protected accounts, discrete app names, and limiting the access to one's health information are ways to improve privacy and confidentiality [37,38].

The key considerations for the evaluation of mHealth interventions identified included determining what intervention components were the most effective and advantageous data collection strategies. The round 1 survey analysis found experts scored recommendations on individual data collection methods such as self-report, paradata, and EMRs as less important than selecting the right one for a specific project. This suggests that a one-size-fits-all approach to mHealth intervention evaluations may not be appropriate and should rather be tailored to fit each project's goals and evolving technology. A clear understanding of the intervention's desired outcomes may help to determine the most appropriate data collection methods.

In regard to real-world implementation of interventions, experts scored seeking youth input about strategies to improve engagement highest, indicating youth input would be the most important factor. Using a multidisciplinary team, including AYA input, from the start of intervention development was regarded as most important among experts. Again, the importance of including AYA input throughout the entirety of the project was recognized as crucial to reach and engage the youth audience the intervention was designed to help. Other highlights from the implementation recommendations primary focus on technology limitations such as high costs and frequent updates to platforms, operating systems, or devices. Understanding and taking into account the challenges facing youth most affected by HIV may help plan for a more successful implementation. For example, multiple forms of contact information, like emails or social media handles, may be helpful to collect since mobile phone disruptions are common in this group [39]. Additionally, staff time for technology maintenance and initiation were highlighted as important aspects to consider for intervention implementation for clinical care teams and community health organizations.

Limitations

While not hosting a face-to-face discussion among experts allowed us to avoid scheduling challenges and reduced participant time burden, it remains a potential limitation to this study [23]. The option for experts to leave comments served as a discussion forum that allowed us to receive every expert's opinion without influence from other experts. Another limitation was the potential for bias among experts invited to participate. Additionally, demographic, employment responsibilities, or years of experience were not collected. For this reason, researchers who occupied diverse roles on research teams were intentionally invited as experts to reduce this bias and gain a more complete understanding of different perspectives of the research design, evaluation, and implementation process.

Conclusions

This study is among the first to propose expert recommendations on the development, evaluation, and implementation of mHealth interventions for HIV prevention among AYAs. With a clear focus on the role of youth in all aspects of the process, these expert opinions may not only help move forward the quality of technology-based research for adolescent HIV prevention but also ensure that successful interventions will be disseminated broadly for the most impact.

Acknowledgments

The authors would like to thank the experts for their time and efforts in the study. This research was supported by the National Institutes of Health, K23 (NIH K23MH102128, Dowshen).

Conflicts of Interest

None declared.

References

1. Kann L, McManus T, Harris WA, Shanklin SL, Flint KH, Queen B, et al. Youth risk behavior surveillance - United States, 2017. MMWR Surveill Summ 2018 Jun 15;67(8):1-114 [FREE Full text] [doi: 10.15585/mmwr.ss6708a1] [Medline: 29902162]
2. HIV Surveillance Report, 2019. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-2018-updated-vol-32.pdf [accessed 2021-12-20]
3. HIV in the United States by age. Centers for Disease Control and Prevention. 2021 Sep. URL: https://www.cdc.gov/hiv/group/age/youth/index.html [accessed 2021-12-21]
4. Hess KL, Hu X, Lansky A, Mermin J, Hall HI. Lifetime risk of a diagnosis of HIV infection in the United States. Ann Epidemiol 2017 Apr;27(4):238-243 [FREE Full text] [doi: 10.1016/j.annepidem.2017.02.003] [Medline: 28325538]
5. Arrington-Sanders R, Hailey-Fair K, Wirtz AL, Morgan A, Brooks D, Castillo M, et al. Role of structural marginalization, HIV stigma, and mistrust on HIV prevention and treatment among young Black Latinx men who have sex with men and transgender women: perspectives from youth service providers. AIDS Patient Care STDS 2020 Jan;34(1):7-15 [FREE Full text] [doi: 10.1089/apc.2019.0165] [Medline: 31944853]
6. Ending the HIV epidemic. HIV.gov. 2020 Oct. URL: https://files.hiv.gov/s3fs-public/ending-the-hiv-epidemic-flyer.pdf [accessed 2021-12-03]
7. National HIV/AIDS strategy: what you need to know. HIV.gov. 2021 Dec. URL: https://hivgov-prod-v3.s3.amazonaws.com/s3fs-public/NHAS-2022-2025-2Pager.pdf [accessed 2021-12-03]
8. Anderson M, Jiang J. Teens, social media and technology 2018. Pew Research Center. 2018 May 20. URL: https://www.pewinternet.org/2018/05/31/teens-social-media-technology-2018/#vast-majority-of-teens-have-access-to-a-home-computer-or-smartphone [accessed 2019-10-04]
9. Cordova D, Alers-Rojas F, Lua FM, Bauermeister J, Nurenberg R, Ovadje L, et al. The usability and acceptability of an adolescent mHealth HIV/STI and drug abuse preventive intervention in primary care. Behav Med 2018;44(1):36-47 [FREE Full text] [doi: 10.1080/08964289.2016.1189396] [Medline: 27223646]
10. Cho H, Powell D, Pichon A, Thai J, Bruce J, Kuhns LM, et al. A mobile health intervention for HIV prevention among racially and ethnically diverse young men: usability evaluation. JMIR Mhealth Uhealth 2018 Sep 07;6(9):e11450 [FREE Full text] [doi: 10.2196/11450] [Medline: 30194060]
11. Saberi P, Siedle-Khan R, Sheon N, Lightfoot M. The use of mobile health applications among youth and young adults living with HIV: focus group findings. AIDS Patient Care STDS 2016 Jun;30(6):254-260 [FREE Full text] [doi: 10.1089/apc.2016.0044] [Medline: 27214751]
12. Anderson-Lewis C, Darville G, Mercado RE, Howell S, Di Maggio S. mhealth technology use and implications in historically underserved and minority populations in the United States: systematic literature review. JMIR Mhealth Uhealth 2018 Jun 18;6(6):e128 [FREE Full text] [doi: 10.2196/mhealth.8383] [Medline: 29914860]
13. Biello KB, Hill-Rorie J, Valente PK, Futterman D, Sullivan PS, Hightow-Weidman L, et al. Development and evaluation of a mobile app designed to increase HIV testing and pre-exposure prophylaxis use among young men who have sex with men in the United States: open pilot trial. J Med Internet Res 2021 Mar 24;23(3):e25107 [FREE Full text] [doi: 10.2196/25107] [Medline: 33759792]
14. Kuhns LM, Hereth J, Garofalo R, Hidalgo M, Johnson AK, Schnall R, et al. A uniquely targeted, mobile app-based HIV prevention intervention for young transgender women: adaptation and usability study. J Med Internet Res 2021 Mar 31;23(3):e21839 [FREE Full text] [doi: 10.2196/21839] [Medline: 33787503]
15. Tanner AE, Song EY, Mann-Jackson L, Alonzo J, Schafer K, Ware S, et al. Preliminary impact of the weCare social media intervention to support health for young men who have sex with men and transgender women with HIV. AIDS Patient Care STDS 2018 Nov;32(11):450-458 [FREE Full text] [doi: 10.1089/apc.2018.0060] [Medline: 30398955]
16. Ampt FH, Mudogo C, Gichangi P, Lim MSC, Manguro G, Chersich M, et al. WHISPER or SHOUT study: protocol of a cluster-randomised controlled trial assessing mHealth sexual reproductive health and nutrition interventions among female sex workers in Mombasa, Kenya. BMJ Open 2017 Aug 18;7(8):e017388 [FREE Full text] [doi: 10.1136/bmjopen-2017-017388] [Medline: 28821530]
17. Garofalo R, Kuhns LM, Hotton A, Johnson A, Muldoon A, Rice D. A randomized controlled trial of personalized text message reminders to promote medication adherence among HIV-positive adolescents and young adults. AIDS Behav 2016 May;20(5):1049-1059 [FREE Full text] [doi: 10.1007/s10461-015-1192-x] [Medline: 26362167]
18. 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]
19. Whiteley L, Brown LK, Mena L, Craker L, Arnold T. Enhancing health among youth living with HIV using an iPhone game. AIDS Care 2018;30(sup4):21-33 [FREE Full text] [doi: 10.1080/09540121.2018.1503224] [Medline: 30626196]
20. Ybarra ML, Prescott TL, Phillips GL, Bull SS, Parsons JT, Mustanski B. Pilot RCT results of an mHealth HIV prevention program for sexual minority male adolescents. Pediatrics 2017 Jul;140(1):e20162999 [FREE Full text] [doi: 10.1542/peds.2016-2999] [Medline: 28659456]

21. Gannon B, Davis R, Kuhns LM, Rodriguez RG, Garofalo R, Schnall R. A mobile sexual health app on empowerment, education, and prevention for young adult men (MyPEEPS Mobile): acceptability and usability evaluation. JMIR Form Res 2020 Apr 07;4(4):e17901 [FREE Full text] [doi: 10.2196/17901] [Medline: 32254043]
22. Sayegh CS, MacDonell KK, Clark LF, Dowshen NL, Naar S, Olson-Kennedy J, et al. The impact of cell phone support on psychosocial outcomes for youth living with HIV nonadherent to antiretroviral therapy. AIDS Behav 2018 Oct;22(10):3357-3362. [doi: 10.1007/s10461-018-2192-4] [Medline: 29948339]
23. Dalkey N, Helmer O. An experimental application of the DELPHI method to the use of experts. Manage Sci 1963 Apr;9(3):458-467. [doi: 10.1287/mnsc.9.3.458]
24. Miller MK, Chernick LS, Goyal MK, Reed JL, Ahmad FA, Hoehn EF, et al. A research agenda for emergency medicine-based adolescent sexual and reproductive health. Acad Emerg Med 2019 Dec;26(12):1357-1368. [doi: 10.1111/acem.13809] [Medline: 31148339]
25. Ranganathan P, Aggarwal R. Study designs: Part 1 - an overview and classification. Perspect Clin Res 2018;9(4):184-186 [FREE Full text] [doi: 10.4103/picr.PICR_124_18] [Medline: 30319950]
26. Program Performance and Evaluation Office (PPEO). Centers for Disease Control and Prevention. 2011. URL: https://www.cdc.gov/eval/guide/cdcevalmanual.pdf [accessed 2021-12-21]
27. Toolkit part 1: implementation science methodologies and frameworks. Fogarty International Center. URL: tinyurl.com/3sep5yjx [accessed 2021-04-01]
28. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
29. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
30. Hightow-Weidman LB, Muessig K, Rosenberg E, Sanchez T, LeGrand S, Gravens L, et al. University of North Carolina/Emory Center for Innovative Technology (iTech) for addressing the HIV epidemic among adolescents and young adults in the United States: protocol and rationale for center development. JMIR Res Protoc 2018 Aug 03;7(8):e10365 [FREE Full text] [doi: 10.2196/10365] [Medline: 30076126]
31. SPNS Initiative: use of social media to improve engagement, retention, and health outcomes along the HIV care continuum, 2015-2019. Ryan White HIV/AIDS Program. 2020 Oct. URL: https://hab.hrsa.gov/about-ryan-white-hivaids-program/spns-social-media [accessed 2021-05-10]
32. Research Portfolio Online Reporting Tools (RePORT). 2021 May. URL: https://report.nih.gov/ [accessed 2021-05-10]
33. Welcome to Medical Subject Headings. National Library of Medicine. 2021 Jan. URL: https://www.nlm.nih.gov/mesh/meshhome.html [accessed 2021-05-10]
34. Castleberry A, Nolen A. Thematic analysis of qualitative research data: is it as easy as it sounds? Curr Pharm Teach Learn 2018 Jun;10(6):807-815. [doi: 10.1016/j.cptl.2018.03.019] [Medline: 30025784]
35. Oetzel JG, Wallerstein N, Duran B, Sanchez-Youngman S, Nguyen T, Woo K, et al. Impact of participatory health research: a test of the community-based participatory research conceptual model. Biomed Res Int 2018;2018:7281405. [doi: 10.1155/2018/7281405] [Medline: 29854784]
36. Kwon S, Tandon SD, Islam N, Riley L, Trinh-Shevrin C. Applying a community-based participatory research framework to patient and family engagement in the development of patient-centered outcomes research and practice. Transl Behav Med 2018 Sep 08;8(5):683-691 [FREE Full text] [doi: 10.1093/tbm/ibx026] [Medline: 30202926]
37. Holloway IW, Winder TJ, Lea CH, Tan D, Boyd D, Novak D. Technology use and preferences for mobile phone-based HIV prevention and treatment among black young men who have sex with men: exploratory research. JMIR Mhealth Uhealth 2017 Apr 13;5(4):e46 [FREE Full text] [doi: 10.2196/mhealth.6436] [Medline: 28408360]
38. Chandler R, Hernandez N, Guillaume D, Grandoit S, Branch-Ellis D, Lightfoot M. A community-engaged approach to creating a mobile HIV prevention app for Black women: focus group study to determine preferences via prototype demos. JMIR Mhealth Uhealth 2020 Jul 24;8(7):e18437 [FREE Full text] [doi: 10.2196/18437] [Medline: 32706723]
39. Avery A, Dowshen N, Thompson M, Aladin B, Kratz J, Smith J, et al. Implementation challenges encountered in the Special Projects of National Significance (SPNS) Social Media Initiatives (SMI). 2018 Dec Presented at: National Ryan White Conference on HIV Care & Treatment; December 2018; Oxon Hill, MD.

Abbreviations

ART: antiretroviral therapy
AYA: adolescent and young adult
EHE: Ending the HIV Epidemic
EMR: electronic medical record
GBMSM: gay, bisexual, other men who have sex with men
MeSH: Medical Subject Headings

mHealth: mobile health
NIH: National Institutes of Health
REDCap: Research Electronic Data Capture
YLH: youth living with HIV

Edited by A Mavragani; submitted 07.01.22; peer-reviewed by D Levine, A Montilha, K Chigudu; comments to author 23.05.22; revised version received 09.06.22; accepted 15.06.22; published 12.07.22.

Please cite as:
Knowles K, Dowshen N, Lee S, Tanner A
Developing a Consensus for Adolescent and Young Adult mHealth HIV Prevention Interventions in the United States: A Delphi Method Study
JMIR Form Res 2022;6(7):e25982
URL: https://formative.jmir.org/2022/7/e25982
doi:10.2196/25982
PMID:35819824

Original Paper

Improving Well-being With a Mobile Artificial Intelligence–Powered Acceptance Commitment Therapy Tool: Pragmatic Retrospective Study

Navot Naor[1], PhD; Alex Frenkel[1], MA; Mirène Winsberg[2], MD

[1]Kai.ai, Tel Aviv, Israel

[2]Kai.ai, San Francisco, CA, United States

Corresponding Author:
Navot Naor, PhD
Kai.ai
Yigal Alon St 114
Tel Aviv, 6744320
Israel
Phone: 972 502183965
Email: navot@kai.ai

Abstract

Background: Research and dissemination of smartphone apps to deliver coaching and psychological driven intervention had seen a great surge in recent years. Notably, Acceptance Commitment Therapy (ACT) protocols were shown to be uniquely effective in treating symptoms for both depression and anxiety when delivered through smartphone apps. The aim of this study is to expand on that work and test the suitability of artificial intelligence–driven interventions delivered directly through popular texting apps.

Objective: This study evaluated our hypothesis that using Kai.ai will result in improved well-being.

Methods: We performed a pragmatic retrospective analysis of 2909 users who used Kai.ai on one of the top messaging apps (iMessage, WhatsApp, Discord, Telegram, etc). Users' well-being levels were tracked using the World Health Organization-Five Well-Being Index throughout the engagement with service. A 1-tailed paired samples *t* test was used to assess well-being levels before and after usage, and hierarchical linear modeling was used to examine the change in symptoms over time.

Results: The median well-being score at the last measurement was higher (median 52) than that at the start of the intervention (median 40), indicating a significant improvement (W=2682927; $P<.001$). Furthermore, HLM results showed that the improvement in well-being was linearly related to the number of daily messages a user sent ($\beta=.029$; $t_{81.36}=4$; $P<.001$), as well as the interaction between the number of messages and unique number of days ($\beta=-.0003$; $t_{81.36}=-2.2$; $P=.03$).

Conclusions: Mobile-based ACT interventions are effective means to improve individuals' well-being. Our findings further demonstrate Kai.ai's great promise in helping individuals improve and maintain high levels of well-being and thus improve their daily lives.

(JMIR Form Res 2022;6(7):e36018) doi:10.2196/36018

KEYWORDS

Acceptance Commitment Therapy; well-being; WHO-5; World Health Organization-Five Well-Being Index; mHealth; mobile health; smartphone; health application; mental health; quality of life; artificial intelligence

Introduction

Psychological distress, which can be broadly understood as an emotional state of suffering from mood- and anxiety-related symptoms [1], is among the leading causes of disability in the world [2]. It has become highly prevalent; it was recently estimated that up to 50% of the world's population experiences symptoms of psychological distress [3]. Knowing how widespread mental health issues are, it is not surprising that they were found to be a central cause for the global burden of diseases worldwide, accounting for some 56.7% of disability-adjusted life years [2]. Furthermore, a recent report estimated that in 2010, poor mental health resulted in a US $2-$5 trillion annual loss worldwide owing to poor health and

reduced productivity. This loss was only projected to increase, expected to approach US $6 trillion by 2030 [4].

While the COVID-19 pandemic imposed numerous hardships on the world, the increase in prevalence and severity of mental distress symptoms reached a level where worldwide organizations declared it a global crisis that needed immediate attention [4,5]. While the number of individuals with mental illnesses has been documented in the second decade of the 21st century, there was no comparable increase in the numbers of individuals receiving treatment for these symptoms over the same period. This highlights the unmet needs for more treatment possibilities [6]. Furthermore, it was recently claimed that despite the substantial investment in the development of new medications and therapeutic protocols, they are only expected to have a limited impact over the long term. In contrast, it is argued that greater attention should be given to the development of prevention strategies and tools [7].

Recently, mobile apps and computer-based applications have become increasingly useful tools for creating and disseminating mental health prevention and treatment programs [8-10]. Mobile interventions introduce a fresh variety of possibilities, including the provision for unlimited interactions and psychological interventions. This creates opportunities to reach patients anytime and anywhere. Furthermore, this kind of platform also has the possibility to enhance efficiency by increasing the treatment intensity and integrating therapeutic strategies into daily life [11].

The use of mobile-based innervations was found to be particularly effective when it was used to deliver Acceptance Commitment Therapy (ACT) protocols [12-15]. Since ACT protocols allow for a low-intensity or high-frequency approach [12], they are a highly effective method to facilitate changes in the flexibility of the thinking process, which is one of ACT's central features [16]. ACT is a form of psychotherapy positioned within the third wave of behavioral therapies. It focuses on teaching patients how to deal with challenging experiences by reflecting on their values [16]. ACT works to develop psychological flexibility, which entails having the ability to stay in contact with the present regardless of any unpleasant thoughts, feelings, or bodily sensations. Recently, ACT was shown to be effective in treating symptoms of both depression [14] and anxiety [13].

In this pragmatic retrospective study, we report the test results of a mobile platform called Kai.ai. Kai.ai is a platform designed to provide ACT-based, artificial intelligence–driven conversational coaching that helps users build habits for healthy living and resilience. It integrates seamlessly with all the top messaging apps (iMessage, WhatsApp, Discord, Telegram, etc) allowing for ecological momentary assessment and ecological momentary interventions. The platform checks in with users throughout the day and leverages users' responses and journaling exercises to deliver tailored ACT skilled coaching. In this study, the platform was evaluated using both a pre-post and longitudinal design to assess changes in measures of well-being among users. Our hypothesis being that using Kai.ai will result in improved well-being.

Methods

Study Design

A pragmatic retrospective design was used for this study, using data collected as part of users' engagement with Kai.ai's services. All participants were located in the United States and had free access to the service, which is seamlessly integrated with all the top messaging apps (iMessage, WhatsApp, Discord, Telegram, etc). This retrospective study included data from 2909 Kai.ai users, 85% of whom were recruited from major social networks such as TikTok, and they are presented with an advertisement inviting them to test how happy they are. Clicking the advertisement would direct users to a landing page (How happy are you? - Kai), where they can complete the World Health Organization-Five Well-Being Index (WHO-5). Upon completing the questionnaire, users were asked whether they would like to continue and engage with Kai.ai in the future.

During the first engagement, users were asked to complete an electronically secured version of the WHO-5. The survey was completely optional, and users could opt to not take part in the assessment phase. As users continued using Kai.ai, they were prompted once every 6 weeks to complete an additional WHO-5 assessment.

Participants

Participants completed the WHO-5 at least twice, with the highest number of completed WHO-5 measurements being 9. All participants in this study were from the United States, but since the data set used in this study was completely anonymized before it was given to the authors, no gender- or age-related information was available.

Ethical Considerations

This report is based on a postfactum analysis, which used anonymized data gathered during users' engagement with Kai.ai's services. As such, it was determined as not being a study involving human subjects by WCG IRB (1-1504102-1).

WHO-5

The WHO-5 is a short self-report to measure well-being [17,18]. Items are rated on a 6-point Likert scale, where participants are asked to rate their experiences in the past 2 weeks from 5=all of the time to 0=at no time. Raw scores are then multiplied by 4 to transform the scale to a percentage-based scale, ranging from 0% to 100%. A transferred score of below 50% reflects poor well-being, and a 10-point change in the translated score is seen as clinically relevant [17]. Owing to its ease of use and high validity, the WHO-5 has been extensively used in mental health– and physical health–related mobile apps [19-23].

Intervention

Overview

Users can choose to interact with Kai.ai through common messaging apps. Each user's first interaction starts with an onboarding process where the WHO-5 is administered. Based on the users' responses, Kai.ai builds 3 routines for the morning, noon, and evening. Each routine includes one or more of the

service's main blocks: (1) *Reflections and Journaling* and (2) *Exercises*.

Reflections and Journaling

For *Reflections and Journaling*, users are invited to think and write about their experiences. To maintain structure and help guide users, Kai.ai uses three kinds of journaling cues: gratitude, learning, and one good thing that happened.

During the gratitude exercise, users are encouraged to identify things for which they are grateful and to list them in writing, where they can easily review this list daily—this gives room to develop flexible thinking patterns. In the learning exercise, users are asked to focus on the lessons they can learn from their experiences and, lastly, in one good thing that *Exercises* users learn to reduce stress and anxiety by focusing on a single task rather than a long to-do list and spreading themselves too thin. Adopting a routine reflection and journaling process is meant to help users achieve a more centered, grounded, joyous, and purposeful state of mind.

Exercises

Two main types of exercises are offered to users: breathing exercises and ACT training. The breathing exercises focus on teaching users how to breathe in a constructive manner; breathing through their noses, using their diaphragm, and noting their posture. These are all meant to ensure a better flow of oxygen to the body, allowing for a reduction in stress and anxiety through the operation of the parasympathetic nervous system [24].

The used ACT training is based on the mindfulness approach aimed at increasing psychological flexibility, and treating pain and discomfort as facts of life can be used for personal growth through a process of acceptance and validation [16].

User Interaction

Users have 2 main ways to interact with Kai.ai: passive and active. In the passive way, the service checks in on the user at least once and not more than 3 times a day, depending on the routine created during onboarding. The check-in prompts can include either *Reflection and Journaling* open-ended questions or closed questions where users are asked to indicate how happy they are. In addition to the passive check-ins, users can actively reach out to Kai.ai by simply texting it whenever they feel the need to.

Data Analyses

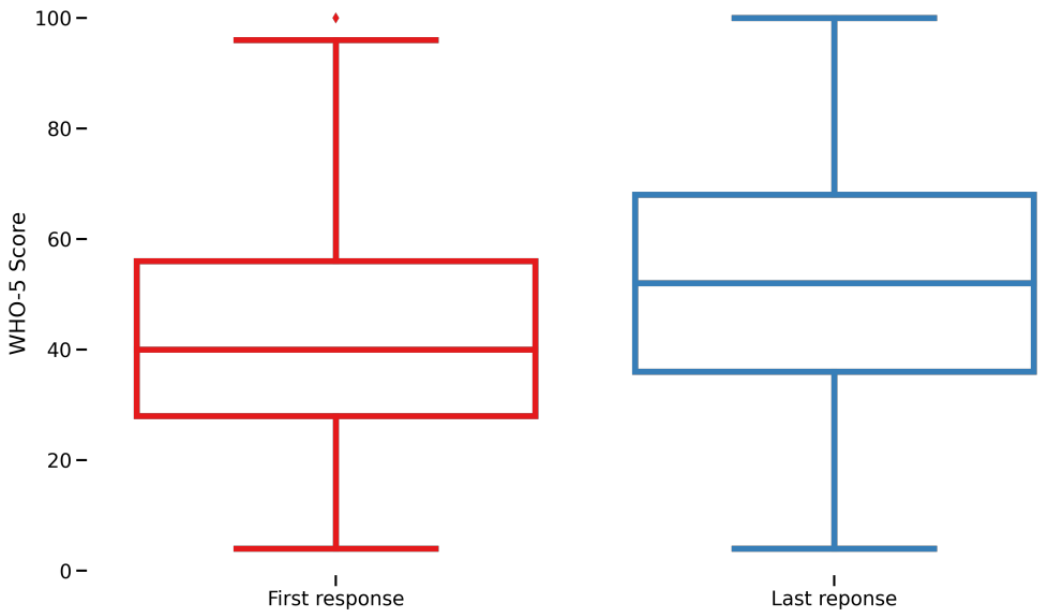
To test Kai.ai's impact on participants' well-being, 2 analytical methods were used. First, a paired samples *t* test was performed to compare participants' baseline WHO-5 scores, which were collected just before the initiation of use, and the last WHO-5 score recorded for each participant. Following the initial analysis, hierarchical linear modeling (HLM) was used to examine the change in the users' well-being over time. HLM was chosen over other repeated measures models, such as repeated measures ANOVA, owing to its unique ability to model incomplete or imbalanced data sets. In the first level, the HLM was modeled to estimate the individual linear slopes and intercepts, and in the second level, the sample's average slope and intercept were estimated.

Results

Paired Samples t Tests

Users' scores on the WHO-5 were compared from the first and last measurement of each user. On average, users' last measurement score was better (median 52) than their first (median 40). A Wilcoxon signed rank test indicated that this difference was significant (W=2682927; P<.001). Figure 1 shows a box plot of the median (IQR) values of change in WHO-5 scores.

Figure 1. Box plot of the median (IQR) values of change in World Health Organization-Five Well-Being Index (WHO-5) scores for users' first and last measurements.



HLMs

In addition to the paired samples *t* test that was used to compare users first and last measurement scores, an HLM was used since it is best suited to deal with incomplete data sets while still accounting for the data set's repeated measures nature [25]. This unique strength of the HLM analysis was this analysis because the number of measurements among users varied between 2 and 9. Some users used the app for only 2 months, others used it for well over a year.

Consequently, our model tested whether differences in users' WHO-5 score between any 2 consecutive measurements could be predicted by the level of engagement during the time frame. As such, we used 2 variables as indications of engagement: the number of unique days during which the user had active interactions with Kai.ai during the specified time frame and the number of daily messages the user sent during that time. Using the R nlme package [26], the effect of the engagement variables on the difference in WHO-5 scores was first tested separately for each measurement time point (second to ninth) and then aggregated over all time points. Our results show that the difference in WHO-5 scores was linearly related to the number of daily messages a user sent (β=.029; $t_{81.36}$=4; *P*<.001) and the interaction between the number of messages and unique number of days (β=–.0003; $t_{81.36}$=–2.2; *P*=.03), but not to the unique number of days on its own (Table 1)

Table 1. Unconditional growth model for changes in the World Health Organization-Five Well-Being Index (WHO-5) scores over time.

WHO-5 score	Estimate (β)	SE	*t* test (*df*)	*P* value
Intercept	.563	1.045	0.536 (81.36)	.59
Total daily messages	.029	.007	4 (81.36)	<.001
Unique active days	–.015	.036	–0.427 (81.36)	.67
Interaction	–.0003	.0001	–2.2 (81.36)	.03

Discussion

Principal Findings

In this paper, we report the results of a pragmatic retrospective study aiming to test the effectiveness of a mobile phone–delivered, ACT-based, artificial intelligence–delivered conversational coaching platform. In the onboarding measurement, half of all participants reported a WHO-5 score of 40 or less, which is well below the cut-off point of 50, and an indicator of poor well-being. However, on the last measurement recorded for each user, half of all participants indicated a score of 52 and above, indicating a change for the better and an overall good well-being. These results highlight the great potential ACT-based mobile apps can have for improving users' daily well-being.

Limitations

This paper describes the use of a novel platform and treatment protocol using mobile messaging apps to deploy ACT-based treatment to improve the well-being of users. Our results, while highlighting Kai.ai's great potential as reflected by the pre-post results, also hints at the wealth of possibilities still left to discover. The results of our HLM model, especially those of the negative interaction between the engagement variables, indicates that further research is needed to fully understand and describe Kai.ai's potential and fit. An example of such an area, which future research should expand on, is Kai.ai's suitability to different genders and age groups. A recent study found differences among these groups in the adaptability of mobile-based coaching tools, as females and younger age groups (18-25 years) showed higher adaptability than males and older ones, respectively [27]. To overcome these adaptability challenges, a number of key features are introduced, which are directly aimed at improving users' engagement, leveraging the motivations of choice and progression. As such, an emphasis was placed on greater choice, as users are given the option of choosing tracks that focus on different central core ideas, such as relationships, mood, and life habits. Furthermore, over time, the daily routines and exercises offered to the user are selected on the basis of past preferences. Lastly, by allowing users to review their progress, allowing users the ability to view their progress time.

Conclusions

The findings reported in this paper show Kai.ai's great promise in helping individuals improve and maintain high levels of well-being and thus improve their daily life. These findings are further emphasized by the great need for mental health interventions, while the availability of and accessibility to treatment remains limited. It seems that the use of such technology, which is found to improve mental well-being per the findings of the WHO-5, is essential, especially nowadays.

Conflicts of Interest

AF is employed by Kai.ai, receives income from Kai.ai, and has been granted equity in Kai.ai. NN and MW receive income from Kai.ai and have been granted equity in Kai.ai.

References

1. Drapeau A, Marchand A, Beaulieu-Prévost D. Epidemiology of Psychological Distress. In: L'Abate L, editor. Mental Illnesses - Understanding, Prediction and Control. Norderstedt: BoD – Books on Demand; 2012.

2. Whiteford HA, Ferrari AJ, Degenhardt L, Feigin V, Vos T. The global burden of mental, neurological and substance use disorders: an analysis from the Global Burden of Disease Study 2010. PLoS One 2015;10(2):e0116820 [FREE Full text] [doi: 10.1371/journal.pone.0116820] [Medline: 25658103]
3. Nochaiwong S, Ruengorn C, Thavorn K, Hutton B, Awiphan R, Phosuya C, et al. Global prevalence of mental health issues among the general population during the coronavirus disease-2019 pandemic: a systematic review and meta-analysis. Sci Rep 2021 May 13;11(1):10173 [FREE Full text] [doi: 10.1038/s41598-021-89700-8] [Medline: 33986414]
4. The Lancet Global Health. Mental health matters. Lancet Glob Health 2020 Nov;8(11):e1352 [FREE Full text] [doi: 10.1016/S2214-109X(20)30432-0] [Medline: 33069297]
5. United Nations. COVID-19 and the Need for Action on Mental Health. UN Executive Office of the Secretary-General (EOSG) Policy Briefs and Papers. 2020 May 21. URL: https://www.un-ilibrary.org/content/papers/27082245/9 [accessed 2022-05-25]
6. Greenberg PE, Fournier A, Sisitsky T, Simes M, Berman R, Koenigsberg SH, et al. The Economic Burden of Adults with Major Depressive Disorder in the United States (2010 and 2018). Pharmacoeconomics 2021 Jun;39(6):653-665 [FREE Full text] [doi: 10.1007/s40273-021-01019-4] [Medline: 33950419]
7. Ormel J, Cuijpers P, Jorm A, Schoevers RA. What is needed to eradicate the depression epidemic, and why. Ment Health Prev 2020 Mar;17:200177. [doi: 10.1016/j.mhp.2019.200177]
8. 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]
9. Lindhiem O, Bennett C, Rosen D, Silk J. Mobile technology boosts the effectiveness of psychotherapy and behavioral interventions: a meta-analysis. Behav Modif 2015 Nov;39(6):785-804 [FREE Full text] [doi: 10.1177/0145445515595198] [Medline: 26187164]
10. Luxton D, McCann R, Bush N, Mishkind M, Reger G. mHealth for mental health: Integrating smartphone technology in behavioral healthcare. Prof Psychol Res Pr 2011 Dec;42(6):505-512. [doi: 10.1037/a0024485]
11. Ebert D, Van Daele T, Nordgreen T, Karekla M, Compare A, Zarbo C, et al. Internet- and Mobile-Based Psychological Interventions: Applications, Efficacy, and Potential for Improving Mental Health. Eur Psychol 2018 May;23(2):167-187. [doi: 10.1027/1016-9040/a000318]
12. Krafft J, Potts S, Schoendorff B, Levin ME. A Randomized Controlled Trial of Multiple Versions of an Acceptance and Commitment Therapy Matrix App for Well-Being. Behav Modif 2019 Mar;43(2):246-272. [doi: 10.1177/0145445517748561] [Medline: 29262693]
13. Landy L, Schneider R, Arch J. Acceptance and commitment therapy for the treatment of anxiety disorders: a concise review. Curr Opin Psychol 2015 Apr;2:70-74. [doi: 10.1016/j.copsyc.2014.11.004]
14. Moore C, Hill J, Feinstein C, Pike C, Delaney K. Implementation of Acceptance Commitment Therapy on a Mood Disorder Unit: A Quality Improvement Project. J Psychosoc Nurs Ment Health Serv 2019 Nov 01;57(11):22-27. [doi: 10.3928/02793695-20190627-01] [Medline: 31305951]
15. Reyes A. The Process of Learning Mindfulness and Acceptance through the Use of a Mobile App Based on Acceptance and Commitment Therapy: A Grounded Theory Analysis. Issues Ment Health Nurs 2022 Jan;43(1):3-12. [doi: 10.1080/01612840.2021.1953652] [Medline: 34346278]
16. Hayes SC. Acceptance and Commitment Therapy, Relational Frame Theory, and the Third Wave of Behavioral and Cognitive Therapies - Republished Article. Behav Ther 2016 Nov;47(6):869-885. [doi: 10.1016/j.beth.2016.11.006] [Medline: 27993338]
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. A. Mastering Depression in Primary Care. Hilleroed: World Health Organization, Regional Office for Europe, Psychiatric Research Unit, Frederiksborg General Hospital; 1998.
19. Andreasson K, Krogh J, Bech P, Frandsen H, Buus N, Stanley B, et al. MYPLAN -mobile phone application to manage crisis of persons at risk of suicide: study protocol for a randomized controlled trial. Trials 2017 Apr 11;18(1):171 [FREE Full text] [doi: 10.1186/s13063-017-1876-9] [Medline: 28399909]
20. Coelhoso C, Tobo P, Lacerda S, Lima A, Barrichello C, Amaro JE, et al. A New Mental Health Mobile App for Well-Being and Stress Reduction in Working Women: Randomized Controlled Trial. J Med Internet Res 2019 Nov 07;21(11):e14269 [FREE Full text] [doi: 10.2196/14269] [Medline: 31697244]
21. Laird E, Ryan A, McCauley C, Bond R, Mulvenna M, Curran K, et al. Using Mobile Technology to Provide Personalized Reminiscence for People Living With Dementia and Their Carers: Appraisal of Outcomes From a Quasi-Experimental Study. JMIR Ment Health 2018 Sep 11;5(3):e57 [FREE Full text] [doi: 10.2196/mental.9684] [Medline: 30206053]
22. Litvin S, Saunders R, Maier M, Lüttke S. Gamification as an approach to improve resilience and reduce attrition in mobile mental health interventions: A randomized controlled trial. PLoS One 2020;15(9):e0237220 [FREE Full text] [doi: 10.1371/journal.pone.0237220] [Medline: 32877425]
23. Stallman H. Efficacy of the My Coping Plan mobile application in reducing distress: A randomised controlled trial. Clin Psychol 2020 Nov 09;23(3):206-212. [doi: 10.1111/cp.12185]

24. Jerath R, Edry J, Barnes V, Jerath V. Physiology of long pranayamic breathing: neural respiratory elements may provide a mechanism that explains how slow deep breathing shifts the autonomic nervous system. Med Hypotheses 2006;67(3):566-571. [doi: 10.1016/j.mehy.2006.02.042] [Medline: 16624497]
25. Maas C, Snijders T. The multilevel approach to repeated measures for complete and incomplete data. Qual Quant 2003;37(1):89. [doi: 10.1023/A:1022545930672]
26. Pinheiro J, Bates D, DebRoy S, Sarkar D, EISPACK, Heisterkamp S, R Core Team. nlme: Linear and Nonlinear Mixed Effects Models. URL: https://cran.r-project.org/web/packages/nlme/nlme.pdf [accessed 2022-05-25]
27. Lungu A, Jun JJ, Azarmanesh O, Leykin Y, Chen CE. Blended Care-Cognitive Behavioral Therapy for Depression and Anxiety in Real-World Settings: Pragmatic Retrospective Study. J Med Internet Res 2020 Jul 06;22(7):e18723 [FREE Full text] [doi: 10.2196/18723] [Medline: 32628120]

Abbreviations

ACT: Acceptance Commitment Therapy
HLM: hierarchical linear modeling
WHO-5: World Health Organization-Five Well-Being Index

Edited by A Mavragani; submitted 28.12.21; peer-reviewed by A Hashim; comments to author 26.01.22; revised version received 15.03.22; accepted 20.05.22; published 12.07.22.

Please cite as:
Naor N, Frenkel A, Winsberg M
Improving Well-being With a Mobile Artificial Intelligence–Powered Acceptance Commitment Therapy Tool: Pragmatic Retrospective Study
JMIR Form Res 2022;6(7):e36018
URL: https://formative.jmir.org/2022/7/e36018
doi:10.2196/36018
PMID:35598216

Original Paper

Online Guided Self-help Cognitive Behavioral Therapy With Exposure to Anxiety and Problem Solving in Type 1 Diabetes Mellitus: Case Study

Dorian Kern[1,2], MSc; Brjánn Ljótsson[2], PhD; Marianne Bonnert[1,3], PhD; Nils Lindefors[1], MD, PhD; Martin Kraepelien[1,2], PhD

[1]Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet & Stockholm Healthcare Services, Stockholm, Sweden
[2]Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
[3]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Dorian Kern, MSc
Centre for Psychiatry Research
Department of Clinical Neuroscience
Karolinska Institutet & Stockholm Healthcare Services
Norra Stationsgatan 69
Stockholm, 11364
Sweden
Phone: 46 723243235
Email: dorian.kern@ki.se

Abstract

Background: Type 1 diabetes mellitus (T1DM) is dependent on self-care to avoid short- and long-term complications. There are several problem areas in diabetes that could be addressed by psychological interventions, such as suboptimal problem-solving strategies and fear of hypoglycemia. There is empirical support for a few psychological interventions, most often cognitive behavioral therapy, with various treatment aims. However, these interventions are largely unavailable in regular diabetes health care. Online guided self-help cognitive behavioral therapy could help achieve greater outreach.

Objective: We tested a manualized treatment in the early stage for further development, with the long-term aim to increase access to care. The purpose of this report was to show the potential of this newly developed online intervention by describing 2 illustrative cases.

Methods: An online guided self-help cognitive behavioral therapy protocol featuring problem solving and exposure was developed. The treatment was administered from a secure online platform and lasted for 8 weeks. Case 1 was a male participant. He had a number of diabetes-related complications and was worried about his future. He reported that he had a general idea that he needed to change his lifestyle but found it difficult to get started. Case 2 was a female participant. She had fear of hypoglycemia and unhelpful avoidance behaviors. She kept her blood glucose levels unhealthily high in order to prevent hypoglycemic episodes. Furthermore, she avoided contact with diabetes health care.

Results: The 2 participants showed clinically significant improvements in their most relevant problem areas. In case 1, the participant's blood glucose levels reduced, and he was able to establish healthy routines, such as increase physical exercise and decrease overeating. In case 2, the participant's fear of hypoglycemia greatly decreased, and she was able to confront many of her avoided situations and increase necessary visits to her diabetes clinic. Treatment satisfaction was high, and no adverse events were reported.

Conclusions: It is possible to deliver a cognitive behavioral therapy intervention aimed at problem areas in diabetes online. Problem solving appears to help with problems in everyday routines and lifestyle choices. Exposure to aversive stimuli appears to be a plausible intervention specifically aimed at the fear of hypoglycemia. Larger and controlled studies are needed.

(JMIR Form Res 2022;6(7):e32950) doi:10.2196/32950

KEYWORDS

fear of hypoglycemia; problem solving; cognitive behavioral therapy; psychology; self-help; diabetes; type 1 diabetes mellitus; high blood glucose

Introduction

Background

Type 1 diabetes mellitus (T1DM) is a chronic condition caused by deterioration in the ability of the pancreas to produce insulin, leading to insufficient insulin levels [1]. Lifestyle factors have not been found to increase the risk of T1DM, although, for example, diet and exercise are important to keep blood glucose stable when the disease is present [2]. Thus, patient self-care behaviors are necessary to maintain a stable and healthy blood glucose level [2]. People with T1DM are required to monitor their blood glucose levels multiple times per day and to plan and adjust insulin doses based on several factors. These factors include daily activities, food intake, exercise, ongoing infection, stress, and others, requiring well-functioning planning and problem-solving skills [3]. Fear of hypoglycemia is a fear specific to insulin-treated diabetes, which can lead to several unhelpful avoidance behaviors [4]. These behaviors generally include keeping blood glucose levels unhealthily high to try to avoid hypoglycemia, as well as avoiding places or situations viewed as potentially dangerous [5].

Psychological interventions aimed at people with T1DM have usually involved different types of cognitive behavioral therapy (CBT), often aimed at emotional coping and behavioral change [4]. One intervention that has been used for people with T1DM is problem solving, which is a technique that teaches a structured approach to solving practical and emotional problems [6]. Another CBT intervention that has been suggested is exposure, an often used and effective treatment component for anxiety disorders. This technique is used to confront feared stimuli, in order to decrease fear and to decrease avoidance of fearful situations [7]. For T1DM, exposure has mainly been proposed as an intervention for fear of hypoglycemia [4]. Exposure has previously been successfully tried for fear of hypoglycemia in a case study of face-to-face CBT [8]. Further evaluation of exposure has been proposed [4]. Exposure has been used successfully in guided self-help treatments for somatic conditions, including irritable bowel syndrome (IBS) [9], atrial fibrillation [10], and asthma [11].

Exposure may theoretically be a sound approach for people with T1DM who display avoidance behaviors and fear of symptoms, and this approach has been successful for similar chronic conditions. Regardless of the helpfulness of CBT in people with T1DM, access to psychologists with relevant knowledge of the diabetes population is severely limited [12]. Over several years, great progress has been made with CBT over the internet. Internet-delivered CBT (internet CBT) has some advantages over traditional CBT. Internet CBT requires less time per therapist and patient, and therefore allows more people to be treated. The format also means that the treatment can be available for people living in rural areas. The effects of internet CBT have been evaluated for a large number of conditions, generally with equal effects as traditional CBT [13]. One recent meta-analysis found 25 randomized controlled trials of CBT for diabetes, although only 5 of these were aimed at type 1 diabetes. The analysis suggested that CBT is an effective intervention to improve blood glucose and quality of life. However, none of the approaches had been completely delivered via the internet, and none had exposure as the main intervention [14]. There is some evidence suggesting the effectiveness of CBT for improving blood glucose and mental health in people with T1DM [14]. Therefore, we believe it is valuable to examine if online CBT for T1DM, including exposure, can be feasible, using a case study as the first step.

Objective

The objective was to develop a manualized CBT treatment delivered online. In this study, we tested a treatment in the very early stage. The participants' course of treatment was explored to get an early indication of what might and might not be practical in this type of treatment. To illustrate this, we describe 2 cases involving 2 participants with T1DM, who completed online guided self-help CBT. The illustrative cases will be used to revise and further adapt the treatment protocol adopted.

Methods

Participants

The details of the below 2 cases have been altered to protect the anonymity of the participants while keeping the functional similarity of the participants' behavioral analysis and the characteristics of the treatment. The participants were recruited via advertisements on social media. They had to have a self-reported T1DM diagnosis, be unsatisfied with their blood glucose, or report other psychological or behavioral problems connected to their diabetes. In telephone interviews, the participants were asked about their diabetes history. The participants were asked about problematic areas with diabetes self-care and life in general. In particular, the participants were asked about specific emotional and behavioral reactions to problematic situations. The interviews were not recorded. See Table 1 for the timeline of the intervention.

Table 1. Timeline of the intervention.

Participant	Assessment	Treatment start	Treatment end
Participant 1 (Mr A)	January 25, 2018	March 06, 2018	April 27, 2018
Participant 2 (Ms B)	April 18, 2018	May 01, 2018	June 23, 2018

Case 1

Mr A was a man in his 40s with a university education. He had self-reported lifestyle problems and an increasing worry about complications. Mr A had found it difficult to engage in physical activity, which he knew was important to his health. He reported that he was very preoccupied with work and that it was challenging to fit physical activity into his schedule. He also found that he ate too much, reporting feelings of hunger even after he had eaten portions of adequate size according to the recommendations of his dietician. Mr A was already experiencing a number of diabetes-related complications (some mild and transient, and others moderately serious and permanent). He reported that he was often preoccupied with worry that the complications would deteriorate and that he would develop additional complications. He reported that it was problematic and a burden to plan the day with consideration of the disease, as well as find time and practical opportunity for self-care activities within daily life. Perfectionistic tendencies were another area that could cause problems, leading to unhelpful thoughts that if Mr A was not able to manage his diabetes perfectly, he might as well give in completely. This could lead to hopelessness and engaging in problematic behaviors, such as eating unhealthy food, reasoning that it did not matter anyway. Worry distracted him from engaging in more enjoyable activities, and could lead to a feeling of helplessness, as well as frustration and anger with the disease, or feeling overwhelmed.

Mr A reported that he had a general idea that he needed to change his behaviors and habits to improve his well-being. However, he had difficulties knowing precisely where to begin, and felt that he needed a structured way forward to fit self-care behaviors into his busy life. Previous approaches to behavioral change had also been hindered by unhelpful thoughts, worry, and perfectionism. We believed that problem solving and exposure would be helpful interventions for Mr A, as this would help him to structure what he already had in mind and help him go against unhelpful emotions to engage in behaviors he felt he needed.

Case 2

Ms B was a woman in her 30s with a university education. She reported high blood glucose levels and fear of hypoglycemia throughout several years. Her fears included worrying about losing consciousness or feeling humiliated in public, losing control, not having anyone there to help her, making serious mistakes, and becoming aggressive toward other people. A particular fear was to experience hypoglycemia during the night, which Ms B felt was a very frightening experience, sharing many characteristics of a panic attack. Furthermore, she reported that she had neglected her medical self-care to some degree when newly diagnosed in her early teens, acting as the disease would simply go away, although she knew that was not a realistic judgement. Stress was a reported problem area, which was accompanied by higher blood glucose. Ms B reported that her blood glucose level was less stable and that she did not find the time for proper self-care when under stress. This could lead to stressful thoughts about her diabetes, increasing stress in a vicious circle. A particularly unhelpful behavior was to eat to keep the blood glucose level high, in order to avoid hypoglycemic episodes during the night. This behavior would ease the fear and anxiety in the short term. In the long run, however, this led to hyperglycemia, with its unpleasant symptoms, such as thirst and fatigue. Likewise, Ms B preferred her blood glucose level to be on the higher side in other situations, such as when attending parties and driving. Another problematic area was that Ms B was afraid of diabetic health care itself. She avoided contacting her hospital diabetes unit as much as possible and reported aversive events where she had been educated in great detail about possible complications. Although the physician had all the best intentions, Ms B reacted with fear. Further mentions of complications made her anxious, and she wished to avoid the subject altogether. Ms B had been in contact with a psychologist about her diabetes before, but she was not satisfied with the intervention, reporting that they mainly talked in general, and that she was told to practice positive thinking, rather than behavioral interventions aimed at specific targets.

Ms B reported that her situation was beginning to be difficult to handle and felt overwhelmed by her fears. Our behavioral analysis showed that her fear had led to many avoidance and safety behaviors, with a negative impact on her blood glucose and general well-being. We hypothesized that it would be useful to have a particular focus on Ms B's fear of hypoglycemia during the treatment. According to our hypothesis, if her fear of hypoglycemia was improved, it would lead to improvement overall.

Ethical Considerations

The exercises in the treatment were discussed with a consultant endocrinologist and were considered safe. All procedures in the study were in accordance with the ethical standards of the regional ethics committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. It was approved on January 04, 2018, by the regional authority Etikprövningsnämnden Stockholm (2017/2278-31). An addendum was approved by the same regional authority on February 12, 2018 (2018/300-32). Another addendum was approved on April 30, 2021, by the Swedish national authority Etikprövningsmyndigheten (2021-02263). Participants provided written informed consent for participation and publication.

Measurements

This study relied on self-reported measures. The measures were filled out as online questionnaires, before and after treatment, and every week in some cases.

Blood Glucose

Glycated hemoglobin (HbA_{1c}) is a well-used biomarker for the blood glucose level over several months [2]. Participants reported their latest known HbA_{1c} value and were encouraged to get a new measurement toward the end of treatment. In what way and how often participants measured their blood glucose level as part of their regular self-care was not controlled for in this study.

Problem Areas in Diabetes

The Problem Areas in Diabetes (PAID) Scale [15] is a scale used to assess areas of diabetes a patient may need guidance for from health care, and to screen for "emotional distress" in patient with diabetes. The PAID Scale consists of 20 items on a scale from 0 ("not a problem") to 4 ("serious problem"). The total score ranges from 0 to 100, and scores over 40 indicate "emotional distress." According to the authors of the scale, people with a score of 40 or higher require attention from their diabetes team, and the score is expected to drop 10-15 points following medical and educational interventions. The scale has been validated and shown to be acceptable in a diabetes population [15]. This scale was measured weekly.

Fear of Hypoglycemia

The Hypoglycemia Fear Survey (HFS) [16] was developed to measure to what degree a patient with diabetes is afraid of hypoglycemia. This scale explores avoidance behaviors and worry in relation to hypoglycemia. It has 23 items with a range from 0 ("never") to 5 ("always"). The total score is 0-92, with a mean score of 25 in a clinical sample. This scale was measured weekly.

Anxiety

Generalized Anxiety Disorder-7 (GAD-7) [17] is commonly used to measure anxiety. It consists of 7 items using a scale from 0 ("not at all") to 3 ("nearly every day"). The total score is 0-21 points. The authors suggest cutoff scores for mild (5), moderate (10), and severe (15) anxiety. The scale has been found to be valid and reliable [17].

Depression

The Patient Health Questionnaire (PHQ-9) [18] consists of 9 items on a scale from 0 ("not at all") to 3 ("nearly every day"), with a total score of 0-27. The authors suggest a cutoff at 5 (possible depression) and 15 (probable depression). The scale has been found to be valid and reliable [18].

Stress Reactivity

The perceived stress scale [19] consists of 14 items on a scale from 0 ("never") to 4 ("very often"), with a total score of 0-40. The authors suggest intervals, with 0-13 indicating low stress, 14-26 indicating moderate stress, and 27-40 indicating high perceived stress. The scale has been shown to be valid and reliable [19].

Credibility/Expectancy

Participants' perception of the credibility of the treatment and their expectation of the final result were assessed with the Credibility/Expectancy Questionnaire (CEQ) [20], using a scale from 1 to 6, with a total score of 0-50. Higher scores suggest better credibility/expectancy. The scale has been found to be valid and reliable. The scale was administered 1 week into treatment.

Treatment Satisfaction

The Client Satisfaction Questionnaire (CSQ) [21] consists of 8 items on a scale from 1 to 4, with a total score of 8-32. The score can be recalculated with a total score from 25 to 100. A higher score suggests a higher satisfaction [21].

Adverse Events

Participants answered questions about whether they had experienced any adverse events, and if so, they were told to elaborate on these experiences. The questionnaire was unpublished.

Subjective Assessment

Participants were asked about subjective blood glucose improvement, using the Subjective Assessment Questionnaire (SAQ), with a 6-point scale ranging from "much declined" to "much improved." This type of scale is often used for other somatic conditions and has been found to be useful and valid [22].

Therapeutic Intervention

The treatment was delivered entirely via the internet, through a secure treatment platform. Participants received homework each week, answering questions about the psychoeducation and doing exercises in their everyday life. Participants received psychoeducation based on material from a previously evaluated CBT group treatment for T1DM [23] and adapted material from an internet-delivered treatment for IBS [9]. The material from the IBS treatment was general information on worry about symptoms and exposure to avoided stimuli, which was easily adapted for T1DM by changing examples. The participants had access to 5 modules of psychoeducation, each approximately corresponding to 10 written A4 pages. The active treatment consisted of established CBT interventions. The main components used were problem solving and exposure. There were also some cognitive interventions, namely information about thinking traps and basic cognitive restructuring, and additional interventions such as assertiveness training and life values. The participants read the same material and tried all interventions. After about half of the treatment, a main focus was chosen according to their specific problem in a discussion with the therapist.

Therapist and Supervisor

The therapist (DK) was a master's student at the time of the study and was trained in CBT. His supervisor for the master's thesis (BL) also acted as a therapy supervisor. The supervisor is a professor of psychology and a licensed psychologist specialized in CBT. The therapist could be reached through text messages in the online platform at any time throughout the treatment. The role of the therapist was to provide guidance and support, provide feedback, and answer questions. The therapist also gave access to the next module after the previous one was completed. Study participants were notified by an automatic text message each time the therapist had made contact via the online platform. If participants were inactive, they were encouraged to continue their assignments via text messages or phone calls.

CBT Model

We mainly used a general cognitive-behavioral model, where we assumed that thoughts, emotions, and behaviors interact and influence each other, which was largely influenced by an exposure-based CBT program for IBS [9]. See Figures 1 and 2 for an overview. This model also assumes that these influence

blood glucose levels directly via behaviors or indirectly via thoughts and emotions that influence behaviors. In some cases, emotions, such as stress, may influence general mood, which may directly influence blood glucose, as suggested by a CBT group treatment for T1DM [4]. Increased awareness may lead to symptom preoccupation, making a person more observant of potential symptoms [9]. We worked from an exposure-based paradigm, where we assumed that humans avoid unwanted experiences that may be aversive short term, but still are important and helpful [9]. Exposing oneself to aversive stimuli or situations decreases anxiety over time, most likely due to extinction of fear [7]. It is possible that symptoms of anxiety can overlap with symptoms of low blood glucose, thus conflating the two [4]. Exposing oneself to feared situations in the presence of anxiety symptoms should, according to this model, decrease fear of hypoglycemic symptoms, and decrease avoidance of feared situations or situations where low blood glucose levels would be unwanted.

The intervention also contained a few approaches that were not necessarily based on this model but did not contradict it, including cognitive restructuring, assertiveness communication skills, and life values. These were components from the previously evaluated CBT group treatment on which this intervention was partially based [4]. These were kept in addition to the main exposure-based model, in order to explore if these could be useful complementary interventions.

Figure 1. General cognitive behavioral therapy model. Adapted from T Anderbro and Susanne Amsberg (unpublished data, 2004) and B Ljótsson and E Hedman-Lagerlöf (unpublished data, 2012).

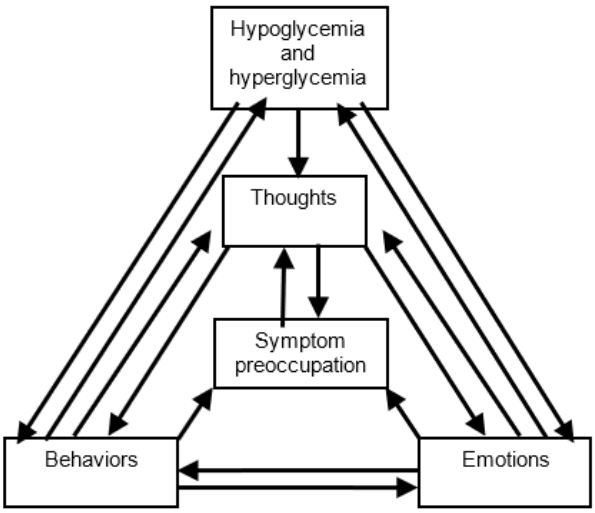
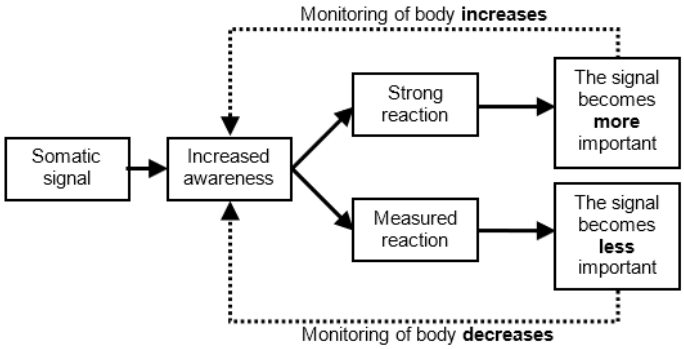


Figure 2. Suggested exposure model for fear of hypoglycemia. Adapted from B Ljótsson and E Hedman-Lagerlöf (unpublished data, 2012).



Content of Treatment Modules

Module 1

The module began with practicalities followed by general education on T1DM. The module also introduced the CBT model and education of general CBT principles. Participants were taught about our proposed interaction of thoughts, emotions, and behaviors. This, in turn, would influence self-care and blood glucose levels. Participants were introduced to the concept of functional behavioral analysis and learned how to identify what prevented them from behaving in accordance with their goals.

Module 2

The module continued the psychoeducation with a focus on stress and negative emotions. The intervention *problem solving* was introduced and presented mainly as an aid for self-care. This was presented as being particularly useful during periods of stress or negative mood, as sometimes self-care behaviors decrease in priority. The potential impacts of negative thoughts, as well as how to identify them were also taught. The participants identified any negative thoughts on their own and discussed how these might interfere with their everyday life, with a focus on diabetes.

Module 3

The module contained psychoeducation about anxiety and worry, and the roles these play in unhelpful behaviors. The participants were educated about the concept of exposure to aversive stimuli, and were provided further information on how to facilitate behavioral change. This module also introduced life values, often used in acceptance and commitment therapy [24], as an additional intervention. The participants considered the various areas in their own life where diabetes could have interfered. The participants were also encouraged to plan and work toward better integrating diabetes self-care into the life they wished to lead. As an additional skill, the participants were taught about assertiveness training.

Module 4

The module was mainly dedicated to exposure exercises. These were based on the problematic behaviors and situations participants had analyzed earlier in treatment. With the help of their therapist, the participants constructed an exposure hierarchy and used that to plan exposure exercises. If exposure to aversive stimuli was not appropriate, the participant would continue to work with behavioral change, for instance, with the help of problem solving and other helpful skills previously taught.

Modules 5-7

There was no additional information at this point in treatment. The participants continued to practice and were guided by their therapist as earlier. If a participant had a slower pace than 1 module per week, this could be a time to catch up.

Module 8

In the final module, the participants constructed a maintenance plan, based on the skills they used the most.

Results

Overview

The overall results in the 2 cases are presented in Table 2.

Table 2. Results in the 2 cases.

Variable	Participant 1 (Mr A)		Participant 2 (Ms B)	
	Pretreatment	Posttreatment	Pretreatment	Posttreatment
HbA_{1c}[a] (mmol/mol)	60	50	62	—[b]
PAID[c] Scale score	50	28	92	53
HFS[d] score	8	7	64	40
GAD-7[e] score	4	4	12	13
PHQ-9[f] score	8	4	8	18
PSS[g] score	32	16	35	34
SAQ[h] score	N/A[i]	6	N/A	4
AEs[j], n	N/A	0	N/A	0
CSQ[k] score	N/A	75	N/A	100

[a]HbA_{1c}: glycated hemoglobin.

[b]Not reported.

[c]PAID: Problem Areas in Diabetes.

[d]HFS: Hypoglycemia Fear Scale.

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

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

[g]PSS: Perceived Stress Scale.

[h]SAQ: Subjective Assessment Questionnaire.

[i]N/A: not applicable.

[j]AEs: adverse events.

[k]CSQ: Client Satisfaction Questionnaire.

Case 1

Problem solving was the primary intervention for Mr A. He was able to try out solutions to numerous problems. For example, he managed to fit squash into his schedule, and had a goal to play once a week with his friends and to walk 8000-12,000 steps per day, in order to get an adequate amount of daily exercise. He made a rule to only take one portion of food for each meal, and he found it effective to go against his worry to do these physical activities, even though they reminded him of his complications. This could possibly be considered a form of exposure to aversive stimuli. Before treatment, Mr A had a score of 50 in the problem areas on the diabetes scale. After treatment, the score had dropped to 28, meaning a decrease of 22 points. Answering a postmeasurement question, Mr A stated that his blood glucose greatly improved. Another

improvement was clinically significantly decreased stress according to the perceived stress scale, with a score of 32 before treatment and 16 after treatment. Fear of hypoglycemia did not seem to be a problematic area for this participant, beginning with a low grade of 8 points compared to an average of 25 in a clinical sample [4]. The value increased somewhat to about 20 through the course of treatment, but gradually decreased again to a score of 7 after treatment. Mr A reported 75 out of 100 points in the CSQ, indicating a high satisfaction with treatment. He stated that he was very satisfied with the treatment overall. Mr A also gave feedback on the psychoeducation material, feeling that the material was too long, and he wished for it to be more concise and to the point.

After treatment, Mr A reported his long-term blood glucose levels to us, which showed a decrease in HbA_{1c} from 60 to 50 mmol/mol, which is not only a clinically significant change in the desired direction, but also within the general goal value of <52 mmol/mol in Swedish diabetes care [25]. Mr A reported that it was the best HbA_{1c} value in his clinical history and that he attributed this to the treatment. It is important to remember that a decrease in HbA_{1c} does not necessarily equal a healthier lifestyle but could be a misleading finding because of repeated hypoglycemic episodes. This was not the case for this participant, however, as neither increased frequency of hypoglycemia nor any other adverse events were reported upon questioning. There were still some hyperglycemic episodes, meaning that there was room for continued decrease in hyperglycemia, but the blood glucose levels were quite satisfactory overall, according to the Swedish guidelines.

Overall, considering Mr A's satisfaction with the treatment, the decreased problematic areas associated with his diabetes condition, and the decreased blood glucose levels, the treatment appeared to be successful.

Case 2

Ms B's treatment consisted mainly of exposure exercises aimed at her fear of hypoglycemia and associated avoidance behaviors. Another important area covered was her fear of visiting the diabetes hospital unit, which could be considered exposure to the emotion of shame, as hospital visits reminded her of her perceived failures as a patient with diabetes. She gradually approached health care, attending a lecture at her diabetes clinic about complications and their prevention. She reported that this increased her anxiety severely, but she was able to stay in the situation until it reduced. As a result of this, Ms B reported that she was able to form a better relationship with her diabetes care staff, which could prove beneficial to her future health. Another important area was the safety behavior of eating extra food before bedtime to decrease the risk of night-time hypoglycemia. Ms B gradually decreased the amount she ate and was eventually able to overcome this safety behavior, with potentially significant results for her overall blood glucose levels. Initially, she reported increased anxiety and difficulty falling asleep. She also reported waking in the middle of the night, feeling anxious, and checking her blood glucose levels. After less than a week, however, she reported that her anxiety started to subside and that she was able to sleep uninterrupted again. Eventually, she was able to stop the safety behavior of excessive eating before bedtime.

Ms B reported a significant decrease in diabetes-specific problems, with a decrease from 92 to 53 on the PAID Scale. This is still a high number, but the decrease was much greater than the expected decrease after education only, according to the authors of the scale. Relevant to her main problems, Ms B reported a significantly lower fear of hypoglycemia of 40 points on the hypoglycemia fear scale compared with 64 points before treatment, which was a clinically significant decrease, and she rated her blood glucose levels as somewhat improved. Ms B reported a score of 100 out of 100 points on the client satisfaction scale, indicating excellent treatment satisfaction. Unfortunately, we did not know if the treatment had any effects on Ms B's blood glucose or HbA_{1c} levels, as she did not report a posttreatment HbA_{1c} value. Nevertheless, we consider the treatment successful because the treatment was able to target her fear of hypoglycemia using exposure to feared stimuli. It should be noted that her depression rating increased after the treatment, as measured by the patient health questionnaire. When asked, she did not think that this was a negative effect of the intervention but attributed her decreased mood to external factors.

Adverse Events

No adverse events or other unexpected events were reported in either of the cases.

Discussion

Summary

In this case study, the online guided self-help CBT treatment of 2 participants has been described. The aim was to test an early version of this treatment, in order to revise and improve the protocol. The 2 participants were treated using the same treatment manual, and they read every module and tried all interventions. However, they both had a personalized focus. In case 1, the participant Mr A chose to focus on problem solving in the later part of the treatment. After treatment, he showed improved blood glucose and other relevant improvements. In case 2, the participant Ms B. chose to focus on exposure in the later part of treatment. She showed a significant decrease in her fear of hypoglycemia and other relevant improvements.

Principal Findings

In the 2 cases, participants had favorable outcomes on their most relevant measurements. The participants completed all modules and were highly satisfied with their treatment. Mr A felt mainly helped by problem solving, which is logical, as this intervention was most suitable to help him accomplish the lifestyle changes desired. Conversely, Mr A found it difficult and redundant to construct exposure exercises, as he did not consider himself to be driven by fear or anxiety. Instead, he preferred to focus on structured problem solving. We believe, however, that there were some elements of exposure to other avoided emotions involved, such as worry and discomfort. Ms B focused on exposure for her fear of hypoglycemia, which was much improved, as well as her general problem areas in diabetes.

Limitations

The results from this case study must be carefully considered as there are limitations. This is a report on only 2 cases, and the participants may have had personal characteristics and environmental factors that could explain their improvements. Furthermore, we were not able to control for attention. We do not know if their improvements are due to the treatment or the fact that someone kept a watchful eye on them. In fact, Mr A stated that he partially attributed attention from the therapist as a factor for his improvement. Finally, we did not receive a report on Ms B's blood glucose level. This is significant missing data, which we have no easy way of retrieving.

Comparison With Prior Work

To our knowledge, this is the first study that tentatively explored an online CBT intervention with an exposure model for T1DM. The outcomes are consistent with those in earlier studies on CBT and T1DM, as CBT has been shown to improve blood glucose and mental health [4,12]. The experiences of these participants indicate that a CBT treatment with multiple interventions may be useful. For example, an entirely exposure-focused treatment would have been less relevant for Mr A. The results of Ms B are in line with the findings in few previous studies that have evaluated exposure to aversive stimuli as an intervention for fear of hypoglycemia [4,8].

Conclusions

These cases provide some insights on how an online-delivered T1DM CBT program could look, and preliminarily suggest that the use of both problem solving and exposure could be useful in CBT for T1DM. In accordance with the feedback of Mr A, the treatment can be further adapted to decrease the amount of text and introduce learning exercises through examples. These changes may improve the treatment and make the information focused and relatable. Looking forward, we suggest a feasibility study with a larger group of participants, to examine the safety, acceptability, and approach of the intervention, as well as its preliminary effects.

Patient Perspective

Mr A was very satisfied with his experience overall and felt that his condition was much improved. As mentioned, he had some critique of the education material provided, but felt that the contact with the therapist made up for it. Ms B was very satisfied and felt that her condition was much improved. She especially appreciated the applied and concrete nature of CBT.

Acknowledgments

We would like to thank Susanne Amsberg, PhD and Therese Anderbro, PhD for allowing us access to their cognitive behavioral therapy–based group treatment manual for type 1 diabetes mellitus, which served as valuable inspiration for the construction of the online intervention. We also would like to thank Eva Toft, MD, PhD for excellent medical expertise.

Authors' Contributions

DK: study design, drafting of the manuscript, and analysis; BL: study design and contributions to the manuscript; MB: contributions to the design of the treatment manual and contributions to the manuscript; NL: contributions to the manuscript; MK: contributions to the manuscript.

Conflicts of Interest

BL owns shares in DahliaQomit, which specializes in online services for symptom assessment outside the submitted work, and licenses a cognitive behavioral treatment manual for irritable bowel syndrome, with royalties paid from Pear Therapeutics. The other authors have no conflicts of interest to declare.

References

1. Amsberg S. Health Promotion in Diabetes Care - Studies on adult type 1 diabetes patients. Open Archive. Stockholm, Sweden: Karolinska Institutet; 2008. URL: https://openarchive.ki.se/xmlui/bitstream/handle/10616/40159/thesis.pdf [accessed 2022-05-14]
2. Mazze RS, Strock ES, Simonson GD, Bergenstal RM. Staged Diabetes Management: A Systematic Approach. Hoboken, NJ: Wiley-Blackwell; 2012.
3. Hillson R. Diabetes Care: A Practical Manual. Oxford, UK: Oxford University Press; 2012.
4. Anderbro T. Behaviour change intervention and fear of hypoglycaemia in type 1 diabetes. Open Archive. Stockholm, Sweden: Karolinska Institutet; 2012. URL: https://openarchive.ki.se/xmlui/bitstream/handle/10616/41244/Thesis_Therese_Anderbro.pdf [accessed 2022-05-14]
5. Wild D, von Maltzahn R, Brohan E, Christensen T, Clauson P, Gonder-Frederick L. A critical review of the literature on fear of hypoglycemia in diabetes: Implications for diabetes management and patient education. Patient Educ Couns 2007 Sep;68(1):10-15. [doi: 10.1016/j.pec.2007.05.003] [Medline: 17582726]
6. Safren SA, Gonzalez JS, Soroudi N. Coping with Chronic Illness: A Cognitive-Behavioral Therapy Approach for Adherence and Depression, Therapist Guide. Oxford, UK: Oxford University Press; 2007.

7. Sewart AR, Craske MG. Inhibitory learning. In: Abramowitz JS, Blakey SM, editors. Clinical handbook of fear and anxiety: Maintenance processes and treatment mechanisms. Washington, DC, USA: American Psychological Association; 2020:265-285.
8. Boyle S, Allan C, Millar K. Cognitive-behavioural interventions in a patient with an anxiety disorder related to diabetes. Behav Res Ther 2004 Mar;42(3):357-366. [doi: 10.1016/j.brat.2003.11.006] [Medline: 14975775]
9. Ljótsson B, Falk L, Vesterlund AW, Hedman E, Lindfors P, Rück C, et al. Internet-delivered exposure and mindfulness based therapy for irritable bowel syndrome--a randomized controlled trial. Behav Res Ther 2010 Jun;48(6):531-539. [doi: 10.1016/j.brat.2010.03.003] [Medline: 20362976]
10. Särnholm J, Skúladóttir H, Rück C, Klavebäck S, Ólafsdóttir E, Pedersen SS, et al. Internet-Delivered Exposure-Based Therapy for Symptom Preoccupation in Atrial Fibrillation: Uncontrolled Pilot Trial. JMIR Cardio 2021 Mar 02;5(1):e24524 [FREE Full text] [doi: 10.2196/24524] [Medline: 33650972]
11. Bonnert M, Andersson E, Serlachius E, Manninen I, Bergström SE, Almqvist C. Exposure-based cognitive behavior therapy for anxiety related to asthma: A feasibility study with multivariate baseline design. Scand J Psychol 2020 Dec;61(6):827-834. [doi: 10.1111/sjop.12674] [Medline: 32706124]
12. Young-Hyman D, de Groot M, Hill-Briggs F, Gonzalez JS, Hood K, Peyrot M. Response to Comment on Young-Hyman et al. Psychosocial Care for People With Diabetes: A Position Statement of the American Diabetes Association. Diabetes Care 2016;39:2126-2140. Diabetes Care 2018 Mar;41(3):e33-e34 [FREE Full text] [doi: 10.2337/dci17-0037] [Medline: 29463672]
13. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11 [FREE Full text] [doi: 10.1002/wps.20083] [Medline: 24497236]
14. Yang X, Li Z, Sun J. Effects of Cognitive Behavioral Therapy-Based Intervention on Improving Glycaemic, Psychological, and Physiological Outcomes in Adult Patients With Diabetes Mellitus: A Meta-Analysis of Randomized Controlled Trials. Front Psychiatry 2020 Jul 28;11:711 [FREE Full text] [doi: 10.3389/fpsyt.2020.00711] [Medline: 32848906]
15. Polonsky WH, Fisher L, Earles J, Dudl RJ, Lees J, Mullan J, et al. Assessing psychosocial distress in diabetes: development of the diabetes distress scale. Diabetes Care 2005 Mar;28(3):626-631. [doi: 10.2337/diacare.28.3.626] [Medline: 15735199]
16. Cox DJ, Irvine A, Gonder-Frederick L, Nowacek G, Butterfield J. Fear of hypoglycemia: quantification, validation, and utilization. Diabetes Care 1987;10(5):617-621. [doi: 10.2337/diacare.10.5.617] [Medline: 3677982]
17. 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]
18. 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]
19. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
20. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. J Behav Ther Exp Psychiatry 2000 Jun;31(2):73-86. [doi: 10.1016/s0005-7916(00)00012-4] [Medline: 11132119]
21. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. J Behav Ther Exp Psychiatry 2000 Jun;31(2):73-86. [doi: 10.1016/s0005-7916(00)00012-4] [Medline: 11132119]
22. Camilleri M. Editorial: is adequate relief fatally flawed or adequate as an end point in irritable bowel syndrome? Am J Gastroenterol 2009 Apr;104(4):920-922 [FREE Full text] [doi: 10.1038/ajg.2009.20] [Medline: 19293789]
23. Amsberg S, Anderbro T, Wredling R, Lisspers J, Lins P, Adamson U, et al. A cognitive behavior therapy-based intervention among poorly controlled adult type 1 diabetes patients--a randomized controlled trial. Patient Educ Couns 2009 Oct;77(1):72-80. [doi: 10.1016/j.pec.2009.01.015] [Medline: 19297117]
24. Hayes SC, Wilson KG, Gifford EV, Follette VM, Strosahl K. Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. Journal of Consulting and Clinical Psychology 1996 Dec;64(6):1152-1168. [doi: 10.1037/0022-006x.64.6.1152]
25. Nationella riktlinjer för diabetesvård Stöd för styrning och ledning. Swedish National Board of Health and Welfare. 2018. URL: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/nationella-riktlinjer/2018-10-25.pdf [accessed 2021-08-05]

Abbreviations

CBT: cognitive behavioral therapy
CSQ: Client Satisfaction Questionnaire
HbA_{1c}: glycated hemoglobin
IBS: irritable bowel syndrome
PAID: Problem Areas in Diabetes
T1DM: type 1 diabetes mellitus

Edited by A Mavragani; submitted 30.08.21; peer-reviewed by E Bernstein, H Alyami, K Matsumoto; comments to author 12.11.21; revised version received 04.02.22; accepted 03.05.22; published 13.07.22.

Please cite as:
Kern D, Ljótsson B, Bonnert M, Lindefors N, Kraepelien M
Online Guided Self-help Cognitive Behavioral Therapy With Exposure to Anxiety and Problem Solving in Type 1 Diabetes Mellitus: Case Study
JMIR Form Res 2022;6(7):e32950
URL: https://formative.jmir.org/2022/7/e32950
doi:10.2196/32950
PMID:35830220

Original Paper

Implementation of a Statewide Web-Based Caregiver Resource Information System (CareNav): Mixed Methods Study

Heather M Young[1], RN, PhD; Janice F Bell[1], MN, MPH, PhD; Orly Tonkikh[1], RN, PhD; Tina R Kilaberia[1], MSW, PhD; Robin L Whitney[2], RN, PhD; Jennifer M Mongoven[1], MPH; Benjamin M Link[1], BS; Kathleen Kelly[3], MPA

[1]Family Caregiving Institute, Betty Irene Moore School of Nursing, University of California Davis, Sacramento, CA, United States

[2]The Valley Foundation School of Nursing, San Jose State University, San Jose, CA, United States

[3]Family Caregiver Alliance, San Francisco, CA, United States

Corresponding Author:
Heather M Young, RN, PhD
Family Caregiving Institute
Betty Irene Moore School of Nursing
University of California Davis
Betty Irene Moore Hall
2570 48th St
Sacramento, CA, 95817
United States
Phone: 1 541 261 0395
Email: hmyoung@ucdavis.edu

Abstract

Background: With the aging population, family caregivers provide increasingly complex and intense care for older adults and persons with disabilities. There is growing interest in developing community-based services to support family caregivers. Caregiving occurs around the clock, and caregivers face challenges in accessing community-based services at convenient times owing to the demands of care. Web-based resources hold promise for accessible real-time support. CareNav (TM), a caregiver resource information system, is a web-based platform designed to support real-time universal caregiver assessment, a record of client encounters, development of a care plan, tailored information and resource content, access to web-based caregiver resources, the capacity to track service authorization and contracts, and secure communications. The assessment includes needs and health conditions of both the care recipient and caregiver; current resources; and priorities for support, information, and referral. In 2019, the California Department of Health Care Services funded the 11 nonprofit California Caregiver Resource Centers (CRCs) to expand and improve family caregiver services and enhance CRC information technology services. Deployment of a statewide information system offered a unique opportunity to examine structures and processes facilitating implementation, providing feedback to the sites as well as lessons learned for similar projects in the future.

Objective: The aim of this paper was to describe the statewide implementation of the comprehensive CareNav system using the Consolidated Framework for Implementation Research as an organizing structure for synthesizing the evaluation.

Methods: This mixed methods study used two major approaches to evaluate the implementation process: a survey of all staff who completed training (n=82) and in-depth qualitative interviews with 11 CRC teams and 3 key informants (n=35). We initially analyzed interview transcripts using qualitative descriptive methods and then identified subthemes and relationships among ideas, mapping the findings to the Consolidated Framework for Implementation Research.

Results: We present findings on the outer setting, inner setting, characteristics of the intervention, characteristics of the staff, and the implementation process. The critical elements for success were leadership, communication, harmonization of processes across sites, and motivation to serve clients in more accessible and convenient ways.

Conclusions: These findings have implications for technology deployment in diverse community-based agencies that aspire to enhance web-based services.

(JMIR Form Res 2022;6(7):e38735) doi:10.2196/38735

KEYWORDS

online assessment; caregiver; technology implementation; Consolidated Framework for Implementation Research; CFIR

Introduction

Background

In the United States, approximately 20% of adults provide unpaid care to a family member or friend [1]. Family caregivers provide the vast majority of long-term care (valued at US $470 billion annually), eclipsing annual governmental spending on long-term care at US $430 billion [2]. Caregiving is often a commitment spanning years and a variety of domains including personal assistance, instrumental aid and emotional support, care coordination, and managing chronic health conditions. With shortened hospital stays, family caregivers now perform complex health care including medical or nursing tasks previously within the purview of health care professionals [3]. Despite the common experience of caregiving, few individuals are prepared for the demands of the role [4]. Many people experience strain, depression, loneliness, deterioration in their own health, and financial distress in the course of providing care [1]. Even so, most are unaware of the existing resources to support them.

Despite their vital role in optimizing function and health for older adults, caregivers are relatively invisible in health care. Caregivers provide valuable information about the person receiving care, yet rarely do health care professionals assess the capacity, readiness, or mental health of caregivers to provide care during routine health encounters [3,5].

Web-Based Supports for Caregivers

In the absence of routine caregiver assessment and support in the clinical setting, there has been growing interest in developing community-based services to support family caregivers. Caregiving occurs around the clock, and caregivers face challenges in accessing community-based services at convenient times owing to the demands of care. Web-based resources hold promise for accessible real-time support, and targeted interventions have been developed for specific audiences such as caregivers for persons with cancer [6], dementia [7-10], and other chronic conditions [11-13].

Most web-based supports emphasize one element such as psychoeducational offerings or stress management. Despite quality issues with many studies, the positive outcomes of targeted interventions have included improved self-efficacy, improved self-esteem, and less strain [14]. Missing from existing web-based interventions are a global assessment of the needs of the caregiver and care recipient and the ability to respond to the priority of the caregiver at the time, whether it is for information, referral to specific local services, or counseling.

Studies on web-based caregiver support have demonstrated 3 tendencies. First, caregivers of persons with dementia are overrepresented because of the demanding nature of caregiving related to issues with memory, thinking, and behavior. Many studies on web-based support for dementia caregivers, such as iSupport [10], focus on psychoeducational interventions that help caregivers cope with stress, anxiety, caregiving burden, quality of life, awareness of stressors and needs, and caregiving competence. Overall, studies of such psychoeducational interventions have reported improvements in the well-being and preparedness of caregivers [15-17]. However, the effect size is medium-small even in randomized controlled trials and quasi-experimental studies [15]. Significant variations in content, structure, outcome measures, and intervention duration further prevent cogent conclusions within web-based support studies for caregivers [18].

Second, the heterogeneity of web-based support types underlies varying results. Studies on web-based support for those caring for persons with cancer [19], posttraumatic stress disorder [20], and psychosis [21,22] offer findings that reflect varied web-based support needs for different caregivers. Disease type, intervention type, dosage (amount of time spent on the web), and duration affect the effectiveness, feasibility, and quality of interventions [23]. The navigation and intuitiveness of web-based sessions may depend on the distinct needs of caregiver subgroups [22].

Third, the breadth and scope of caregiver support in the web-based modality affords subgroup-specific knowledge rather than exhaustive knowledge on implementation strengths and challenges. For example, in an implementation study of a video health technology intervention to improve self-care of caregivers of persons with heart failure, Hirschman et al [24] found adaptation challenges related to hardware, software, and network connectivity. In a behavioral intervention for dementia caregivers that had significant improvements for caregivers, Nichols et al [16] discerned as important the clinical success, leadership and staff support, and ongoing need for modifications while maintaining fidelity, linkage to the organizational context, and fiscal health. In mapping directions for research, Lindeman et al [25] underscored, among others, matters of equity, inclusion, and access; privacy and security; and the influence of political and regulatory factors on interoperability. To date, research has not examined broad-based web-based resources that include caregivers involved across different health conditions and provide a full array of supports, from information and referral to individual and group support, to legal and respite service provision. The aim of this paper was to describe the statewide implementation of the comprehensive CareNav (TM) system using the Consolidated Framework for Implementation Research (CFIR) as an organizing structure for synthesizing the evaluation.

Supporting Caregivers in California

CareNav, a caregiver resource information system, is a web-based platform designed to support interactive universal caregiver assessment, a record of client encounters, development of a care plan, tailored information and resource content, access to web-based caregiver resources, capacity to track service authorization and contracts, and secure communications. The assessment includes both care recipient and caregiver needs and health conditions; current resources; and priorities for support, information, and referral. The assessment can be self-administered on the web or administered by a staff member over the phone or in-person. After assessment, a staff consultant meets with the caregiver to prioritize and develop a care plan that might include contracting services such as respite, referral to educational offerings or a support group, or vouchers for legal aid or counseling.

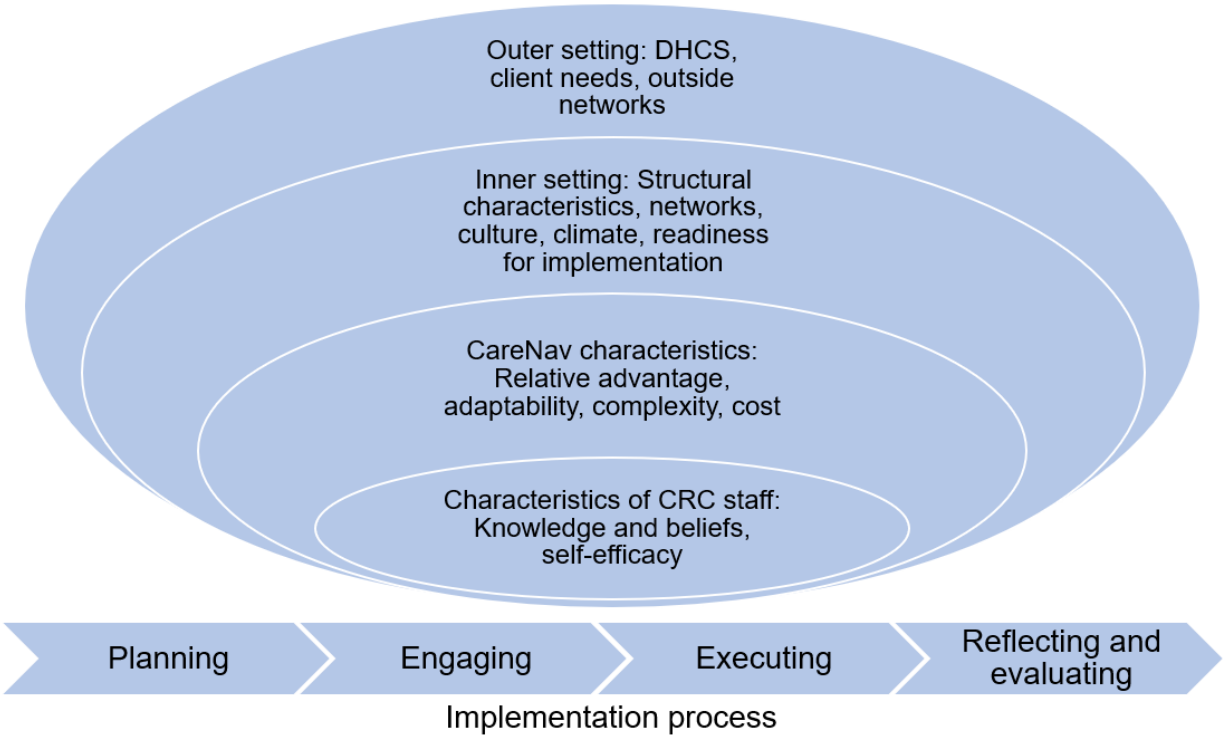
The Family Caregiver Alliance (FCA), one of the 11 California Caregiver Resource Centers (CRCs; California State System of Support for Caregivers), pioneered CareNav with private funding and deployed this system across 3 CRCs that served as pilot sites. In 2019, the California Department of Health Care Services funded 11 nonprofit CRCs to expand and improve family caregiver services and enhance CRC information technology services, deploying CareNav as a common data set across the state over a 3-year period (2019-2022).

The FCA led the implementation team for CareNav in partnership with the technology developer Quality Process (QP). The FCA contracted with the University of California Davis Family Caregiving Institute to conduct an evaluation of the implementation process as well as an analysis of the statewide data to determine program effectiveness and quality improvement opportunities. The implementation began in January 2020, and the system was fully deployed by September 2020.

Organizing Framework

Implementation evaluation was guided by the health-focused CFIR [26]). The CFIR includes five domains that organize our mapping of the results (Figure 1): the outer setting, the inner setting, the intervention characteristics, staff characteristics, and the process of implementation. The outer setting includes the social and economic context within which the statewide CRC system resides, particularly considering relationships with outside organizations, client needs, and the effects of the COVID-19 pandemic. The inner setting refers to the levels and characteristics of the organization, in this case the CRC statewide system, focusing on structural characteristics, culture, the implementation climate, networks and communications, and readiness for implementation. The intervention characteristics acknowledge the perception of the stakeholders about the key attributes of CareNav including relative advantage, adaptability, complexity, and cost. The characteristics of individuals recognize staff knowledge and attitudes toward CareNav as well as beliefs about the capabilities required for the implementation and stage of change. The implementation process considers 4 major aspects: planning, engaging, executing, and reflecting and evaluating [26].

Figure 1. The Consolidated Framework for Implementation Research model for CareNav implementation. CRC: Caregiver Resource Center; DHCS: Department of Health Care Services.



Methods

Design Overview

This mixed methods study used two major approaches to evaluate the implementation process: in-depth interviews with key informants at the CRCs, and surveys of all staff who completed training. The in-depth interviews explored all aspects of the implementation process from multiple perspectives, while the surveys were used to characterize the readiness and self-efficacy of the staff implementing the system. This approach will enable the evaluation team to assess readiness and satisfaction over time using quantitative scores while developing a deeper understanding of the dynamics of change and appreciation of multiple perspectives afforded by qualitative interviews.

Qualitative Interviews

The evaluation team conducted focus group interviews with teams of each CRC, and individual interviews with key

informants from the implementation team. All current leaders and staff of the 11 CRCs were eligible to participate in this study. A focus group was established for each CRC to include all interested staff from the same site. The interviews elicited perspectives on implementation and training activities, including perceptions of benefits and concerns regarding CareNav. We used a semistructured interview guide for both focus groups and individual interviews asking about the implementation process, challenges and facilitators, anticipated system and client outcomes, and satisfaction with the process. Owing to the widespread implications of the COVID-19 pandemic on service delivery, we asked all participants about how the COVID-19 pandemic affected both the implementation process and experiences of their clients as caregivers. Interviews were conducted using Zoom, and audio was recorded with the consent of the participants and transcribed.

Pretraining and Posttraining Surveys

The evaluation team designed a pretraining and posttraining readiness survey to determine readiness, preparation, and confidence regarding the implementation process and to identify self-efficacy and perceived benefits and concerns before and after training. We invited all staff to complete surveys before and immediately after a full day of formal training on CareNav.

The pretraining readiness survey included 10 items rated on a 5-point scale, where 1 represents the most positive response. Cronbach α for the current sample was .83. The survey also assessed whether the participants are familiar with CareNav, know its purpose, and how to do an intake and assessment. Open-ended questions identified benefits and concerns about CareNav.

The posttraining survey reassessed participants' knowledge, preparedness for implementation, confidence, and knowing where to get help with CareNav. The posttraining survey also assessed whether the training met participants' needs and their willingness to take actions that could support CareNav implementation, such as encouraging staff or coworkers to use CareNav and ensuring new staff members are educated on how to use CareNav.

Data Analysis

The transcribed interviews were imported into the Dedoose qualitative data analysis software. Qualitative descriptive methods were used to analyze the transcripts and open-ended responses to the surveys [27,28]. Two members of the research team (HMY and TRK) reviewed the transcripts and developed initial codes and definitions. They independently coded the same transcript and then compared coding decisions, refined definitions, and arrived at a consensus about coding and documenting changes to the definitions. All the transcripts were subsequently coded into agreed-upon general categories. A second analysis (by HMY and OT) identified the subthemes and relationships among ideas, mapping the findings to the CFIR.

Quantitative data were analyzed using descriptive statistics. To enable meaningful interpretation and visualization, the 10-item readiness scale was recoded to a 5-point scale in which a higher score represents better readiness. Pretraining and posttraining scores were compared using paired 2-tailed *t* tests for continuous variables and McNemar tests for dichotomous variables. Quantitative analyses were performed using the SPSS statistical package (version 27; IBM Corporation).

Ethical Considerations

This study was determined to be exempt from ethics approval by the UC Davis institutional review board. We collected no identifying information about the participants in the survey and focus groups. Participants were informed about the purpose of the study and the voluntary nature of participation, providing assent by completing the survey or continuing with the recorded Zoom interview.

Results

Participants

Pretraining and Posttraining Readiness Surveys

In total, 86 staff members completed training, with 82 (95%) participants contributing pretraining data and 56 (65%) participants contributing posttraining data. The staff included CRC directors, family consultants who interact with caregiver clients and provide resources and support, and analysts who manage client and financial data. The surveys were anonymous, with no identifying data of the staff members. The results of readiness and self-efficacy are presented in subsequent sections within the CFIR model as characteristics of the staff.

Qualitative Interviews

Between May and August 2020, we conducted 11 focus group interviews (ranging in size from 2 to 6 participants) and 5 focused individual interviews with key informants, totaling 35 participants. Participants represented all 11 sites and the implementation team and included all roles (directors, clinical, and technical staff). To protect the privacy of the participants, we did not collect demographic data associated with these interviews.

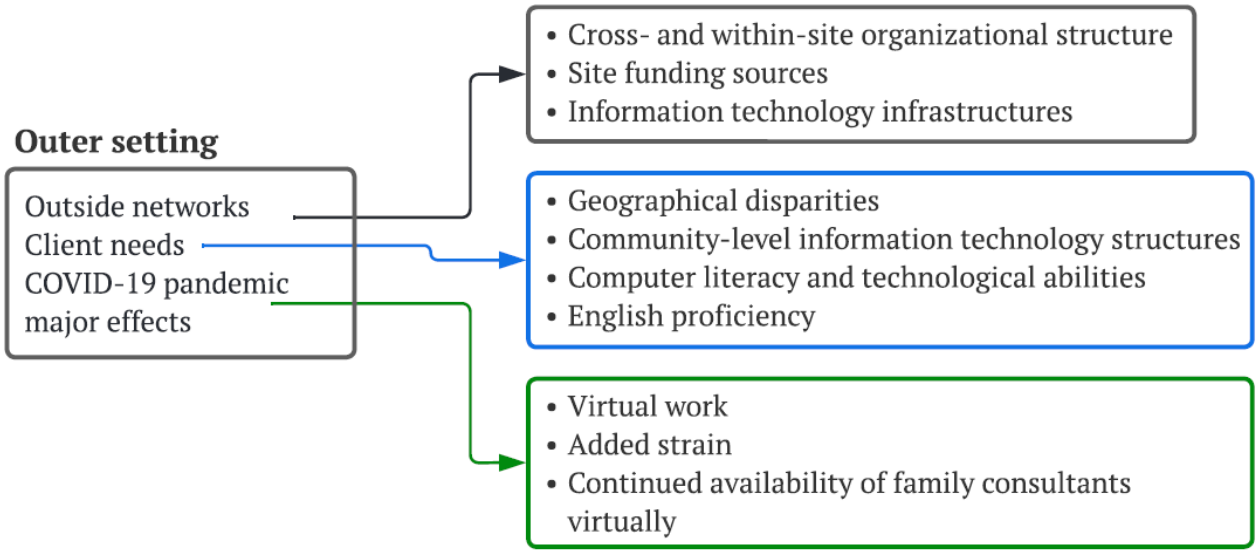
Evaluation of Implementation

We mapped the findings of this evaluation to the following CFIR domains: outer setting, inner setting, characteristics of the intervention, staff characteristics, and implementation process. The specific findings related to these domains have been expanded upon in the subsequent sections.

Outer Setting

Interviews across CRC sites revealed site-level variability and heterogeneity in relationships and networking with external organizations as well as the diversity of client characteristics and needs (Figure 2). Implementation occurred during the COVID-19 pandemic, shaping additional contextual factors.

Figure 2. Outer setting.



Outside Network

Some CRCs exist as standalone organizations. More complex sites are embedded or hosted in larger health care systems with their own information technology platforms, which requires additional efforts to integrate CareNav and to be compliant with additional health system privacy and security policies. Moreover, the funding sources and constellations of services vary between sites and influence both documentation and administrative requirements. This challenged the transition from the existing local databases to the CareNav platform. For example:

> *Because it's a hospital it has security really well locked down. And so, the website that I need to access to see if the data is correct in their initial upload in our system is blocked. So, I'm working within my own organization to try to get that website unblocked.*

Client Needs

Geographic disparities are reflected in rural sites serving communities that lack the adequate technology structures and broadband connection necessary for reliable access to an internet platform. Clients are also diverse in their English proficiency and their capacity to use technology. Several participants mentioned low computer literacy and lack of technological and internet safety skills as barriers for older caregivers to use technology. Nonetheless, some of the staff reported being surprised by the amount of participation by clients whom they assumed would not ordinarily use technology. This observation was partially attributed to the consequences of the COVID-19 pandemic:

> *But I think a lot of them are open to technology. I think this pandemic has forced that issue where a lot of family caregivers are using more technology now...And I'm surprised as to the ones that I thought wouldn't know how to use technology. The ones that before COVID there were a lot that I never, never thought that they could do it as a video with me. I never thought that because I just, you generalize people...because of age...a lot of them are facetiming family, they are video conferencing with their family care navigator...so we just have to start thinking differently as far as how we get them used to CareNav.*

COVID-19 Pandemic: Major Effects

The COVID-19 pandemic, which emerged simultaneously with the implementation kickoff, amplified the existing risks and threats to the well-being of caregivers. Family consultants reported higher levels of stress, more financial concerns, increased housing and food insecurity, job loss, and escalating costs of caregiving supplies (such as gloves and masks) among their clients. The strain was exacerbated by loss of support resources, including availability of adult day care, home help, in-person support groups, and a reluctance to consider assisted living or skilled nursing alternatives at this time as residential care facilities were widely reporting COVID-19 outbreaks and, in some cases, higher mortality rates for older persons. Many caregivers faced additional home demands, caring for multiple family members and children being schooled at home. The pandemic increased their sense of isolation, with family consultants reporting more symptoms of depression and anxiety among their clients. Finally, caregiver health was further compromised by the inability to quarantine when a family member was ill, thus increasing the exposure of the caregiver. Caregivers also expressed reluctance to access formal services owing to fear of contagion or not having the time to devote to their own health:

> *We've seen kind of that increased risk, I think, across the board because a lot of our support systems are not available right now. [...] I think what most impacts caregivers, or can impact their health overall and wellbeing, is what level of support they have available to them.*

The pandemic created a particularly challenging climate for the implementation of a new major system**.** Additional strain was placed on staff members who were already facing a learning

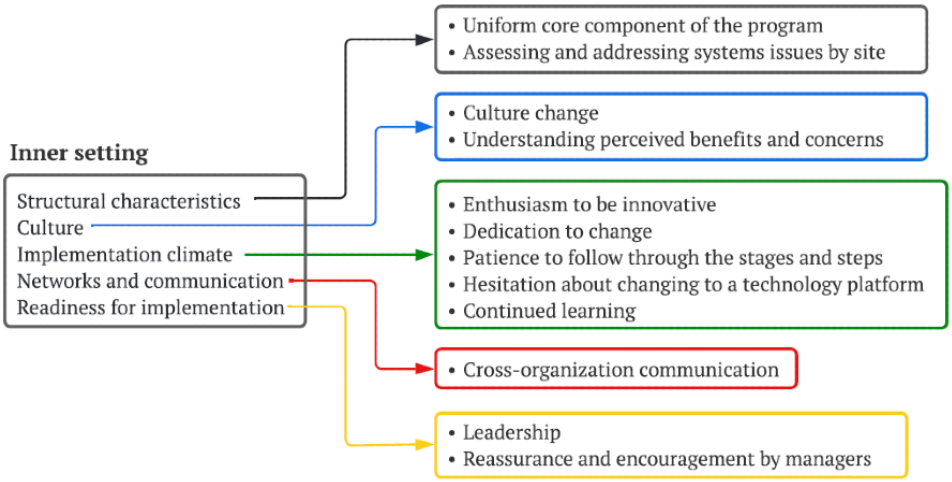
curve to execute the new system, as well as dramatic changes in their workflow. Simultaneously, it accelerated the need for a web-based platform and rapidly demonstrated the usefulness and feasibility of the web-based environment (eg, continued availability of family consultants virtually). Interview participants highlighted how the pandemic demonstrated the benefits of CareNav:

> *I think the Coronavirus kind of shows how helpful it is to have everything online. There really hasn't been too much of a disruption to the services we're able to provide our caregivers. I think that's been a huge plus to be able to see how nicely it can transition when things do happen, that we're still able to access what we need.*

Inner Setting

The inner setting for CareNav implementation is the state-level CRC system and 11 individual CRCs (Figure 3).

Figure 3. Inner setting.



Structural Characteristics

California launched a CRC system in 1984 with the enactment of the Comprehensive Act for Families and Caregivers of Brain-Impaired Adults, thereby establishing a network of support for caregivers. At present, the California Department of Health Care Services (DHCS) funds 11 CRCs that provide support to family caregivers affected by chronic health conditions, including Alzheimer disease and related dementias and other degenerative diseases.

CRCs serve as a point of entry for services available to caregiving families, covering every county in California. While each center tailors its services to its geographic area, all CRCs have a core component of programs that provide uniform caregiver assessment and information, education, and support for caregivers. Individual CRCs also receive funds from county contracts, foundations, business partners, and donations to provide additional services. For over three decades, CRCs have supported caregivers in their regions, relying on staff intake interviews with caregivers and tailored referrals to relevant resources. With fluctuating funding, the programs became decentralized, and over the past decade, the 11 CRCs had been operating relatively independently. Funding to implement the statewide web-based platform provided an opportunity for new collaborations and connections.

Culture

CRCs provide services across income categories, and the original enabling legislation included middle-income families who are often overlooked and for whom few services are targeted. CRCs are united by shared values emphasizing choice, collaboration, innovation, quality, participation, respect, and diversity.

The implementation prompted culture change in three major ways: formalizing the system of CRCs across the state by uniting loosely affiliated sites, instituting standardized assessments, and changing to a new way of delivering services to clients virtually. The funding and subsequent process fostered a shared goal among CRC directors to serve the entire state and collaborate with one another to do so. Committing to a shared technology platform involved greater discussion and information exchange among the groups as well as the recognition that the whole is greater than the sum of the parts. A welcome culture change was the comradery and mutual aid that solidified over the initial year of implementation:

> *I feel like just within the CRC system the genuine enthusiasm for the project. Just to embrace it, even with all this other kind of craziness [COVID-19 pandemic] going on in the world.*

Implementation Climate

In-depth interviews revealed a collective positive attitude with a willingness to change, staff dedication, continued learning, and patience to follow through the stages and steps of the process. All interview participants expressed enthusiasm for the adoption of CareNav and a belief in the positive potential of this change in their practices:

> *I think it's been well received. We've been anticipating this for a long time. And have been very excited about it...We have some folks that were really excited and got right in there and started playing with it and*

working with it. Others I think are a little bit, you know, we're learning as we go.

Participants expressed hesitation about changing to a technology platform but recognized that caregivers' needs are changing in society and that the systems must evolve to meet the changing needs.

The idea is to be innovative and for caregivers, because, as things continue our younger population is very used to being more self-directed and being online. So, it has potential for growth and change over time. We'll continue to do that with a lot of streamlining of systems.

Networks and Communication

Communication was vital at all levels, from the implementation team to the sites, from the directors to their staff, and among sites. Communication involved developing a shared vision for the process and the outcomes, as well as coordinating the logistical aspects of the implementation. The process of preparing for technology deployment revealed workflow and processes that were not initially evident or taken for granted, and these had to be addressed as technology was applied to automate processes. Communication among the teams was vital to understanding work processes and establishing new ways of operating as needed.

Communication and support from FCA and among CRC sites played a major role in successful implementation, particularly as sites faced delays or barriers and were able to benefit from lessons learned elsewhere:

Something that's new is now we are communicating and reaching out to the other CRCs, which I believe hasn't happened in years. So it's really great that if we have a question or wondering, a different way of doing something in our program, I can just reach out and to anyone you know, in California and get their help and opinion.

Both the implementation process and the shared platform facilitated deeper collaboration among sites and the ability to elevate local issues with colleagues across the state, enhancing the strength of recommendations to address caregiver needs more comprehensively.

Readiness for Implementation

At the beginning of the implementation effort, 3 CRC sites, including FCA, were already using CareNav. The subsequent sites benefited from reliability and usability testing, as well as refinements and improvements that had been made during the initial, more limited deployment by FCA and QP.

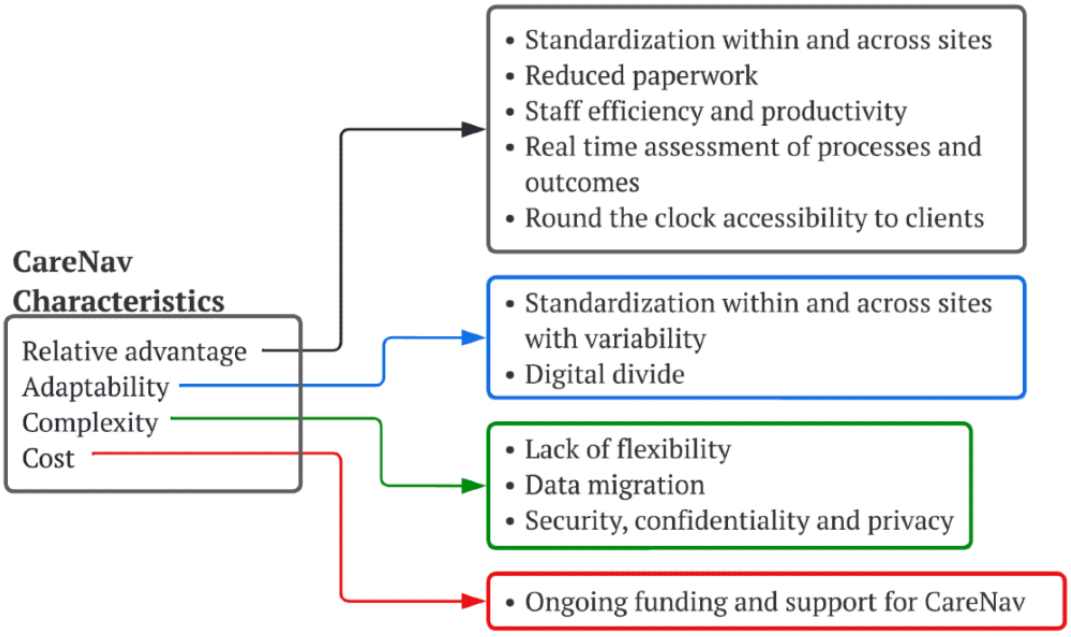
Leadership was an essential condition for success from the implementation team and at the CRC site level. Strong leadership, reassurance, and encouragement by managers, coupled with effective communication, established the overall vision, a shared understanding of the goal, the anticipated process and outcomes, and motivated staff across the sites to engage in implementation:

I think there's been leadership behind this and a sense of vision...FCA is [a] very unique organization and that [sic] they are really sitting in this point where they're both policy wonks as well as clinical experts. And I think that this allows FCA to bring vision and be forward looking of where we're going. And I think that the other CRCs are benefiting from that.

CareNav Characteristics

Focused interviews point to CareNav attributes that both facilitate and inhibit its implementation (Figure 4).

Figure 4. CareNav characteristics.



Relative Advantage

A key factor for the intervention is its relative advantage over previous data collection tools and software used by CRC sites to gather and aggregate site-level data on the caregivers they served and the programs they administered. Using CareNav reduced paperwork, enabled easier documentation, more efficient and secure charting, and a more environment-friendly approach. Additional advantages included increased CRC capacity and round-the-clock accessibility for clients. Several participants emphasized how the design of CareNav supports the process of care:

> *Now when a new person starts it's...here's the computer system that we use. This is the workflow, because I think it's really laid out nicely, the toolbar at the top of care journey. There's a workflow that you follow. It's just all right there in one place and as long as you have a secure connection, you can access it, so it's going to save a lot of time. I think people are going to end up really happy that they have that information at their fingertips, and they don't have to write it down on paper or search for it, laboriously, in their client file.*

Both clinicians and managers reported the advantages of CareNav for management functions. Standardization of the system across CRCs was perceived as allowing sites to collaborate for consultation, evaluation, and advocacy. Using a uniform tool supports administrative and management functions, including case data record-keeping as well as assessing the process and outcomes of the organization in real time:

> *For a manager being able to pull up reports without it being having to dump it into Excel and then sort through all the different things that you don't need. It looks like data collection is just going to be a lot simpler and easier...as a manager it's easier to measure staff productivity on a system like that because you can go in their notes.*

Adaptability

The most frequently identified benefit and concern about CareNav implementation is related to the adaptability of CareNav to the outer-setting variability among sites and populations served. Sites varied along multiple dimensions including population served, geographic characteristics, funding sources, relationships with host organizations, and size. The ability of CareNav to provide remote access, particularly in rural areas, was perceived as a benefit even with the variability in geographic distribution of clients in several CRCs and the impact of the COVID-19 pandemic.

Nonetheless, the digital divide (the issue of equitable access to technology and broadband connection) constitutes an ongoing barrier, particularly for rural, low-income, and non–English-speaking clients. Moreover, there are some caregivers in remote areas who value self-sufficiency and view government programs with suspicion, with reluctance to share information on the web regardless of access to broadband connection. Some communities in the catchment area do not have reliable broadband connections, and many clients may not be able to afford internet services or the associated technology. The platform is provided in English, limiting access to caregivers who do not speak English or have low literacy:

> *But language for me is probably my biggest concern, because we do have a big, Spanish speaking population and Vietnamese. It's a very significant group.*

Complexity

Sites hosted within a larger health care and information system experienced difficulties associated with the lack of flexibility and data migration, adding complexity and increasing the time and effort demands for staff. For instance, double data entry was necessary for some CRCs to ensure timely and accurate data for fiscal reporting and reimbursement across various funding agencies.

Confidentiality presents another level of complexity. While CareNav is a secure and private system, there is a learning curve for staff to understand the inherent privacy and security protection of the care team and clients. To assure a secure environment, written direct communication with clients changed with CareNav, requiring user authentication processes to maintain clients' privacy:

> *In the beginning, one of the issues that I was having, for example, when I was doing intake, if I sent emails, I was sending emails through CareNav. But people were not replying back and it turns out because it would send them a message into their actual email, but it would go to Spam....People just prefer for me to directly email them to their email.*

Cost

The threat of delay in funding and losing funding during the budget negotiations raised concerns about ongoing funding and support for CareNav. Several sites expressed frustration with the delays associated with state contracting and system readiness. Time-limited funding without a clear path to sustainability posed challenges for directors in their long-term planning.

Characteristics of the CRC Staff

Overview

The pretraining survey was an early opportunity to evaluate CRC staff knowledge about CareNav, self-efficacy, and readiness for change (Figure 5). Overall, the participants had very positive attitudes toward the implementation of CareNav (Table 1), with a total readiness score of 3.8 (SD 0.6) and average responses to almost all items in the positive range.

Figure 5. Characteristics of the Caregiver Resource Center (CRC) staff.

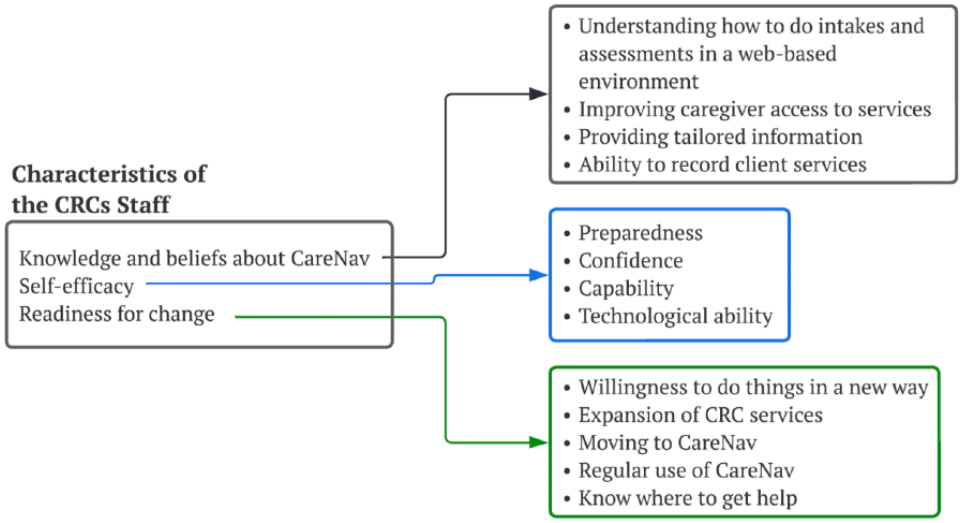


Table 1. Pretraining Readiness Survey (N=82).

Item	n (%)
Knowledge and beliefs about CareNav	
Seen or heard about CareNav[a]	73 (89)
Understand how to do an intake and assessment in CareNav[a]	30 (37)
CareNav is designed to improve caregiver access to services[a]	79 (98)
CareNav will improve the ability to record services[b]	64 (74)
CareNav provide tailored and accessible information for caregivers[b]	59 (68)
Self-efficacy	
Prepared to implement CareNav[c]	43 (49)
Confident to implement CareNav[c]	73 (84)
Capable to implement CareNav[c]	59 (68)
Readiness for change	
Positive with the expansion of CRC services[c]	71 (82)
Positive with moving to CareNav[c]	66 (76)
Willing about doing new things[c]	72 (83)
It will take time to ensure regularly use of CareNav[b] by everyone on the staff	73 (90)
Know where to obtain help[b]	47 (54)

[a]Staff members reported “yes.”

[b]Staff members reported “strongly agree” or “somewhat agree.”

[c]Staff members reported “very positive/willing/prepared/confident/capable” or “somewhat positive/willing/prepared/confident/capable.”

Knowledge and Beliefs About CareNav

Almost 90% (73/82) of staff had seen or heard about CareNav before the training and almost all (79/82, 98%) believed that CareNav is designed to improve caregiver access to services. Before training, one-third (30/82, 37%) of the staff members reported understanding how to complete the intake and assessment in CareNav. The staff members also believed that CareNav would enable better structure and process of care including improved ability to record services (mean 4.3, SD 0.9) and providing tailored and accessible information for caregivers (mean 4.0, SD 0.9).

Self-efficacy

The staff expressed a high level of self-efficacy to implement CareNav in terms of feeling prepared (mean 3.3, SD 1.2), confident (mean 4.5, SD 0.9), and capable (mean 3.9, SD 1.2).

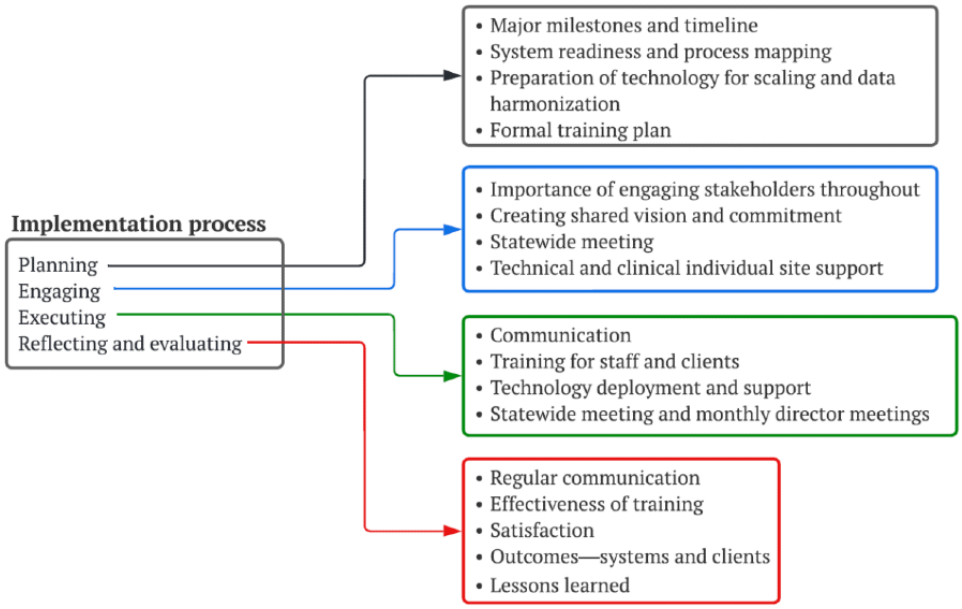
Readiness for Change

Regarding readiness for change, participants expressed a strong willingness to do new things (mean 4.5, SD 0.7) but believed that it would take time for everyone on the staff to become familiar with using CareNav (mean 1.5, SD 0.8).

Implementation Process

The elements of implementation included developing an overall project plan, engaging stakeholders, preparing technology for scaling, general training, and providing technical expertise and support at each CRC site (Figure 6).

Figure 6. Implementation process.



Planning

The implementation team consisting of staff from the FCA and QP developed a project management plan that addressed culture change, establishing site readiness, staff training, and technical implementation and support.

System Readiness and Process Mapping

Building on the demonstrated success and utility of CareNav in 3 sites, the implementation process involved planning for scaling the technology across 8 additional sites. As described in the inner and outer setting sections, the sites were diverse in several respects: the clients they served, the constellation of services and funding sources, and their relationship with a parent organization providing information technology. Early preparation involved assessing the requirements of each system for technology compatibility and interoperability, security, and compliance with applicable laws (eg, the Health Insurance Portability and Accountability Act). The technology team engaged with each site in mapping the data processes, identifying data sources, and reporting requirements.

Preparing the Technology for Scaling

The overall project plan involved determining major milestones along with the requisite resources and coordination to ensure progress. A major task of the technology team was to prepare the technology for scaling by harmonizing the data across the sites and mapping source data fields to the CareNav platform to assure standardization and data integrity. As the sites reviewed the CareNav software, they requested customization to fit their particular programs, funding requirements, and workflow, thus requiring site-specific revisions to the platform.

Engaging

The FCA team recognized the importance of engaging stakeholders throughout the process. They appreciated that scaling CareNav involved cultural change and made significant efforts to create a shared vision and commitment to engage in a new way. In a statewide kickoff meeting in January 2020, directors and staff of the CRCs came together to build relationships, develop a deeper understanding of CareNav and its deployment, and generate excitement regarding the effort. CRC sites shared best practices with one another and began to develop a stronger sense of collective resources and commitment to meet the needs of diverse caregivers across California.

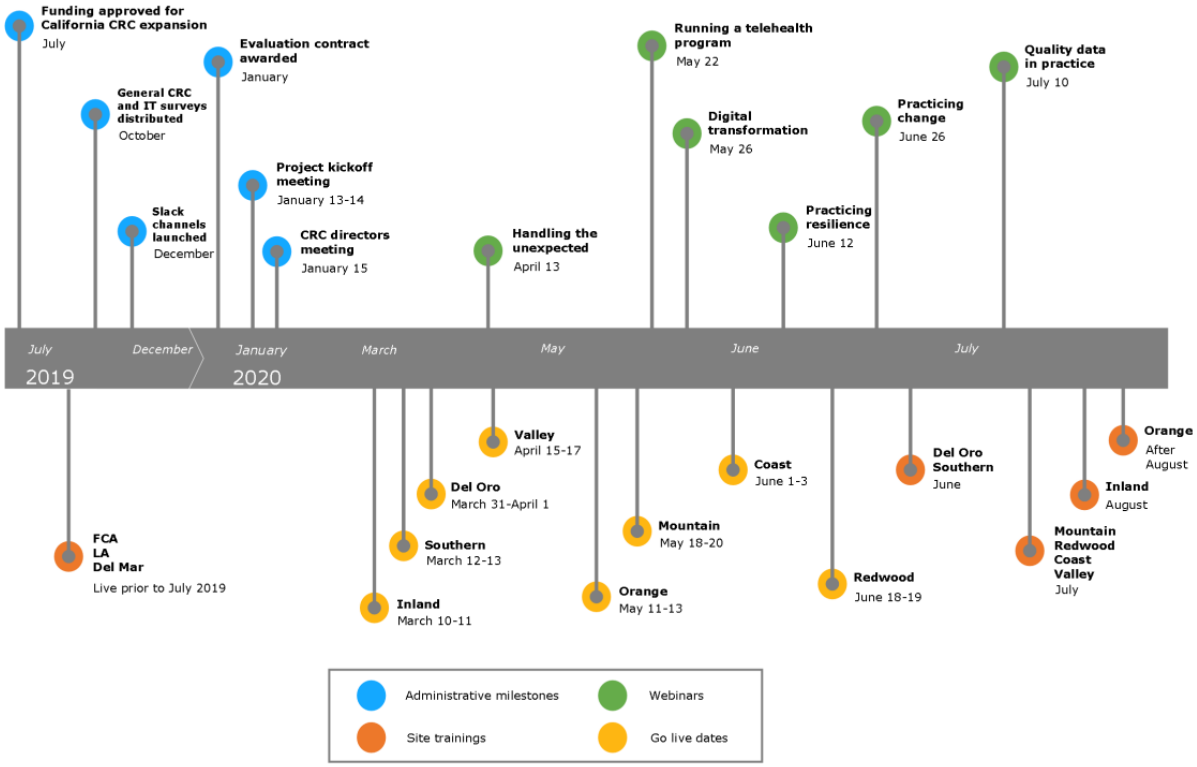
General training was accomplished through an initial in-person kickoff meeting followed by statewide webinars. The webinars were widely attended and offered topics to address overarching issues, such as managing change, using telehealth and technology-enabled assessment and supports, and using data for quality improvement.

Technical and clinical support was provided by individual site-level trainings to prepare the staff for technology deployment, addressing site-specific workflow and learning needs. Starting in March 2020, these 2-day sessions included the opportunity for hands-on practice using CareNav, reviewing new work processes, creating Caregiver Action Plans and service authorizations, generating reports, and using the library of resources available within CareNav. The implementation team provided extensive individual coaching and problem solving to the sites as the implementation proceeded.

Executing

Implementation, including all activities related to training, communication, and installation of CareNav at the 8 sites began in March 2020, and the last site was onboarded in late July 2020 (see timeline in Figure 7). The initial implementation schedule was affected by delays in finalizing state contracts, fund transfers, host agency site requirements for hiring staff, and information technology issues to enable the process to commence. In March 2020, owing to shelter-in-place and physical distancing measures to slow the spread of the COVID-19 pandemic, the sites responded by moving to remote operations with staff working from home.

Figure 7. Project timeline. CRC: Caregiver Resource Center; FCA: Family Caregiver Alliance; LA: Los Angeles.



Training and Communication

In addition to the statewide launch meeting followed by training at each individual site, regular statewide web-based training was organized to address best practices, quality improvements, and other topics of broad interest. Owing to state contracting delays, the overall implementation schedule lagged. In some cases, training occurred in March 2020, with delays occurring in technology deployment until July 2020. This necessitated refresher training and ongoing support. Overall, the training was viewed by interview participants as excellent, responsive, customized, and personal.

Technology Deployment and Support

The technology team developed and implemented a sophisticated process of CareNav deployment across sites. On the basis of the initial assessment of site-specific technology issues and requirements, they collaborated with staff members to map the data fields from existing data sources to CareNav. System-level requirements for security and interoperability with the existing information technology platform for sites hosted within health systems presented complex challenges with several levels of review and approval. For some sites, particularly those that are more rural, broadband connectivity was an issue for staff working remotely. Following the mapping process, the technology team worked with the sites to perform the data migration and verification. Throughout the process, the implementation team provided both clinical and technical support to assist the sites in making the transition effectively and assisted local staff to become resources for peers.

All sites praised the support received from the implementation team, including their responsiveness and ability to address challenges unique to the site and to their proactive approach to problem solving. It was clear that across all the organizations, both the site directors and staff were highly committed and dedicated to making the implementation a success. Ongoing communication enabled continuous access to the evolving information and knowledge.

The long-term continuous maintenance of the system has been emphasized as a key structural component required to maintain the operational activity of CRCs:

> *I think it would be great to have a source to go to when things break down...And we're not gonna go in and fix the system ourselves...So someone maintaining the system is gonna be real important for us to keep it going...that's part of the whole licensing process-we're licensed to use the software and support.*

Beyond staff training, several interview participants highlighted the importance of training clients on how to use the CareNav system. Currently, the staff members send information to clients about how to log on to the system but thought that this could

be enhanced to assist clients who are less experienced with the web environment or face linguistic barriers to a program provided in English. A few sites have identified the importance of marketing and outreach to assist caregivers in finding CRC services to facilitate access:

> *To me it's like okay so how do we sell this now and I think it's a lot it's gonna be getting them used to it. When my family care navigators schedule video zoom for family consultation...it's gonna be a lot of that education piece and getting them used to that. I think a change that we have to make internally to help the caregivers use it.*

Reflecting or Evaluating

Overview

Weekly data quality meetings were held with attendance by the evaluation researchers, CareNav application developer, and the FCA Client Services Director. These meetings focused on refining definitions and operationalizing evaluation metrics, defining data filters for the evaluation, identifying and addressing data entry discrepancies overall and by CRC site, and reconciling counts of activities in the evaluation data set with those in reports generated from CareNav. Each quarter, the data quality team met with each of the 11 CRC sites individually to share activity counts and service grant data for comments and any identified data entry issues particular to the site. The team met with the CRC directors approximately 4 times each year to share progress.

Another element of reflection and evaluation is the effectiveness of the training and outcomes for clients and staff. The posttraining survey and in-depth discussions with staff at all CRCs provided insight into the effectiveness of the training, participants' satisfaction with the process of implementation, as well as early outcomes of CareNav implementation both at the system and the client levels.

Comparing pretraining and posttraining readiness for implementation survey data shows significant increase in staff beliefs about their knowledge how to use CareNav, where to obtain help if needed and their preparedness. The proportion of staff who reported that they understood how to complete an intake and assessment in CareNav significantly increased from 27% before training to 97% after training (P<.001 based on the McNemar test). Both before and after training, the staff believed that the system would improve caregiver access to services. Given the high proportion endorsing this survey item in the pretraining and posttraining periods, this change was not statistically significant. Paired 2-tailed t test analysis showed significant improvement in feeling prepared to implement the CareNav score (pretraining mean 3.23, SD 1.24 vs posttraining mean 3.63, SD 0.91; P=.04) and in knowing where to obtain help (pretraining mean 3.52, SD 1.16 vs posttraining mean 4.32, SD 0.96; P<.001). A nearly significant decrease in the mean confidence score was found (pretraining mean 4.39, SD 0.99 vs posttraining mean 4.11, SD 0.93; P=.05). It is likely that the training fostered a greater awareness of the system and recognition that it would take time to learn new ways of working and using technology.

Almost all (n=57, 93%) staff reported willingness to adapt current services or provide services using CareNav. All participants will ask for support if they have questions about how to use CareNav, and 55 participants (90%) will offer support to coworkers if they have questions; 52 (85%) participants will ensure new coworkers are educated on how to use CareNav, and 50 participants (82%) will encourage coworkers to use CareNav.

Respondents found the training useful and helpful: 82% (n=50) rated the training as useful, 14% (n=9) were neutral, and 5% (n=3) did not find the training useful. In response to a question about whether the training met their needs, 76% (n=46) of the respondents agreed, 11% (n=7) were neutral, and 13% (n=8) did not feel that the training met their needs. Finally, almost 80% (n=48) felt that they had sufficient time to practice new skills.

Satisfaction

Focused interviews explored satisfaction with the system and implementation process. Overall, the interview participants expressed a high level of satisfaction with the implementation process, affirming the vision as important for the next phase of the work of the CRCs and embracing change as a positive force. Interview participants remarked on how easy it was to learn and use the new system and how readily they could see the benefits to their workflow and clients. Many expressed surprise at how well the process had gone despite the unforeseen challenges:

> *I mean I think for rolling out a new platform of this magnitude, it's gone amazingly well.*

> *It has been a pretty positive experience. So, I think it's been easy for me to navigate and I think people who are used to doing everything online appreciate that from the system.*

Interview participants from the 2 sites that were the last to implement the program were less satisfied and more skeptical about the long-term benefits. Unlike other sites, these sites encountered system barriers including integration with the information system at the broader host organization and difficulty harmonizing their existing records with the new platform. Despite these frustrations and delays, the sites remained open-minded and were willing to move forward:

> *I think it just further kind of reaffirms that our staff is flexible and we'll just roll with it most of the time and yeah it's frustrating and that sort of thing that we're going to make the best of it, regardless of the situation and that's just kind of our attitude about most things in general.*

Effectiveness of Training

The interview participants recognized that many of the features of the platform were even more relevant in the context of the pandemic and that physical distancing guidelines accelerated adoption by agencies and clients at an unexpected rate. Importantly, throughout the pandemic, CRCs continued to innovate and serve their clients who needed them more than ever:

> *I think what we've all learned that we've been very surprised about is we have discovered, so many new opportunities because of the pandemic. And so, our repertoire of services or our availability is expanded. My staff continue to express great satisfaction with their relationships with their clients and, and they are now having satisfying phone conversations where they're having satisfying online support groups, we're continuing to have our educational workshops and online and people are satisfied with that. So yes there are disappointments and frustrations, but at the same time we've discovered so many opportunities. And we've become empowered, and we are saying to the community, "We're here. We're open we didn't go anywhere, call us, we're here for you" and so that was a nice surprise, and my staff, myself, the entire organization, we just feel more confident now of knowing how to work remotely, knowing how to use zoom, knowing how to use our new databases, we just feel like, we got it together. I always say to my team they're rock stars they just impress me every day.*

The focused interviews also explored early outcomes of the implementation of the web-based platform, both at the system and client levels. Finally, the participants shared the lessons learned during the process.

System Outcomes

The most commonly reported benefits were improved access and convenience. This was particularly salient during the pandemic with staff working from home. Interview participants cited the benefits of having a single paperless location for client data, accessible by any staff member from any location. Family consultants also valued the ability to track services more completely:

> *It's allowed us to work remotely, it has allowed at the time of COVID to work safely as a group and still support our caregivers. We've been able to continue to offer services. We've continued to be able to document and work together.*

The directors valued the potential of the system to generate data to evaluate services, improve quality, guide program decisions, and provide evidence for advocacy. At the site level, directors anticipated using the data to guide strategic directions and identify service gaps. Statewide, they valued the potential to aggregate data to inform planning and policy. For example, data on caregivers can inform the implementation of the California Master Plan on Aging:

> *The other big benefit eventually will be that we will have statewide information data that is consistent across all 11 Caregiver Resource Centers. And to me that's huge because it helps with government planning and policymaking as it relates to caregiving.*

A positive consequence of this initiative has been bringing the sites together, with a shared vision to serve caregivers across the entire state, learning from one another and sharing resources. Several directors shared a vision to collaborate in new ways, assuring that caregivers across the state had access to culturally and linguistically congruent resources. For example:

> *Could we survey our sites to see who has family consultants in different languages? In our County we have a pretty sizable Korean population. We currently don't have a Korean family consultant, but if another site did, could they serve them? And could we serve if they had a Vietnamese client? In other words, how do we leverage our statewide network around language barriers? I think there's a potential opportunity to use the statewide network differently with this common platform.*

Client Outcomes

The greatest client benefit reported was having a platform that clients could access at their convenience, around the clock and from any location. In addition, the consultants found it beneficial that the platform enabled them to tailor resources for specific clients based on their assessment, which allowed for immediate delivery from the web-based library. Clients could also access information about authorized services, such as respite or counseling on the web. Several consultants voiced their advantage to clients with disabilities who find web-based navigation more accessible than other alternatives:

> *I have clients with disabilities, for example, I had a few who were completely deaf and—we communicated through (CareNav) and through email. So, that was a wonderful tool for her because she said in the past it was really challenging to get services.*

The new platform provided an avenue for clients to connect during the COVID-19 pandemic and address social isolation in a way that had not been previously available:

> *One of our staff members said, "my online Spanish support group are really having a hard time with the isolation and they so appreciate being together at least once a month, but they want to meet two times. Can I do my support group two times a month?" I said of course!*

Lessons Learned

The most commonly stated lesson was *patience*. All the interview participants reported resilience and flexibility as they approached and engaged in the implementation process. Many affirmed the positive nature of the change and a commitment to ongoing learning associated with this platform enhancement. They expressed heightened appreciation for the importance of communication and collaboration within sites, across sites, and in the community. For those on the clinical and technical support teams, they recognized in even greater detail the variability and diversity across the sites:

> *Having to realize that people really think differently, and they learn at different speeds, and they communicate differently, both individuals and as organizations, and the variability in even their data or how they store their data and how they manage their programs. Yeah, I think there was just a little bit of variability everywhere.*

Discussion

Principal Findings

Our findings suggest 4 factors as essential underpinnings of successful implementation, with implications for future replication [29]. First, key informants emphasized the importance of effective leadership as a key inner-setting component. The participants cited creativity, perseverance, unwavering vision, clear communication, relevant expertise, and deep knowledge of all aspects of the intervention and the inner and outer settings. Effective leadership facilitates a positive implementation climate, which in turn enhances implementation effectiveness [30]. Beyond the inner setting, leadership was described as continually interacting and potentially affecting other CFIR domains [31]. The CareNav implementation team has long tenure in the system and a history of successful implementation of other prior initiatives across the CRC system that may have increased staff self-efficacy. The expertise of program managers, information technology staff, and site directors and staff, as well as leaders' deep knowledge of CareNav and the CRCs outside networks contributed to continuous interaction with stakeholders throughout the technology implementation process [31]. The leaders' deep understanding of unique CRC site-level structures and processes was described as particularly important for enhancing both customized and overall implementation efforts. These findings echo recent technology implementation studies in multisite settings, highlighting the lack of senior leadership endorsement as impeding successful implementation [32].

Second, the interview participants stressed the key role of training and ongoing support to increase their self-efficacy using the technology platform. Ongoing training and support were suggested as effective strategies to attract and involve key stakeholders in implementing or using the innovation [33,34]. Our results in the implementation process domain provide insight into the participants' thoughts on the ongoing training performed. The implementation leadership team gathered information about the system user experience from the training and applied the learning for system refinement and quality improvement. As reflected in both the surveys and interviews, participants appreciated the pace and cadence of the training as well as the opportunity to provide feedback. This feedback loop, in turn, improved platform usability, further inspiring staff self-efficacy. Of note, this 2-way communication for continuous improvement is ongoing with quarterly meetings of the QP development team, program evaluators, and CareNav end users to review program productivity data and share issues, concerns, and feedback—providing an open forum for problem solving in the future.

Third, CareNav represented culture change in two major ways: by bringing together decentralized sites and by introducing novel technology that provides customization, round-the-clock availability, and enhanced workflow for staff. In addition, CareNav reflects culture changes societally. In both surveys and interviews, participants expressed a high degree of willingness to change and an openness to new approaches to delivering and documenting interactions with caregivers. This willingness to adopt technology was a key factor in the rapid and successful deployment of the platform.

Finally, the interview participants raised a critical issue of ongoing funding to support the CareNav initiative going forward. Over the last few decades, state funding for the CRC system has waxed and waned, along with the political will to support family caregivers. In the leanest years, CRCs functioned largely as independent entities, drawing on local funding and scaling back services accordingly. In the current phase of funding, state-level investment for CareNav was significant, covering the technology roll out across 11 sites and supporting the QP CareNav development team to provide ongoing technical support, system refinement, reporting capability, and building system elements that expand the functionality of the user interface of the portal through which caregivers can complete intake and assessment forms and access information around the clock. Without ongoing funds to support this vital work, the impact of this initial investment will be vastly diminished, and none of the stakeholders (caregivers, CRCs, and funders) will be able to capitalize on its full potential. Indeed, the long-term benefits of a uniform assessment, data collection, and reporting system are yet to fully materialize, with great potential to disseminate interventions to support family caregivers in the future.

Context of the COVID-19 Pandemic

This study echoes the recent implementation literature underscoring the COVID-19 pandemic as an outer-setting construct inconclusively affecting implementation [35-37]. In fact, the pandemic may have facilitated and accelerated the technology implementation process with significant effects on CareNav characteristics, the inner setting, characteristics of the CRC staff, and the outer setting. CareNav proved to be an agile system during the unexpected national and global circumstances of the COVID-19 pandemic and showed promise of helpfulness in sudden onset or unforeseen conditions, especially in California, given the likelihood of natural disasters such as earthquakes or fires and their impacts on caregivers and the caregiving role. Indeed, the vision for a shared technology platform proved to be prescient as the COVID-19 pandemic gained momentum and the CRCs were able to respond quickly to provide uninterrupted services to their clients by ramping up outreach and adapting supports for caregivers. As examples related to the inner setting, the circumstances of the pandemic may have improved the organizational will to implement by spotlighting the perceived benefits of the technology platform, given its ability to support web-based work; improved the implementation climate through increased staff enthusiasm for the platform; increased staff patience to follow through the stages and steps of the implementation process; reduced hesitation about changing to the technology platform; and increased the need for cross-site communication to best serve the needs of clients. Related to the outer setting, CRC staff members were acutely aware of the toll the pandemic took on their clients, as their needs increased and access to many services and supports—including respite care, adult day care, and institutional placement—became more limited.

At the same time, CRC staff members themselves experienced added strain, juggling their own competing demands outside work, covering for one another owing to illness, coping with staff turnover, and pivoting to remote work from home to provide web-based family consultations to their clients. These factors may have either increased or decreased the willingness of the staff to use the CareNav platform to do their work in new ways, their readiness for change, and their perceived need to expand services. Moreover, the pandemic might have slowed the ability to integrate CareNav with other systems, specifically for sites hosted within a larger health care system who were themselves focusing on care delivery priorities. Similarly, the implementation of other web-based technologies encountered significant barriers, explained by time constraints and competing priorities associated with the COVID-19 pandemic [38]. Effective leadership engagement, as demonstrated in our study, might have played a mediating role in buffering the COVID-19 pandemic slowing effect by maintaining commitment and resources to the implementation and ultimately reducing behavioral resistance to change [35,36,39]. Taken together, these factors must be considered in the interpretation of the findings of this evaluation, particularly in light of the perceived success of the implementation from the staff perspective.

Comparison With Prior Work

This implementation evaluation also offers the knowledge needed to address gaps in prior research. A systematic review of internet-based supportive interventions for caregivers of patients with dementia recommended individual tailoring of supports for the success of digital interventions for caregivers [23]. Underlying this recommendation is the heterogeneity of web-based support systems regarding intervention type, dosage (amount of time spent on the web), and duration, which affect intervention effectiveness, feasibility, and quality [23,40]. In addition, caregivers use web-based supports targeting caregiving for distinct illnesses such as cancer [41], dementia [10], posttraumatic stress disorder [20], and psychosis [21,22] suggesting illness-specific web-based support needs for caregivers. As described, CareNav enables either self-administered or staff-administered caregiver and care recipient assessment, appraisal of current resources, delineation of priority supports, and development of a care plan to contract appropriate services such as respite, educational sessions, support groups, and vouchers for legal aid or counseling across all health conditions, providing a more comprehensive resource for caregivers. These services can be individually tailored to caregivers and illness-specific caregiving needs. CareNav includes a tailored library of resources for caregivers based on the assessment, with the ability to link to resources such as the World Health Organization's iSupport site [10].

Studies in the past have shown that web-based support can increase confidence, self-efficacy, and self-esteem and reduce depression and strain among caregivers [14,23,42]. This study showed that the capacity to enhance access, convenience, and customized support for clients held a universal appeal for CRC staff. CRC sites highly valued the potential of the data collected in CareNav to provide real-time feedback on who they were serving and the effectiveness of their interventions. Future analyses should explore changes in client confidence and mental health outcomes. As users gain expertise and competence, they will likely see greater benefits to full use of the CareNav platform features, enhancing client engagement and improving decision-making for care consultants, particularly for clients at the highest risk for negative outcomes.

Limitations

This study was an evaluation of an implementation process in established CRCs across the state. There were trade-offs in data collection and reporting to protect the anonymity of participants who could otherwise be readily identified. Therefore, this study does not provide demographic characteristics of participants nor is it able to link findings to the specific roles of staff. However, the decision to permit anonymity facilitated broader participation and a more complete perspective on implementation, as evidenced by the high participation rate of staff in both the surveys and interviews. We applied the CFIR during the data analysis stages of the implementation and did not use the framework to its full potential in guiding implementation design. This study focused on the implementation process and its immediate impact on staff and clients, and did not examine changes in caregiver health and well-being.

Conclusions

In this study, we applied the CFIR [29] to examine the implementation of a web-based platform to engage and support family caregivers across 11 regional sites in California, the most populous state in the United States. The findings identified factors that contributed to success and were relevant for future scaling and dissemination of this innovative platform designed to support global assessment, intervention, and service delivery to address the unmet needs of family caregivers. Importantly, the platform is accessible to both staff and clients regardless of geographic region or care recipient medical condition, and supports timely and local responses to caregiver priorities. Using the CFIR facilitated a more systematic understanding of the multilevel experience of technology implementation across multiple sites. Future implementation projects would benefit from using the CFIR to organize implementation strategies and identify potential blind spots in planning.

As noted, the value of long-term care provided by family caregivers eclipses that provided by government funding [2]. Caregivers have become indispensable in health care delivery. Although their contributions have remained largely invisible, CareNav elevates caregivers' contributions as part of the fabric of health care. The global assessment of the caregiver and the ability to respond to the priority of the caregiver in real time with information or referral responds to myriad health-related urgencies and imbeds caregivers as elemental in the fluid and ongoing process of care. CareNav additionally supports caregivers in the community. Studies show that the caregiver community is helpful over and above the benefits related to caregivers' coping [43]. Information gathering, reminiscing, legacy building, and giving back to the community of caregivers maintain the spirit of commitment [42,43]. Although in our study, the mechanism of culture change is primarily associated with modernizing business practices using technology, the CRC vision and long-range outcomes align with broader societal

implications of acknowledging and supporting caregivers as a crucial link in providing supports across communities.

Among the 5 bold goals of California's Master Plan on Aging for 2030 is *caregiving that works* [44]. The California CRCs took on the impressive goal to harmonize their information technology platform and implement the new platform across the state in all 11 sites in 6 months. Other state plans, such as those of New York, have similar priorities for caregivers and CRCs [45]. Future studies could compare California state data with both other states and national populations. This could produce greater knowledge of the similarities and differences in the support needs for the nation's caregivers. With the opportunity for longitudinal analysis, future studies could examine the predictive factors of intake for negative caregiver outcomes, facilitating earlier interventions. Further analysis of use and outcomes could elucidate which services are most helpful for subpopulations of caregivers.

The next phase of implementation should involve developing a deeper understanding of the rich information available through CareNav, determining both site-specific and statewide reports that would be most helpful in evaluating adoption and dissemination of the programs into the community, effectiveness and gaps in service, quality of delivery, and the impact of the services and supports on caregiver outcomes. The data hold the power to drive individual and system changes. At the individual level, the assessment can be the basis for determining risk and matching services to caregiver needs. At the system level, the data can drive strategy for priority program development, funding, and advocacy. Future studies should evaluate the impact of the CRC system on caregiver health and well-being, as well as develop a deeper characterization of the trajectory of caregiving and how interventions can improve outcomes for caregivers and the persons in their care.

A determined and committed group of leaders and staff dedicated to improving the lives of caregivers began a journey together. They are well on their way to actualizing a vision for the future of caregivers in California.

Acknowledgments

This work is a collective effort with many dedicated individuals sharing a vision and contributing their time, energy, and effort to serve the caregivers of California. The evaluation team worked closely with the Family Caregiver Alliance team (represented by Christina Irving and KK) and the Quality Process team (represented by Brad Silen) throughout the evaluation design, data collection, and data verification process. The authors appreciate the leadership and commitment of the 11 Caregiver Resource Center directors and the dedicated staff who participated in the implementation while delivering high quality services.

The authors received funding from the California Department of Health Care Services and the Gordon and Betty Moore Foundation (Family Caregiving Institute and Heather M Young postdoctoral fellows TRK and OT).

Conflicts of Interest

None declared.

References

1. Caregiving in the U.S. 2020. The National Alliance for Caregiving. Washington, DC, USA: National Alliance for Caregiving and AARP; 2020. URL: https://www.caregiving.org/caregiving-in-the-us-2020/ [accessed 2021-08-14]
2. Reinhard S, Feinberg LF, Houser A, Choula R, Evans M. Valuing the Invaluable 2019 Update: Charting a Path Forward. American Association of Retired Persons. 2019. URL: https://www.aarp.org/ppi/info-2015/valuing-the-invaluable-2015-update.html [accessed 2021-08-15]
3. Reinhard SC, Young HM, Levine C, Kelly K, Choula R, Accius J. Home Alone Revisited: Family Caregivers Providing Complex Care. American Association of Retired Persons. 2019 Apr. URL: https://www.aarp.org/ppi/info-2018/home-alone-family-caregivers-providing-complex-chronic-care.html [accessed 2020-02-10]
4. Keeton VF, Trask J, Whitney R, Bell JF. Overburdened and underprepared: medical/nursing task performance among informal caregivers in the United States. J Gerontol Nurs 2020 Sep 01;46(9):25-35. [doi: 10.3928/00989134-20200811-05] [Medline: 32845344]
5. Bell JF, Whitney RL, Young HM. Family caregiving in serious illness in the United States: recommendations to support an invisible workforce. J Am Geriatr Soc 2019 May;67(S2):S451-S456. [doi: 10.1111/jgs.15820] [Medline: 31074854]
6. Kaltenbaugh DJ, Klem ML, Hu L, Turi E, Haines AJ, Hagerty Lingler J. Using Web-based interventions to support caregivers of patients with cancer: a systematic review. Oncol Nurs Forum 2015 Mar;42(2):156-164. [doi: 10.1188/15.ONF.156-164] [Medline: 25806882]
7. Kales HC, Gitlin LN, Stanislawski B, Marx K, Turnwald M, Watkins DC, et al. WeCareAdvisor™: the development of a caregiver-focused, Web-based program to assess and manage behavioral and psychological symptoms of dementia. Alzheimer Dis Assoc Disord 2017;31(3):263-270 [FREE Full text] [doi: 10.1097/WAD.0000000000000177] [Medline: 27849639]
8. Zhao Y, Feng H, Hu M, Hu H, Li H, Ning H, et al. Web-based interventions to improve mental health in home caregivers of people with dementia: meta-analysis. J Med Internet Res 2019 May 06;21(5):e13415 [FREE Full text] [doi: 10.2196/13415] [Medline: 31066680]

9. Leng M, Zhao Y, Xiao H, Li C, Wang Z. Internet-based supportive interventions for family caregivers of people with dementia: systematic review and meta-analysis. J Med Internet Res 2020 Sep 09;22(9):e19468 [FREE Full text] [doi: 10.2196/19468] [Medline: 32902388]
10. iSupport for dementia: training and support manual for carers of people with dementia. World Health Organization. 2019 May 9. URL: https://www.who.int/publications/i/item/9789241515863 [accessed 2022-05-15]
11. Lorca-Cabrera J, Grau C, Martí-Arques R, Raigal-Aran L, Falcó-Pegueroles A, Albacar-Riobóo N. Effectiveness of health Web-based and mobile app-based interventions designed to improve informal caregiver's well-being and quality of life: a systematic review. Int J Med Inform 2020 Feb;134:104003. [doi: 10.1016/j.ijmedinf.2019.104003] [Medline: 31790857]
12. Ploeg J, Markle-Reid M, Valaitis R, McAiney C, Duggleby W, Bartholomew A, et al. Web-based interventions to improve mental health, general caregiving outcomes, and general health for informal caregivers of adults with chronic conditions living in the community: rapid evidence review. J Med Internet Res 2017 Jul 28;19(7):e263 [FREE Full text] [doi: 10.2196/jmir.7564] [Medline: 28754652]
13. Wasilewski MB, Stinson JN, Cameron JI. Web-based health interventions for family caregivers of elderly individuals: a scoping review. Int J Med Inform 2017 Jul;103:109-138. [doi: 10.1016/j.ijmedinf.2017.04.009] [Medline: 28550996]
14. Ploeg J, Ali MU, Markle-Reid M, Valaitis R, Bartholomew A, Fitzpatrick-Lewis D, et al. Caregiver-focused, Web-based interventions: systematic review and meta-analysis (part 2). J Med Internet Res 2018 Oct 26;20(10):e11247 [FREE Full text] [doi: 10.2196/11247] [Medline: 30368439]
15. Etxeberria I, Salaberria K, Gorostiaga A. Online support for family caregivers of people with dementia: a systematic review and meta-analysis of RCTs and quasi-experimental studies. Aging Ment Health 2021 Jul;25(7):1165-1180. [doi: 10.1080/13607863.2020.1758900] [Medline: 32363901]
16. Nichols LO, Martindale-Adams J, Burns R, Zuber J, Graney MJ. REACH VA: moving from translation to system implementation. Gerontologist 2016 Feb;56(1):135-144. [doi: 10.1093/geront/gnu112] [Medline: 25398828]
17. Pleasant M, Molinari V, Dobbs D, Meng H, Hyer K. Effectiveness of online dementia caregivers training programs: a systematic review. Geriatr Nurs 2020;41(6):921-935. [doi: 10.1016/j.gerinurse.2020.07.004] [Medline: 32703628]
18. Egan KJ, Pinto-Bruno Á, Bighelli I, Berg-Weger M, van Straten A, Albanese E, et al. Online training and support programs designed to improve mental health and reduce burden among caregivers of people with dementia: a systematic review. J Am Med Dir Assoc 2018 Mar;19(3):200-6.e1. [doi: 10.1016/j.jamda.2017.10.023] [Medline: 29306605]
19. Male DA, Fergus KD, Stephen JE. The continuous confrontation of caregiving as described in real-time online group chat. J Palliat Care 2015;31(1):36-43. [doi: 10.1177/082585971503100106] [Medline: 26399089]
20. Ferrell EL, Russin SE, Hardy RM. Informal caregiving experiences in posttraumatic stress disorder: a content analysis of an online community. J Community Psychol 2019 May;47(4):757-771. [doi: 10.1002/jcop.22151] [Medline: 30592051]
21. Glynn SM, Randolph ET, Garrick T, Lui A. A proof of concept trial of an online psychoeducational program for relatives of both veterans and civilians living with schizophrenia. Psychiatr Rehabil J 2010;33(4):278-287. [doi: 10.2975/33.4.2010.278.287] [Medline: 20374986]
22. Sin J, Henderson C, Norman I. Usability of online psychoeducation for siblings of people with psychosis. Int J Technol Assess Health Care 2014 Oct;30(4):374-380. [doi: 10.1017/S0266462314000488] [Medline: 25394550]
23. Boots LM, de Vugt ME, van Knippenberg RJ, Kempen GI, Verhey FR. A systematic review of Internet-based supportive interventions for caregivers of patients with dementia. Int J Geriatr Psychiatry 2014 Apr;29(4):331-344. [doi: 10.1002/gps.4016] [Medline: 23963684]
24. Hirschman KB, Bowles KH, Garcia-Gonzalez L, Shepard B, Walser TJ, Thomas GL, et al. Lessons learned from the implementation of a video health coaching technology intervention to improve self-care of family caregivers of adults with heart failure. Res Nurs Health 2021 Feb;44(1):250-259 [FREE Full text] [doi: 10.1002/nur.22100] [Medline: 33341950]
25. Lindeman DA, Kim KK, Gladstone C, Apesoa-Varano EC. Technology and caregiving: emerging interventions and directions for research. Gerontologist 2020 Feb 14;60(Suppl 1):S41-S49 [FREE Full text] [doi: 10.1093/geront/gnz178] [Medline: 32057082]
26. 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]
27. Bradshaw C, Atkinson S, Doody O. Employing a qualitative description approach in health care research. Glob Qual Nurs Res 2017 Nov 24;4:2333393617742282 [FREE Full text] [doi: 10.1177/2333393617742282] [Medline: 29204457]
28. 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]
29. Keith RE, Crosson JC, O'Malley AS, Cromp D, Taylor EF. Using the Consolidated Framework for Implementation Research (CFIR) to produce actionable findings: a rapid-cycle evaluation approach to improving implementation. Implement Sci 2017 Feb 10;12(1):15 [FREE Full text] [doi: 10.1186/s13012-017-0550-7] [Medline: 28187747]
30. Helfrich CD, Weiner BJ, McKinney MM, Minasian L. Determinants of implementation effectiveness: adapting a framework for complex innovations. Med Care Res Rev 2007 Jun;64(3):279-303. [doi: 10.1177/1077558707299887] [Medline: 17507459]

31. Melder A, Robinson T, McLoughlin I, Iedema R, Teede H. Integrating the complexity of healthcare improvement with implementation science: a longitudinal qualitative case study. BMC Health Serv Res 2022 Feb 19;22(1):234 [FREE Full text] [doi: 10.1186/s12913-022-07505-5] [Medline: 35183164]
32. Fujioka JK, Bickford J, Gritke J, Stamenova V, Jamieson T, Bhatia RS, et al. Implementation strategies to improve engagement with a multi-institutional patient portal: multimethod study. J Med Internet Res 2021 Oct 28;23(10):e28924 [FREE Full text] [doi: 10.2196/28924] [Medline: 34709195]
33. Perry CK, Damschroder LJ, Hemler JR, Woodson TT, Ono SS, Cohen DJ. Specifying and comparing implementation strategies across seven large implementation interventions: a practical application of theory. Implement Sci 2019 Mar 21;14(1):32 [FREE Full text] [doi: 10.1186/s13012-019-0876-4] [Medline: 30898133]
34. Longacre ML, Keleher C, Chwistek M, Odelberg M, Siemon M, Collins M, et al. Developing an integrated caregiver patient-portal system. Healthcare (Basel) 2021 Feb 10;9(2):193 [FREE Full text] [doi: 10.3390/healthcare9020193] [Medline: 33578838]
35. Piat M, Wainwright M, Cherkas D, Leblanc S, Sofouli E, Rivest MP, et al. Identifying and understanding the contextual factors that shaped mid-implementation outcomes during the COVID-19 pandemic in organizations implementing mental health recovery innovations into services. Implement Sci Commun 2021 Sep 15;2(1):101 [FREE Full text] [doi: 10.1186/s43058-021-00206-w] [Medline: 34526136]
36. Becker SJ, Garner BR, Hartzler BJ. Is necessity also the mother of implementation? COVID-19 and the implementation of evidence-based treatments for opioid use disorders. J Subst Abuse Treat 2021 Mar;122:108210 [FREE Full text] [doi: 10.1016/j.jsat.2020.108210] [Medline: 33509413]
37. Husain A, Cohen E, Dubrowski R, Jamieson T, Kurahashi AM, Lokuge B, et al. A clinical communication tool (loop) for team-based care in pediatric and adult care settings: hybrid mixed methods implementation study. J Med Internet Res 2021 Mar 03;23(3):e25505 [FREE Full text] [doi: 10.2196/25505] [Medline: 33656445]
38. Levinson AJ, Bousfield J, Douglas W, Ayers S, Sztramko R. A novel educational prescription Web-based application to support education for caregivers of people living with dementia: development and usability study with clinicians. JMIR Hum Factors 2020 Dec 04;7(4):e23904 [FREE Full text] [doi: 10.2196/23904] [Medline: 33275103]
39. Lorenzi NM, Riley RT. Managing change: an overview. J Am Med Inform Assoc 2000;7(2):116-124 [FREE Full text] [doi: 10.1136/jamia.2000.0070116] [Medline: 10730594]
40. Powell J, Chiu T, Eysenbach G. A systematic review of networked technologies supporting carers of people with dementia. J Telemed Telecare 2008;14(3):154-156. [doi: 10.1258/jtt.2008.003018] [Medline: 18430288]
41. Male DA, Fergus KD, Stephen JE. Professional online support group facilitators: guarantors of maximal group utility. J Group Psychother 2017;67(3):314-336 [FREE Full text] [doi: 10.1080/00207284.2016.1240587]
42. Lewis ML, Hobday JV, Hepburn KW. Internet-based program for dementia caregivers. Am J Alzheimers Dis Other Demen 2010 Dec;25(8):674-679 [FREE Full text] [doi: 10.1177/1533317510385812] [Medline: 21131674]
43. Anderson JG, Hundt E, Dean M, Keim-Malpass J, Lopez RP. "The church of online support": examining the use of blogs among family caregivers of persons with dementia. J Fam Nurs 2017 Feb;23(1):34-54. [doi: 10.1177/1074840716681289] [Medline: 27920340]
44. California Master Plan on Aging. California Department for Aging. 2021. URL: https://mpa.aging.ca.gov
45. 2019-2023 New York State Plan on Aging. New York Office for the Aging. 2019. URL: https://aging.ny.gov/system/files/documents/2019/10/state_plan_2019-2023_070119_final2com.pdf [accessed 2022-05-08]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
CRC: Caregiver Resource Center
FCA: Family Caregiver Alliance
QP: Quality Process

Edited by A Mavragani; submitted 14.04.22; peer-reviewed by A Garcia, O Ogundaini; comments to author 08.05.22; revised version received 23.05.22; accepted 24.05.22; published 13.07.22.

Please cite as:
Young HM, Bell JF, Tonkikh O, Kilaberia TR, Whitney RL, Mongoven JM, Link BM, Kelly K
Implementation of a Statewide Web-Based Caregiver Resource Information System (CareNav): Mixed Methods Study
JMIR Form Res 2022;6(7):e38735
URL: https://formative.jmir.org/2022/7/e38735
doi:10.2196/38735
PMID:35830234

Original Paper

A Privacy-Preserving Audit and Feedback System for the Antibiotic Prescribing of General Practitioners: Survey Study

Kassaye Yitbarek Yigzaw[1], MSc, PhD; Taridzo Chomutare[1], MSc, PhD; Rolf Wynn[2,3], MD, PhD; Gro Karine Rosvold Berntsen[1,4], MD, PhD; Johan Gustav Bellika[1,2], MSc, PhD

[1]Norwegian Centre for E-health Research, University Hospital of North Norway, Tromsø, Norway
[2]Institute of Clinical Medicine, UiT The Arctic University of Norway, Tromsø, Norway
[3]Divison of Substance Use and Mental Health, University Hospital of North Norway, Tromsø, Norway
[4]Department of Community Medicine, UiT The Arctic University of Norway, Tromsø, Norway

Corresponding Author:
Kassaye Yitbarek Yigzaw, MSc, PhD
Norwegian Centre for E-health Research
University Hospital of North Norway
Sykehusvneien 23
Tromsø, 9019
Norway
Phone: 47 92438720
Email: kassaye.yitbarek.yigzaw@ehealthresearch.no

Abstract

Background: Antibiotic resistance is a worldwide public health problem that is accelerated by the misuse and overuse of antibiotics. Studies have shown that audits and feedback enable clinicians to compare their personal clinical performance with that of their peers and are effective in reducing the inappropriate prescribing of antibiotics. However, privacy concerns make audits and feedback hard to implement in clinical settings. To solve this problem, we developed a privacy-preserving audit and feedback (A&F) system.

Objective: This study aims to evaluate a privacy-preserving A&F system in clinical settings.

Methods: A privacy-preserving A&F system was deployed at three primary care practices in Norway to generate feedback for 20 general practitioners (GPs) on their prescribing of antibiotics for selected respiratory tract infections. The GPs were asked to participate in a survey shortly after using the system.

Results: A total of 14 GPs responded to the questionnaire, representing a 70% (14/20) response rate. The participants were generally satisfied with the usefulness of the feedback and the comparisons with peers, as well as the protection of privacy. The majority of the GPs (9/14, 64%) valued the protection of their own privacy as well as that of their patients.

Conclusions: The system overcomes important privacy and scaling challenges that are commonly associated with the secondary use of electronic health record data and has the potential to improve antibiotic prescribing behavior; however, further study is required to assess its actual effect.

(JMIR Form Res 2022;6(7):e31650) doi:10.2196/31650

KEYWORDS

learning health care system; feedback; antimicrobial stewardship; quality improvement; privacy; electronic health record; antibiotics; prescription; patient privacy; clinical setting

Introduction

Antibiotic stewardship is an issue that has recently come under scrutiny in Norway [1] and internationally [2], especially in general practices, which is where most antibiotic prescribing occurs [3]. The inappropriate use of antibiotics is thought to be an important factor in the development of antimicrobial resistance. Although a number of guidelines define when antibiotic use is warranted, defining *appropriate use* is difficult, unless there is clear agreement on the etiology of an infection [4]. There are diverse and complex reasons for the overuse of antibiotics, and consequently, several strategies have been designed to combat this problem [5]. There have been increasing calls by government agencies, such as the US Centers for

Disease Control and Prevention, to embrace regular tracking and reporting (audits and feedback) as two of the core elements of stewardship programs [6].

Audits and feedback [7] can incorporate behavioral science elements, that is, clinical performance can be reviewed and compared among peers, and this has been widely reported as effective in reducing the inappropriate prescribing of antibiotics [1,2,8,9]. Despite strong support in the literature, there are at least two important factors that make audits and feedback hard to implement in real-life settings or make them ineffective altogether. First, without proper privacy safeguards, the risk of breaching regulatory compliance with regard to health data confidentiality increases. Additionally, underperforming general practitioners (GPs) may feel exposed and view the exercise as punitive [10]. The second factor, scalability, is important for efficiently analyzing vast amounts of distributed electronic health record (EHR) data to facilitate the regular scheduling of audit and feedback (A&F) programs. Regular scheduling is an important factor, since evidence shows that if audits and feedback are stopped, their benefits are likely to dissipate [11,12].

This study evaluates a scalable, privacy-preserving A&F system [13] in clinical settings. The system was deployed in three GP offices and used to generate feedback for 20 GPs on their antibiotic prescribing for selected respiratory tract infections (RTIs), which was viewed in comparison to the average performance of peers. The objective is to evaluate the GPs' sentiments toward (1) the privacy protections that the system offers, (2) the accuracy and appropriateness of using personal feedback as the basis for comparisons with peers, and (3) self-efficacy and the potential of the intervention to improve the prescribing of antibiotics.

Methods

Privacy-Preserving System for Audits and Feedback

The privacy-preserving system contains software components that are deployed at each health institution. The system extracts data daily from the local EHR system and transforms and loads the data into a database that conforms to a common data model [14].

A third party (denoted as a *coordinator*) aids the system without learning private information and is trusted to follow protocol specifications. This is a standard security model known as the *honest-but-curious adversarial model* [15]. In this study, the Norwegian Centre for E-health Research acted as the coordinator. The coordinator accepts queries for the aggregated performance indicators of GPs across health care institutions. The system uses privacy-preserving distributed data mining techniques [16-18] for executing statistical queries on the combined data of health care institutions without allowing any party to view the private data that health care institutions compute locally. Statistical results on the combined data of multiple health care institutions are not considered sensitive information and are therefore stored at the coordinator and retrieved through a web service.

The feedback report is generated locally by combining local personal indicators and the aggregated indicators retrieved from the coordinator. Access to the feedback report is restricted to the respective GP and is provided through a web client or email as an encrypted PDF file, and the decryption key is sent to a mobile phone. We refer interested readers to a study by Yigzaw et al [13] for an elaborated description of the system.

Study Setting

In 2019, we deployed the system in three GP offices in Norway, and 20 GPs received a single feedback report on their prescribing of antibiotics for selected RTIs. The GPs were then asked to fill out a web-based questionnaire for assessing their perceptions of the feedback received and how the feedback was presented.

Norwegian GPs are responsible for treating a set of patients and refer only those who need more specialized health services to hospitals [19]. Therefore, the 20 GPs had a total of around 19,345 patients on their lists. The number of registered patients varied during the study period because patients sometimes change their GP.

Audits and Feedback

The GPs received a single feedback report on their prescribing of antibiotics for selected RTIs for which antibiotics are generally not recommended if the patient is otherwise healthy (eg, during the first presentation of a current episode of an RTI or when the patient has no significant underlying comorbidity) [20]. Based on the *International Classification of Primary Care, Second Edition* [21], the selected RTIs were acute upper respiratory infection (R74), acute sinusitis (R75), acute laryngitis/tracheitis (R77), acute otitis media/myringitis (H71), and unspecified respiratory infections (R83). Acute bronchitis (R78) was also included because it is most often a viral infection, and the use of an antibiotic is rarely recommended.

Feedback was provided for the selected RTI cases between 2015 and 2018; both combined and separate feedback were provided for each of the diagnoses. All statistics were stratified by both year and antimicrobial spectrum (broad- and narrow-spectrum antibiotics). Based on the Anatomical Therapeutic Chemical classification system [22], narrow-spectrum antibiotics include β-lactamase–sensitive penicillins (J01CE), and broad-spectrum antibiotics include tetracyclines (J01A); β-lactam antibacterials and penicillins (J01C, excluding J01CE); other β-lactam antibacterials (J01D); sulfonamides and trimethoprim (J01E); macrolides, lincosamides, and streptogramins (J01F); and quinolone antibacterials (J01M).

The feedback report contains the number of cases that were diagnosed and treated with antibiotics. A set of performance indicators was presented as a time series graph that compares the performance indicators of a GP with the average indicators of all participating GPs. Performance indicators were based on the indicators proposed by the European Surveillance of Antimicrobial Consumption Network [23], as follows:

- The percentage of diagnosed cases treated with an antibacterial drug for systemic use

- The percentage of cases treated with narrow-spectrum antibiotics among all diagnosed cases treated with antibiotics
- The percentage of cases treated with broad-spectrum antibiotics among all diagnosed cases treated with antibiotics

Questionnaire

A panel of experts with clinical and medical informatics backgrounds developed the questionnaire. As shown in Multimedia Appendix 1, the questionnaire broadly covers aspects that were derived from the following three theoretical constructs of the Theory of Planned Behavior [24]: (1) attitude toward privacy and the prescribing of antibiotics, (2) subjective norms (reflection on the accuracy and appropriateness of using feedback as the basis for comparisons with peers), and (3) self-efficacy regarding changing prescribing behavior.

Statistical Analysis

Descriptive statistics were calculated to summarize the results of the categorical questions, including age groups and years of experience.

Ethical Considerations

This study was based on an anonymous questionnaire and therefore did not require ethics review board approval, according to Norwegian regulations. However, all participants gave their informed consent to participate in this study.

Results

Participants

The retrospective data that were generated by using the A&F system show that approximately 20% (2924/14,396, 20.3%) of all cases of the selected RTIs were treated with antibiotics, as illustrated in Figure 1.

A total of 14 GPs responded to the questionnaire, representing a 70% (14/20) response rate. Of these GPs, 36% (5/14) were aged 50 years or older, 57% (8/14) had more than 15 years of experience as a GP, and 79% (11/14) had a specialization in family medicine. The survey results on attitudes toward privacy, audits, and feedback are summarized in Table 1.

Figure 1. Antibiotic prescribing for all participating general practitioners between 2015 and 2018.

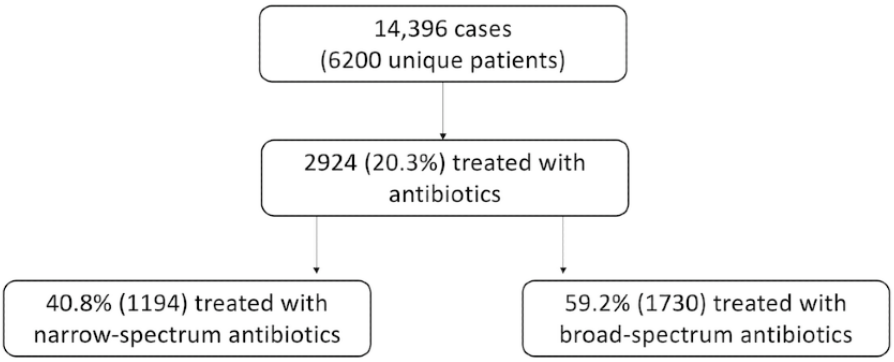


Table 1. Participant attitudes toward privacy, audits, and feedback (N=14).

	Very useful, n (%)	Useful, n (%)	Not useful, n (%)
Participant attitudes toward patients' privacy	10 (71)	4 (28)	0 (0)
Participant attitudes toward physicians' privacy	3 (21)	6 (42)	5 (35)
Participant attitudes toward feedback on own prescribing of antibiotics	9 (64)	4 (28)	1 (7)
Participant attitudes toward comparison with peers	6 (42)	7 (50)	1 (7)

Privacy Protection

We can observe from Table 1 that the GPs were unanimous about privacy protection for their patients, and 64% (9/14) of GPs considered the protection of their own privacy to be very useful or useful.

Feedback Report

Most of the GPs (13/14, 92%) reported that feedback on their prescribing of antibiotics and the ability to compare their performance indicators with those of their peers were useful or very useful. The clinicians preferred to receive feedback at regular intervals, such as half-yearly (5/13, 38%) and yearly (7/13, 53%; 1 response was invalid). Most of the clinicians (13/14, 92%) preferred to receive an encrypted feedback report through email rather than having it integrated into the EHR (3/14, 21%), receiving it on a secure website (1/14, 7%), or receiving it on a mobile app (0/14, 0%).

Effects of Audits and Feedback

Of the 14 GPs, 1 (7%) indicated the intention to change their antibiotic prescribing behavior based on the feedback received, 2 (14%) indicated that they did not intend to change, and the rest (n=11, 78%) were unsure of whether they would change.

Feedback Stratification

Most of the GPs (11/14, 78%) wanted their feedback to be stratified by the characteristics of their patient populations and diagnoses. Age and gender were the commonly requested patient characteristics for stratification. The clinicians also requested

stratification by chronic conditions, such as chronic obstructive pulmonary disease, asthma, cystic fibrosis, other airway comorbidities, acne, perioral dermatitis, and heart failure.

Discussion

Privacy Protection

Perhaps the major highlight of this study relates to privacy preservation, since the A&F system resolves most of the privacy concerns raised in similar studies. Our results show that the GPs value the protection of their own privacy as well as that of their patients, and this finding is in line with those of existing studies [25]. Our results appear to support our initial assumption that protecting the privacy of clinicians may increase their willingness to participate in quality improvement activities.

GPs in Norway often work in small practices with few or no peers; therefore, comparisons with peers across multiple institutions could be especially useful, and the privacy safeguards serve as further incentives for participation.

Feedback Report

Our results show that the GPs were in favor of receiving encrypted feedback reports through their emails. This contrasts with the common belief that integrating a clinical decision support system with EHR systems is an important success factor [26]. It is plausible that the GPs assumed that integration with the EHR would trigger distractive feedback alerts more frequently than those triggered by the desired half-yearly or yearly periodic feedback.

Effects of Audits and Feedback

The GPs appeared to value both seeing their antibiotic prescribing statistics and being able to compare these with those of their peers. Although this study did not measure whether the clinicians actually changed their prescribing behavior following feedback, it is likely that most of the participants had a close-to-average prescribing practice and would therefore not have an incentive to change their prescribing behavior. In a larger sample of clinicians, there would be a higher number of individuals that deviate from the mean and have an incentive to change their prescribing behavior.

Feedback Stratification

The GPs requested the stratification of feedback by patient characteristics, such as demographics and chronic diseases, of which both are known to influence the decision to prescribe antibiotics. Stratification provides important information to clinicians on whether their prescribing behavior can be justified based on patient characteristics and enables peer comparisons among clinicians with similar patient populations.

Limitations

An important limitation is that we could not evaluate GPs' responses based on their clinical performance, since the questionnaire was anonymous and access to feedback reports was limited to the respective clinician. For example, we were not able to assess how the self-reported intention to change related to an individual's current clinical performance.

Another limitation is that the feedback reports provided to GPs were not adjusted for comorbidities and age. These two variables are known to influence the decision to prescribe antibiotics, and this may have affected their responses to the survey.

We also noted that the average performance of peers might be skewed if there are outlier clinicians who are far from the mean on 1 side. Therefore, it might be necessary to exclude outliers or provide more statistics, such as SDs and percentiles, to enable detailed comparisons with peers.

The sample size in this study can be considered small but can also be viewed as appropriate for an in-depth case analysis for a pilot implementation of this type.

Recommendations, Implications for Practice, and Impacts

The findings from this study have important implications for practice, especially implications related to quality improvement programs at health care institutions. We single out 3 elements that we recommend for improving systems that support antibiotic stewardship programs. The first and basic element is protecting the privacy of clinicians as well as that of patients. Second, scalability should be considered so as to enable comparisons among peers across multiple health care providers. Comparisons across health care providers are especially useful for small practices like GP offices, since such comparisons require a wide pool of similar patient groups and can be adjusted for comorbidities and age. Finally, human factors should be considered. For example, the frequency of feedback should be considered, since clinicians may not want distractive feedback alerts.

In terms of impact, our system offers a solution to key challenges that hamper antibiotic stewardship programs. It provides privacy guarantees to patients, clinicians, and health care institutions and provides the scalability required to ensure that long-term audits and feedback can be sustainable parts of stewardship programs.

The impact for Norway might be limited, since Norway is a small country with relatively lower antibiotic prescribing rates compared to those of other European countries [27]. For larger countries with differently organized general practices, a higher effect can be expected, since audits and feedback are the most effective in places where the baseline performance is poorer compared to that of the best general practices [7].

Conclusion

Our results show generally positive sentiments among GPs regarding the potential impact of our system, but it should be noted that this was an initial study with a relatively small number of participants. Therefore, the absolute number of clinicians with very high prescribing rates was likely to be low. In larger samples, we expect that a higher number of physicians would receive an incentive to change their practice.

Acknowledgments

We are very thankful for the general practitioners who participated in this study. We also would like to thank Peder A. Halvorsen, Torje Henriksen, Andrius Budrionis, and Luis Marco-Ruiz for their contributions to the design of the questionnaire that was used in this study. This work was supported by the European Commission's H2020 research and innovation program (project number: 826093) and the Research Council of Norway (grant 248150/O70). The funding bodies did not have any role in this study and the writing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire.
[DOC File , 170 KB - formative_v6i7e31650_app1.doc]

References

1. Gjelstad S, Høye S, Straand J, Brekke M, Dalen I, Lindbæk M. Improving antibiotic prescribing in acute respiratory tract infections: cluster randomised trial from Norwegian general practice (prescription peer academic detailing (Rx-PAD) study). BMJ 2013 Jul 26;347:f4403 [FREE Full text] [doi: 10.1136/bmj.f4403] [Medline: 23894178]
2. Gerber JS, Prasad PA, Fiks AG, Localio AR, Grundmeier RW, Bell LM, et al. Effect of an outpatient antimicrobial stewardship intervention on broad-spectrum antibiotic prescribing by primary care pediatricians: a randomized trial. JAMA 2013 Jun 12;309(22):2345-2352. [doi: 10.1001/jama.2013.6287] [Medline: 23757082]
3. Handlingsplan mot antibiotikaresistens i helsetjenesten. Ministry of Health and Care Services. URL: https://www.regjeringen.no/contentassets/915655269bc04a47928fce917e4b25f5/handlingsplan-antibiotikaresistens.pdf [accessed 2022-06-29]
4. Robinson JL. Commentary on 'Interventions to improve antibiotic prescribing practices in ambulatory care'. Evid Based Child Health 2006 Jun;1(2):693-694 [FREE Full text] [doi: 10.1002/ebch.28] [Medline: 32313517]
5. McDonagh M, Peterson K, Winthrop K, Cantor A, Holzhammer B, Buckley DI. Improving antibiotic prescribing for uncomplicated acute respiratory tract infections. In: Comparative Effectiveness Reviews, No. 163. Rockville, MD: Agency for Healthcare Research and Quality (US); Jan 2016.
6. CDC. Core elements of outpatient antibiotic stewardship | Antibiotic use | CDC. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/antibiotic-use/core-elements/outpatient.html [accessed 2021-06-11]
7. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev 2012 Jun 13(6):CD000259. [doi: 10.1002/14651858.CD000259.pub3] [Medline: 22696318]
8. Arnold SR, Straus SE. Interventions to improve antibiotic prescribing practices in ambulatory care. Cochrane Database Syst Rev 2005 Oct 19(4):CD003539 [FREE Full text] [doi: 10.1002/14651858.CD003539.pub2] [Medline: 16235325]
9. Meeker D, Linder JA, Fox CR, Friedberg MW, Persell SD, Goldstein NJ, et al. Effect of behavioral interventions on inappropriate antibiotic prescribing among primary care practices: A randomized clinical trial. JAMA 2016 Feb 09;315(6):562-570 [FREE Full text] [doi: 10.1001/jama.2016.0275] [Medline: 26864410]
10. Patel SJ, Saiman L, Duchon JM, Evans D, Ferng YH, Larson E. Development of an antimicrobial stewardship intervention using a model of actionable feedback. Interdiscip Perspect Infect Dis 2012;2012:150367 [FREE Full text] [doi: 10.1155/2012/150367] [Medline: 22500166]
11. Linder JA, Meeker D, Fox CR, Friedberg MW, Persell SD, Goldstein NJ, et al. Effects of behavioral interventions on inappropriate antibiotic prescribing in primary care 12 months after stopping interventions. JAMA 2017 Oct 10;318(14):1391-1392 [FREE Full text] [doi: 10.1001/jama.2017.11152] [Medline: 29049577]
12. Gerber JS, Prasad PA, Fiks AG, Localio AR, Bell LM, Keren R, et al. Durability of benefits of an outpatient antimicrobial stewardship intervention after discontinuation of audit and feedback. JAMA 2014 Dec 17;312(23):2569-2570. [doi: 10.1001/jama.2014.14042] [Medline: 25317759]
13. Yigzaw KY, Budrionis A, Marco-Ruiz L, Henriksen TD, Halvorsen PA, Bellika JG. Privacy-preserving architecture for providing feedback to clinicians on their clinical performance. BMC Med Inform Decis Mak 2020 Jun 22;20(1):116 [FREE Full text] [doi: 10.1186/s12911-020-01147-5] [Medline: 32571306]
14. Bellika JG, Henriksen T, Hurley J, Budrionis A, Marco-Ruiz L, Yigzaw KY, et al. Requirements to the data reuse application programming interface for electronic health record systems. Norwegian Centre for E-health Research. 2017. URL: https://ehealthresearch.no/files/documents/Rapporter/NSE-rapport_2017-10_Requirements-to-the-data-reuse-application-for-ehr.pdf [accessed 2022-06-29]
15. Lindell Y, Pinkas B. Secure multiparty computation for privacy-preserving data mining. J Priv Confid 2009;1(1):59-98 [FREE Full text] [doi: 10.29012/jpc.v1i1.566]

16. Hailemichael MA, Yigzaw KY, Bellika JG. Emnet: a system for privacy-preserving statistical computing on distributed health data. 2015 Jun Presented at: SHI 2015, The 13th Scandinavien Conference on Health Informatics; June 15-17, 2015; Tromsø, Norway.
17. Yigzaw KY, Bellika JG, Andersen A, Hartvigsen G, Fernandez-Llatas C. Towards privacy-preserving computing on distributed electronic health record data. 2013 Presented at: Middleware '13: 14th International Middleware Conference; December 9-13, 2013; Beijing, China. [doi: 10.1145/2541534.2541593]
18. Yigzaw KY, Michalas A, Bellika JG. Secure and scalable deduplication of horizontally partitioned health data for privacy-preserving distributed statistical computation. BMC Med Inform Decis Mak 2017 Jan 03;17(1):1 [FREE Full text] [doi: 10.1186/s12911-016-0389-x] [Medline: 28049465]
19. Llanwarne N, Newbould J, Burt J, Campbell JL, Roland M. Wasting the doctor's time? A video-elicitation interview study with patients in primary care. Soc Sci Med 2017 Mar;176:113-122 [FREE Full text] [doi: 10.1016/j.socscimed.2017.01.025] [Medline: 28135690]
20. Harris AM, Hicks LA, Qaseem A, High Value Care Task Force of the American College of Physicians and for the Centers for Disease Control and Prevention. Appropriate antibiotic use for acute respiratory tract infection in adults: Advice for high-value Care from the American College of Physicians and the Centers for Disease Control and Prevention. Ann Intern Med 2016 Mar 15;164(6):425-434 [FREE Full text] [doi: 10.7326/M15-1840] [Medline: 26785402]
21. PH3C Primary Health Care Classification Consortium. PH3C. URL: http://www.ph3c.org/4daction/w3_CatVisu/en/Articles.html?wCatIDAdmin=8 [accessed 2022-03-10]
22. WHOCC - Home. WHO Collaborating Centre for Drug Statistics Methodology. URL: https://www.whocc.no/ [accessed 2022-03-10]
23. Adriaenssens N, Coenen S. Disease-specific antibiotic prescribing quality indicators report. European Centre for Disease Prevention and Control. 2010 Sep 10. URL: https://ecdc.europa.eu/sites/portal/files/media/en/healthtopics/antimicrobial-resistance-and-consumption/antimicrobial-consumption/publications-documents/Documents/ESAC-Net-archive-report_disease_specific_antibiotic_prescribing_quality_indicators.pdf [accessed 2022-06-29]
24. Liu C, Liu C, Wang D, Deng Z, Tang Y, Zhang X. Determinants of antibiotic prescribing behaviors of primary care physicians in Hubei of China: a structural equation model based on the theory of planned behavior. Antimicrob Resist Infect Control 2019 Jan 30;8:23 [FREE Full text] [doi: 10.1186/s13756-019-0478-6] [Medline: 30733857]
25. El Emam K, Mercer J, Moreau K, Grava-Gubins I, Buckeridge D, Jonker E. Physician privacy concerns when disclosing patient data for public health purposes during a pandemic influenza outbreak. BMC Public Health 2011 Jun 09;11:454 [FREE Full text] [doi: 10.1186/1471-2458-11-454] [Medline: 21658256]
26. Marcos M, Maldonado JA, Martínez-Salvador B, Boscá D, Robles M. Interoperability of clinical decision-support systems and electronic health records using archetypes: a case study in clinical trial eligibility. J Biomed Inform 2013 Aug;46(4):676-689 [FREE Full text] [doi: 10.1016/j.jbi.2013.05.004] [Medline: 23707417]
27. Adriaenssens N, Coenen S, Versporten A, Muller A, Minalu G, Faes C, ESAC Project Group. European Surveillance of Antimicrobial Consumption (ESAC): outpatient antibiotic use in Europe (1997-2009). J Antimicrob Chemother 2011 Dec;66 Suppl 6:vi3-v12. [doi: 10.1093/jac/dkr453] [Medline: 22096064]

Abbreviations

A&F: audit and feedback
EHR: electronic health record
GP: general practitioner
RTI: respiratory tract infection

Edited by A Mavragani; submitted 01.07.21; peer-reviewed by M Salimi, M Davis; comments to author 02.11.21; revised version received 21.03.22; accepted 14.04.22; published 13.07.22.

Please cite as:
Yigzaw KY, Chomutare T, Wynn R, Berntsen GKR, Bellika JG
A Privacy-Preserving Audit and Feedback System for the Antibiotic Prescribing of General Practitioners: Survey Study
JMIR Form Res 2022;6(7):e31650
URL: https://formative.jmir.org/2022/7/e31650
doi:10.2196/31650
PMID:35830221

Original Paper

Active Virtual Reality for Chronic Primary Pain: Mixed Methods Randomized Pilot Study

Natalie Tuck[1,2], PhD; Catherine Pollard[1,3], MHlthPrac; Clinton Good[3], MPhil; Caitlin Williams[1], BHlthSc; Gwyn Lewis[1], PhD; Murray Hames[3], DipPT, DipMT; Tipu Aamir[3], MBBs, MRCPsych, FFPMANZCA; Debbie Bean[1,2], PhD

[1]The Health and Rehabilitation Research Institute, Auckland University of Technology, Auckland, New Zealand

[2]The Pain Management Unit, Department of Anaesthesiology and Perioperative Medicine, Waitematā District Health Board, Auckland, New Zealand

[3]The Auckland Regional Pain Service, Auckland District Health Board, Auckland, New Zealand

Corresponding Author:
Natalie Tuck, PhD
The Health and Rehabilitation Research Institute
Auckland University of Technology
90 Akoranga Drive
Northcote
Auckland, 0627
New Zealand
Phone: 64 21 2669363
Email: natalie.tuck@aut.ac.nz

Abstract

Background: The modern management of chronic pain is largely focused on improving functional capacity (often despite ongoing pain) by using graded activation and exposure paradigms. However, many people with chronic pain find functional activation programs aversive, and dropout rates are high. Modern technologies such as virtual reality (VR) could provide a more enjoyable and less threatening way for people with chronic pain to engage in physical activity. Although VR has been successfully used for pain relief in acute and chronic pain settings, as well as to facilitate rehabilitation in conditions such as stroke and cerebral palsy, it is not known whether VR can also be used to improve functional outcomes in people with chronic pain.

Objective: This study aimed to assess the feasibility of conducting an adequately powered randomized controlled trial (RCT) to test the efficacy of VR in a chronic pain treatment center and assess the acceptability of an active VR treatment program for patients in this setting.

Methods: For this mixed methods pilot study, which was designed to test the feasibility and acceptability of the proposed study methods, 29 people seeking treatment for chronic pain were randomized to an active VR intervention or physiotherapy treatment as usual (TAU). The TAU group completed a 6-week waitlist before receiving standard treatment to act as a no-treatment control group. The VR intervention comprised twice-weekly immersive and embodied VR sessions using commercially available gaming software, which was selected to encourage movement. A total of 7 VR participants completed semistructured interviews to assess their perception of the intervention.

Results: Of the 99 patients referred to physiotherapy, 53 (54%) were eligible, 29 (29%) enrolled, and 17 (17%) completed the trial, indicating that running an adequately powered RCT in this setting would not be feasible. Despite this, those in the VR group showed greater improvements in activity levels, pain intensity, and pain interference and reported greater treatment satisfaction and perceived improvement than both the waitlist and TAU groups. Relative effect sizes were larger when VR was compared with the waitlist (range small to very large) and smaller when VR was compared with TAU (range none to medium). The qualitative analysis produced the following three themes: VR is an enjoyable alternative to traditional physiotherapy, VR has functional and psychological benefits despite continued pain, and a well-designed VR setup is important.

Conclusions: The active VR intervention in this study was highly acceptable to participants, produced favorable effects when compared with the waitlist, and showed similar outcomes as those of TAU. These findings suggest that a confirmatory RCT is warranted; however, substantial barriers to recruitment indicate that incentivizing participation and using a different treatment setting or running a multicenter trial are needed.

(JMIR Form Res 2022;6(7):e38366) doi:10.2196/38366

KEYWORDS

chronic pain; virtual reality; VR; rehabilitation; serious games; physiotherapy; pain management; acceptability; intervention; feasibility

Introduction

Background

Chronic pain is a leading cause of disability worldwide and represents a significant burden to individuals, societies, and health care systems [1,2]. Virtual reality (VR) has been successfully used to supplement conventional care across diverse medical and rehabilitation settings, including stroke and cerebral palsy [3], and VR technologies are increasingly being used in pain treatment settings to provide pain relief and facilitate rehabilitation [4-6].

Within the field of pain, VR has been predominantly used in acute pain settings as a nonpharmacological approach to pain relief in people undergoing painful medical procedures such as needle insertion and burn care [7,8]. In addition, the past decade has seen growth in the use of VR in chronic pain settings, where studies have generally focused on using VR to provide immediate pain relief via distraction or relaxation [4,9-14]. These interventions typically involve participants being immersed in a pleasant and distracting setting in which they can interact with a simulated environment. For example, in a proof-of-principle study on 13 people with chronic pain, an immersive VR game designed to teach mindfulness-based stress reduction led to reductions in pain scores immediately after a 12-minute session [9]. In another study, 30 people with chronic pain played a VR game specifically designed for pain management, with results showing a reduction in pain scores during gameplay and from the pre- to posttreatment time points [12]. Similarly, a randomized controlled crossover study of 20 people with chronic pain found that an immersive VR game led to greater reductions in pain scores than the baseline and control conditions [10]. Despite representing a useful start, studies focusing on the application of VR for short-term pain relief have not assessed long-term outcomes. Although findings suggest that applications designed to be relaxing or distracting can provide immediate analgesia, whether this approach leads to sustained improvements in other pain outcomes is not known.

The successful management of chronic pain relies on maintaining or improving physical activity, often despite ongoing pain [15]. In addition to short-term pain relief, VR may be used as a tool to encourage healthy movement [16]. VR has been used to support functional rehabilitation for spinal cord injury [17], traumatic brain injury [18], Parkinson disease [19], stroke [20,21], and phantom limb pain [22]. In these contexts, VR has been used to help participants increase their range of movement using neuromodulation and exposure paradigms, and recent work has begun to explore whether these principles can also be applied to other chronic pain conditions. For example, in a well-designed feasibility study, 52 participants with chronic low back pain were randomized to either VR or no-treatment control [23]. The VR group played 15 minutes of virtual dodgeball, designed to gradually increase lumbar flexion for 3 consecutive days. The results indicated that the VR group had greater increases in lumbar flexion during gameplay, with participants reporting a strong positive response to the game. Although there were no between-group differences in pain and functional outcomes 4 to 6 days after treatment, the authors attributed this to the brevity of the intervention, and a randomized controlled trial (RCT) testing the efficacy of a 9-week VR intervention is currently in progress [24].

In related work exploring whether VR can be used in chronic pain settings to encourage activity, 2 recent RCTs compared VR with physiotherapy treatment as usual (TAU) among people with chronic neck pain [25,26]. Both studies were based on 44 participants and administered VR interventions comprising 8 training sessions over 4 weeks. One of the studies used a game designed specifically to improve neck function [25], whereas the other used commercially available games [26]. Both studies reported pain-relevant outcomes immediately and at follow-up, with both reporting that the VR group showed at least equivalence to TAU across outcomes [25,26].

These studies suggest that VR can be used to encourage healthy movement using conditioning and exposure paradigms; however, to date, studies have typically focused on specific pain sites using targeted exercises [24-27], often with specifically designed health applications [23,25,28]. Although specifically designed games are appropriate in some treatment settings [28], there are also circumstances where commercially available games have advantages, which may be true for chronic pain treatment settings where people tend to present with widespread pain or pain in multiple sites. In many cases, general function and activities of daily living are more important therapeutic targets than training specific joints or muscle groups [29,30]. In these instances, commercially available VR games may be useful for supporting sustained behavior change and adherence to physical activity more broadly [31,32].

Overall, it seems likely that commercially available, immersive, and embodied games that encourage full-body movement may be a less threatening and more enjoyable way for people with diverse pain conditions to improve their general function. However, the authors are only aware of 1 pilot study that has tested this possibility. In this study, 16 veterans with chronic pain participated in daily VR sessions using commercially available VR games over a 3-week period. The authors reported improvements in kinesiophobia, pain intensity, pain catastrophizing, and pain-specific functioning in some participants [33]; however, whether this approach is also suitable for patient groups in clinical settings has not been examined. Moreover, the perspectives of end users have not been adequately assessed [28], and perceptions of whether commercially available VR games are considered an acceptable treatment option among people seeking treatment for chronic pain are not known.

Study Aims

This mixed methods pilot study aimed to test the planned methods and approach by identifying whether it would be feasible to conduct an active VR RCT in a hospital-based chronic pain treatment center and whether active VR treatment would be considered acceptable to patients in this setting. The criteria for feasibility were as follows: (1) 30 participants recruited within a 6-month time frame; (2) ≥70% retention rate; and (3) effect sizes (ESs) for primary outcomes (pain intensity and interference) of ≥0.5, indicating that a sample of approximately 60 participants per treatment arm would be sufficient to detect an effect [34]. The criteria for acceptability to patients were that (1) session rating scales for enjoyment are ≥6 on Likert scales ranging from 0 to 10; (2) treatment satisfaction and perceived improvement are ≥4 on Likert scales ranging from 0 to 7; (3) no serious adverse events or increases in pain specific to the VR treatment; (4) outcomes for VR are at least equivalent to TAU; and (5) interviews indicate that participants enjoy the VR treatment, find it beneficial, and believe that it is an acceptable and appropriate intervention for chronic pain rehabilitation.

Methods

Ethics Approval

Ethics approval was granted by the Health and Disability Ethics Committee (New Zealand Ministry of Health; HDEC ref 19/CEN/106), and locality approval was granted by the Auckland District Health Board in July 2019. The trial was registered with the Australian New Zealand Clinical Trials Registry (ACTRN12619001170112; universal trial number U1111-1234-0487). Eligible patients were given a written information sheet, an informed consent form, and instructions for accessing the study website. Those interested in participating were advised to access the study website where the information sheet and consent form were replicated, and informed consent was indicated by clicking a link at the end of the web-based consent form. This action opened the baseline questionnaire, and the completion of this questionnaire enrolled participants in the study. Eligible participants who expressed interest in the clinic but did not complete the baseline questionnaire were contacted via telephone to address any questions or concerns they may have had about the study. Participants were able to complete informed consent procedures and baseline questionnaires in the clinic in a separate appointment if they preferred.

Study Design

A mixed methods randomized pilot study was conducted with assessment points before and after 3 treatment arms to assess the feasibility and acceptability of a planned RCT, comparing an active VR intervention with a waitlist control and physiotherapy TAU.

Participants

The study took place at a hospital-based interdisciplinary chronic pain center, The Auckland Regional Pain Service (TARPS), Auckland District Health Board, New Zealand. Patients attending TARPS and referred to physiotherapy between October 1, 2019, and November 30, 2020, were invited to participate. Inclusion criteria were having musculoskeletal pain, being aged 18 to 70 years, and being able to communicate in English. Participants were excluded if they had a severe medical or psychiatric condition, were receiving treatment for pain elsewhere, or if their health care was being funded by the Accident Compensation Corporation (New Zealand's accident compensation provider).

Procedures

Overview

Participants completed baseline questionnaires before being randomized using a web-based random number generator to either VR or waitlist followed by TAU. All participants wore activPAL activity monitors (PAL Technologies Ltd) and completed questionnaires at the start and end of each 6-week treatment period. Participants in the VR group completed session evaluations at each appointment, and 7 attended the semistructured interviews.

VR Intervention (Experimental Group)

Participants attended twice-weekly VR appointments for 6 weeks, supervised by a physiotherapist with 4 years of experience in using VR for chronic pain. The HTC Vive immersive VR system (HTC Corporation) was used with a head-mounted display and accompanying hand sensors. The VR software programs were run via a wall-mounted desktop display that allowed the physiotherapist to view the participant's visual field. Games that encouraged full-body movements were selected, and participants were guided to perform physically active tasks within the virtual environment and progressed through VR games at the discretion of the treating physiotherapist (Multimedia Appendix 1).

Waitlist

Patients assigned to TAU completed a 6-week no-intervention waitlist to act as a no-treatment control before receiving standard physiotherapy treatment. Outcome measures were completed before and after the waitlist period.

TAU Group

Following the 6-week waitlist, the participants attended 6 weeks of physiotherapy treatment with 1 of 2 physiotherapists, each with >10 years of experience working in this chronic pain treatment setting. Therapy included education related to pain neuroscience, fear avoidance, and deconditioning, and participants were given home-based exercise regimens and tailored gym-based activity programs focused on graded activation and exposure therapy.

Measures

Brief Pain Inventory

The Brief Pain Inventory (BPI) [35,36] measures pain intensity and interference on 11-point Likert scales. The BPI has adequate reliability and validity in populations with chronic pain [37] and is a recommended outcome measure for chronic pain clinical trials [38]. A 2-point reduction is considered the minimal clinically important difference (MCID) in pain intensity scores [38,39].

Kinesiophobia

The Tampa Scale of Kinesiophobia-13 (TSK-13) is a measure of fear of movement, injury, and reinjury [40,41]. Participants respond to items on 4-point Likert scales, with scores ranging from 13 to 52, with higher scores indicating greater fear of movement and (re)injury. The TSK-13 has sufficient reliability and validity in samples with chronic pain [42,43]. In this study, Cronbach α was .72, indicating adequate internal consistency. An 18% reduction in TSK-13 scores is considered an MCID [44].

Activity

Daily activity and step counts were collected using activPAL activity monitors during weeks 1 and 6 of VR, waitlist, and TAU. ActivPAL monitors have been validated in samples with chronic pain [45-46], and the average daily number of steps and activities (total minutes spent standing, walking, and cycling) were extracted.

Global Impression of Change

The Patient Global Impression of Change scale [47] is a single-item recommended pain outcome measure [38] of perceived improvement with treatment on a 7-point Likert scale ranging from 1 (*not at all improved*) to 7 (*very much improved*). Participants also indicated their satisfaction with treatment on a scale from 1 (*not at all satisfied*) to 7 (*very satisfied*).

Session Evaluation Questionnaire

Participants completed questionnaires at each VR treatment session to assess changes in pain intensity from the pre- to postsession time points, as well as the degree to which they found the VR session enjoyable and immersive, using 10-point Likert scales.

Semistructured Interviews

The qualitative component of this study was designed in accordance with the Critical Appraisal Skills Program (CASP) guidelines [48]. The first 7 participants to complete 8 VR treatment sessions attended 30 to 60-minute semistructured interviews designed to assess their expectations and perceived usefulness of the VR intervention, changes in pain and function, and suggestions for improvement.

Data Analysis

Quantitative data were analyzed using SPSS (version 27). Baseline demographics and clinical characteristics were summarized using descriptive statistics (Table 1). Recruitment and retention rates, pain outcomes, enjoyment, and immersion scores were calculated. As this was a feasibility study, significance testing was not conducted. Instead, Hedges g relative ESs were used to estimate the necessary sample sizes for an RCT using G*Power software, with an α of .05 and power of 0.80. Ratings of pain, enjoyment, and immersion collected at the end of each VR session were averaged across the total number of attended sessions to create mean pain change, enjoyment, and immersion scores.

For qualitative data, interviews were transcribed and analyzed using reflexive thematic analysis [49,50] by a member of the research team (CW), with support from 2 additional team members (NT and DB). The 5 steps of reflexive thematic analysis were followed, where phase 1 (familiarization) involved reading and rereading the data while taking notes on features of interest. In phase 2, data were coded by applying brief unique labels to all quotes that appeared relevant to the research question. For phase 3, initial themes were generated by grouping quotes with similar codes and examining the resulting sets of quotes. In phase 4, 3 themes were developed and reviewed by looking at how each code fit with the proposed theme, the data supporting each code and theme, and the degree to which each theme helped answer the research question. Finally, in phase 5, 3 themes were named and defined. As noted earlier, qualitative analysis was conducted primarily by CW, a female physiotherapy student with training in reflexive thematic analysis. Qualitative analysis was supported by NT and DB, both of whom have PhDs in health psychology and prior experience working in an interdisciplinary pain center. Both NT and DB currently hold positions as senior research fellows, with training in reflexive thematic analysis.

Table 1. Demographic and baseline clinical characteristics (N=20).

Characteristics	Total (n=20)	Treatment	
		VR[a] (n=10)	Waitlist and TAU[b] (n=10)
Age (years), mean (SD)	40.1 (16.2)	41.3 (17.7)	38.7 (15.3)
Sex (female), n (%)	13 (65)	8 (80)	5 (50)
Ethnicity, n (%)			
New Zealand European	12 (60)	6 (60)	6 (60)
Māori	2 (10)	1 (10)	1 (10)
Indian	3 (15)	2 (20)	1 (10)
Fijian	1 (5)	0 (0)	1 (10)
Other European	2 (10)	1 (10)	1 (10)
Employment, n (%)			
Full-time	5 (25)	2 (20)	3 (30)
Part-time	3 (15)	2 (20)	1 (10)
Retired	1 (5)	1 (10)	0 (0)
Unemployed	9 (45)	3 (30)	6 (60)
Student	2 (10)	2 (20)	0 (0)
Pain duration, n (%)			
<12 months	0 (0)	0 (0)	0 (0)
12-24 months	3 (15)	2 (20)	1 (10)
2-5 years	5 (25)	1 (10)	4 (40)
>5 years	12 (60)	7 (70)	5 (50)
Main pain location, n (%)			
Neck	2 (10)	0 (0)	2 (20)
Back	8 (40)	4 (40)	4 (40)
Stomach	2 (10)	1 (10)	1 (10)
Chest	1 (5)	1 (10)	0 (0)
Hips	1 (5)	1 (10)	0 (0)
Pelvis or groin	1 (5)	1 (10)	0 (0)
Knee	3 (15)	1 (10)	2 (20)
Leg	1 (5)	1 (10)	0 (0)
Foot	1 (5)	0 (0)	1 (10)
BPI[c] intensity, mean (SD)	8.3 (1.5)	8.4 (1.8)	8.1 (1.2)
BPI interference, mean (SD)	7.3 (1.6)	7.5 (1.7)	7.1 (1.5)
TSK-13[d], mean (SD)	33.5 (5.4)	32.3 (5.4)	34.6 (5.4)

[a]VR: virtual reality.

[b]TAU: treatment as usual.

[c]BPI: Brief Pain Inventory.

[d]TSK-13: Tampa Scale of Kinesiophobia-13.

Results

Quantitative Results

Recruitment and Retention

Between October 1, 2019, and November 30, 2020, a total of 99 non–Accident Compensation Corporation patients were assessed at TARPS and referred to physiotherapy. Of these 99 patients, 53 (54%) met the inclusion criteria, and 29 (29%) were enrolled in the study. Of the 29 patients, 13 withdrew (n=5 from the VR group, and n = 4 from each of the waitlist and TAU groups), resulting in a final sample of 20 participants, of whom 10 were in the VR group, 10 in the waitlist group, and 6 in the TAU group. The recruitment rate of 2 participants per month, as well as the dropout rate of 45%, indicated that it would not be feasible to recruit approximately 60 participants per trial arm in this clinical setting.

Intervention Parameters

All participants in the VR arm completed ≥7 appointments, and 30% (3/10) completed 12 appointments. There were no adverse events or increases in pain specific to the VR intervention. Immersion scores ranged from 8.4 to 9.6, and enjoyment scores ranged from 8.0 to 9.9, thereby meeting the acceptability criteria.

Pain Intensity and Interference (BPI)

There were medium ESs favoring VR over the waitlist for a reduction in BPI pain intensity (ES=0.52) and BPI pain interference (ES=0.50). This suggests that for an α of .05, and a power of 0.80, a sample of 60 and 64 per group would be sufficient to detect significant effects for pain intensity and interference, respectively (Table 2).

Table 2. Mean change scores and relative effect sizes for pain-relevant outcomes, perceived improvement, and satisfaction with treatment.

Outcomes	Waitlist (n=10)		TAU[a] (n=6)		VR[b] (n=10)		VR versus waitlist, effect size (95% CI)	VR versus TAU, effect size (95% CI)
	Participants, n (%)[c]	Values, mean (SD)	Participants, n (%)[c]	Values, mean (SD)	Participants, n (%)[c]	Values, mean (SD)		
ΔBPI[d] intensity	10 (100)	–0.30 (1.57)	6 (100)	–0.17 (2.32)	9 (90)	–1.00 (0.87)	–0.52 (–1.39 to 0.67)	–0.49 (–1.47 to 0.51)
ΔBPI interference	10 (100)	–1.10 (2.13)	6 (100)	–1.00 (1.39)	9 (90)	–2.06 (1.46)	–0.50 (–1.37 to 0.39)	–0.70 (–1.70 to 0.32)
ΔTSK-13[e]	10 (100)	–2.90 (5.17)	6 (100)	–4.00 (4.56)	9 (90)	–1.56 (5.57)	0.24 (–0.63 to 1.10)	0.44 (–0.55 to 1.42)
Change in daily steps	7 (70)	212 (2394)	6 (100)	1127 (2784)	8 (80)	852 (2934)	0.22 (–0.74 to 1.18)	–0.09 (–1.08 to 0.90)
Change in daily activity (minutes)	7 (70)	2.15 (59.03)	6 (100)	–21.05 (91.49)	8 (80)	19.45 (64.50)	0.26 (–0.70 to 1.22)	0.49 (–0.52 to 1.49)
Satisfaction (score: range 1-7)	9 (90)	4.78 (1.20)	6 (100)	5.83 (0.98)	9 (90)	6.11 (0.93)	1.18 (0.20 to 2.14)	0.28 (–0.71 to 1.25)
Improvement (score: range 1-7)	9 (90)	4.78 (0.83)	6 (100)	5.67 (1.03)	9 (90)	5.89 (0.78)	1.31 (0.31 to 2.28)	0.24 (–0.75 to 1.21)

[a]TAU: treatment as usual.

[b]VR: virtual reality.

[c]Number of valid participants with complete data.

[d]BPI: Brief Pain Inventory.

[e]TSK-13: Tampa Scale of Kinesiophobia-13.

Kinesiophobia

Small ESs favored the waitlist over VR (ES=0.24) and TAU over VR (ES=0.44). None of the groups met the criteria for MCID in kinesiophobia change scores (≥18% reduction; Table 2) [44].

Step Counts and Activity (Minutes Spent Standing, Walking, and Cycling)

There were small ESs favoring VR over the waitlist for both step count (ES=0.22) and activity scores (ES=0.26). Both VR and TAU met the criteria for an MCID in daily step counts (≥600-1100 step increase; Table 2) [51].

Treatment Satisfaction and Global Impression of Change

There were very large ESs favoring VR over the waitlist for treatment satisfaction (ES=1.18) and perceived improvement (ES=1.31) and small ESs favoring VR over TAU for treatment satisfaction (ES=0.28) and perceived improvement (ES=0.24; Table 2).

Qualitative Results

Analysis of the transcribed interviews generated three themes: (1) VR is an enjoyable alternative to traditional physiotherapy, (2) VR leads to functional and psychological benefits despite continued pain, and (3) the importance of a well-designed VR setup.

Theme 1: VR Is an Enjoyable Alternative to Traditional Physiotherapy

Participants were enthusiastic about trying VR rehabilitation and had positive expectations of treatment; for example, one of the participants said the following:

> *I had really high hopes...I thought it might actually take my pain away* [Male, 51 years, Fijian-Indian, chest and shoulder pain]

Participants also described VR as more enjoyable, accessible, and achievable than their previous experiences of physiotherapy and described VR as *fun* and *exciting*:

> *It's a really good way to incorporate fun activity into your life on a regular basis. And for someone who struggles to find the mental and physical energy to do anything like that, it's a really good pull to get you up.* [Female, 25 years, European, back pain]

Participants also described VR as improving activity levels by distracting them from pain. Several participants described VR as *brain training* and believed that taking part in the VR treatment changed the way that their nervous system processed pain:

> *By doing a movement and not having that immediate thought that it's going to hurt, it's already training my brain so that it doesn't go—nope you're not doing that!—and then sending pain signals everywhere.* [Female, 23 years, European, pelvic pain]

Despite the perceived benefits, some participants expressed uncertainty about whether VR could be considered a legitimate treatment. They explained that the enjoyable nature of VR was inconsistent with their expectations of physiotherapy:

> *With VR you don't really feel like it's treatment...you have an expectation that when you go to [physical therapy] they give you exercises or manual treatment so when you leave it feels like you've done something. But when you're doing VR, I don't feel like I've necessarily achieved any treatment.* [Female, 21 years, European, hip pain]

Overall, the participants had positive expectations about VR, the rationale for VR treatment was clear to them, and they found VR to be highly agreeable. For some participants, the enjoyable nature of VR meant that it did not feel like a legitimate treatment approach.

Theme 2: Functional and Psychological Benefits Despite Continued Pain

Participants said that they experienced pain relief during the VR sessions; however, this was not sustained after treatment. Despite ongoing pain, participants reported an overall increase in daily physical activity, improved strength, reduced stiffness, improvements in sitting and standing tolerances, and greater confidence in engaging in activities of daily living. Participants said that VR changed their perspectives of activity from something unpleasant and difficult to something that could be enjoyable:

> *It helped me understand that I can move and do more activity. I can go for a walk outside and enjoy it and not have to focus on being in pain all the time. So, I think it just made me realise that that was actually an option.* [Female, 21 years, European, hip pain]

The participants also felt that the VR sessions improved their moods. They described feeling happier and more content following the sessions and explained that VR provided relief from daily stressors and an opportunity to do something enjoyable:

> *Mentality wise, it made a big difference. I looked forward to coming to the VR sessions. I'd be like—yeah, I've got this pain but at least something's happening, and I have fun when I'm there.* [Female, 23 years, European, pelvic pain]

Overall, participants described the intervention as having a positive impact on their mood and function and reported increased confidence in participating in physical activities despite continued pain.

Theme 3: The Importance of a Well-Designed VR Setup

Most participants found the VR equipment easy to use and the games straightforward to understand and play. The exceptions were those whose first language was not English and who described difficulties in understanding game instructions. Participants emphasized the importance of a comfortable and adjustable headset, and the physical space was considered important, with all participants feeling constrained by the size of the room. There were a range of opinions regarding the importance of having a trained physiotherapist deliver the VR; however, all participants highlighted the importance of being supervised by someone with whom they could form a therapeutic alliance:

> *I think it makes you feel better that it's a trained physiotherapist. You knew they had that background and it just fills you with confidence a bit more.* [Female, 64 years, European, back pain]

Overall, this theme demonstrates the importance of investing in a good-quality VR setup that participants can use easily and comfortably in a safe and supervised environment and the value of developing a strong therapeutic relationship.

Discussion

Principal Findings

This mixed methods pilot study assessed the acceptability and feasibility of a VR RCT in a hospital-based chronic pain treatment center. The findings indicated that the VR intervention was highly acceptable to patients, with session rating scores, treatment satisfaction, and perceived improvement all surpassing the a priori acceptability criteria. Those in the VR treatment showed improvements in pain intensity, pain interference, step counts, and activity scores, with preliminary findings suggesting that VR may be superior to no treatment and equivalent to TAU. Qualitative data supported the quantitative findings and indicated that participants enjoyed the VR treatment and found it beneficial.

Despite the benefits, poor recruitment and high dropout rates suggest that it would not be feasible to conduct an adequately powered RCT in this setting. ESs for primary outcomes in this study indicate that a reasonable sample size of approximately 60 per group (power=0.80) would be adequate to detect effects. These ESs are consistent with recommended ESs for motor interventions [34] and comparable with another VR RCT currently in progress [24]; however, the recruitment rate of approximately 2 participants per month means that an adequately powered RCT would not be feasible. In addition, any future RCT assessing multiple pain-relevant outcomes would need to carefully select the most relevant outcome measures and correct for the family-wise error rate. Attrition was higher than expected, with approximately 45% of enrolled participants withdrawing from the study. Similar rates of withdrawal across all treatment arms suggest that attrition was not specific to the treatment, with the most frequently cited reasons for withdrawal being (1) comorbid mental health problems, (2) inability to attend regular appointments, and (3) loss of contact following SARS-COV-2 lockdowns. High attrition is common in chronic pain settings [11], and future work may benefit from encouraging participation by offering treatment in community settings, ensuring adequate parking, and providing transportation and remuneration for attendance. Finally, although the findings suggest that the involvement of a qualified physiotherapist is important, remotely delivered VR therapies have shown promise [11], and future RCTs may benefit from adapting active VR treatments for home use.

Comparison With Prior Work

Most prior studies exploring the role of VR in chronic pain have focused on specifically designed health applications [23,25], targeted singular pain sites [26], or tested the efficacy of VR among veterans with chronic pain [33]. This is the first study to assess the feasibility of conducting an RCT that tests whether commercially available active VR games may improve outcomes among people attending a tertiary-level chronic pain treatment center. Consistent with previous studies [23], high levels of engagement and treatment satisfaction were reported, and the findings suggest that active VR may be equivalent to standard physiotherapy [25,26]. When considering potential mechanisms, it has been hypothesized that VR might improve functional outcomes by reducing kinesiophobia, and prior work has documented reductions in kinesiophobia following VR interventions [26,33,52]. However, our findings did not support this hypothesis. Instead, qualitative data suggest that increases in general well-being and pain self-efficacy are more likely mechanisms. Positive mood states may be a protective factor in chronic pain [53,54], and a previous study has shown that a VR intervention among people with fibromyalgia led to improvements in general mood state, positive emotions, motivation, and self-efficacy [55]. Future RCTs may benefit from the inclusion of measures of pain self-efficacy [56] along with positive affect, sleep, and other pain-relevant quality of life measures to identify the mechanisms by which VR interventions improve chronic pain outcomes.

When considering game design and selection, commercially available games were used for this pilot study because of their benefits in terms of access, convenience, and cost. Using commercially available games in chronic pain rehabilitation also makes theoretical sense, as when pain is attributed to nociplastic mechanisms, the key interventional targets are general physical activity, stress reduction, and pain self-management [57] rather than focusing on specific movements, muscles, or joints. Despite the benefits of using specifically designed apps in some health settings [28,56], games that target specific pain sites, such as neck pain or back pain, would only be suitable for a subset of people attending chronic pain treatment centers, whereas commercially available games are more likely to be applicable to people with a range of pain conditions and pain sites and varying degrees of functional impairment. In addition, commercially available games appear to be at least equivalent to conventional therapy in diverse physical rehabilitation settings [32]. Overall, whether commercially available or specifically designed games are used, it is important that games are developed using fundamental design principles of reward, goals, challenge, and meaningful play [58], and it seems likely that combining active VR treatment with other established interventions such as pain neuroscience education, as well as strategies to downregulate autonomic arousal, may strengthen the positive effects of VR in chronic pain management and enhance our understanding of how VR can supplement conventional care.

Limitations

This study had several limitations. Primarily, comparing active VR with a no-treatment control and physiotherapy TAU allowed for comparisons between active VR and standard physiotherapy-led activity programs but did not provide insight into the most salient elements of VR or gaming interventions. Future work administering an active non-VR control, such as computer-based games using similar movements but without a head-mounted display, or comparing active VR with passive VR treatments that facilitate downregulation of autonomic arousal, would help to clarify the unique benefits specific to immersive, embodied, and active VR protocols over and above other VR and gaming platforms. Furthermore, this study deviates from a standard 3-arm trial design as, for practical purposes (given the small pool of participants), a single group covered 2 intervention arms (waitlist and TAU). Future work would benefit from recruiting separate groups of participants for the TAU and no-treatment arms. In addition, it is likely that patients interested in VR would have self-selected for the trial and potentially overreported the perceived benefits. Another limitation is that session rating scales were not collected in the TAU arm, meaning that between-group comparisons for session enjoyment and changes in pain scores immediately after treatment could not be made and that VR and TAU were delivered at different doses, with VR offered twice weekly and TAU delivered once weekly; thus, it is not known whether VR delivered at the same dose of TAU would have produced the benefits seen here. Finally, end users were not consulted in the study design phase, and future RCTs would benefit from engaging in a co-design process.

Conclusions

Despite these limitations, this mixed methods pilot study indicates that active VR is an acceptable treatment for patients

attending a tertiary-level chronic pain treatment center. Qualitative data suggest that participants enjoyed the VR treatment, found it beneficial, and believed that it was an acceptable component of chronic pain rehabilitation. Although outcomes for VR in this pilot study were superior to no treatment, and appeared to be at least equivalent to standard physiotherapy, these findings should only be interpreted as a basis for designing future adequately powered clinical trials. In particular, the finding that the ESs comparing VR with standard treatment were generally small suggests that there may not be clinically meaningful differences between these groups in a larger trial. Despite this, an adequately powered RCT appears justified as, if equivalence is found, then VR may be a useful adjunct or alternative to standard treatment for some people with chronic pain. Although a future RCT is warranted, low recruitment and poor retention rates indicate that this would not be feasible in the present setting. Future RCTs would benefit from incentivizing study participation and actively reducing dropout rates while considering a broader range of outcome measures to identify likely mechanisms. VR technology is increasingly affordable and accessible and may improve chronic pain outcomes by encouraging participation in activities in a novel and enjoyable way. Adequately powered RCTs with long-term follow-ups examining the recommended pain-relevant outcomes and potential mechanisms of action are warranted.

Acknowledgments

Internal Auckland University of Technology research funding was used to purchase activPAL devices. No other funding sources were used for this study.

Authors' Contributions

NT was involved in the conceptualization, methodology, formal analysis, project administration, and supervision of the study and in the writing of the original draft and review and editing of the manuscript. CP was involved in the methodology, formal analysis, and investigation of the study and in the writing of the original draft. CG was involved in the methodology and investigation of the study and in the review and editing of the manuscript. CW was involved in the formal analysis, investigation, and data curation of the study and in the writing of the original draft. GL was involved in the conceptualization, methodology, and supervision of the study and in the review and editing of the manuscript. MH was involved in the conceptualization and supervision of the study and in the review and editing of the manuscript. TA was involved in the conceptualization and supervision of the study and in the review and editing of the manuscript. DB was involved in the conceptualization, methodology, project administration, supervision, and funding acquisition of the study and in the review and editing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Virtual reality software programs graded from levels 1 to 6.
[DOCX File , 15 KB - formative_v6i7e38366_app1.docx]

References

1. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015 Aug 22;386(9995):743-800. [doi: 10.1016/S0140-6736(15)60692-4] [Medline: 26063472]
2. Buchbinder R, Blyth FM, March LM, Brooks P, Woolf AD, Hoy DG. Placing the global burden of low back pain in context. Best Pract Res Clin Rheumatol 2013 Oct;27(5):575-589. [doi: 10.1016/j.berh.2013.10.007] [Medline: 24315140]
3. Tieri G, Morone G, Paolucci S, Iosa M. Virtual reality in cognitive and motor rehabilitation: facts, fiction and fallacies. Expert Rev Med Devices 2018 Feb 10;15(2):107-117. [doi: 10.1080/17434440.2018.1425613] [Medline: 29313388]
4. Tack C. Virtual reality and chronic low back pain. Disabil Rehabil Assist Technol 2021 Aug 20;16(6):637-645. [doi: 10.1080/17483107.2019.1688399] [Medline: 31746250]
5. Mallari B, Spaeth EK, Goh H, Boyd BS. Virtual reality as an analgesic for acute and chronic pain in adults: a systematic review and meta-analysis. J Pain Res 2019;12:2053-2085 [FREE Full text] [doi: 10.2147/JPR.S200498] [Medline: 31308733]
6. Trost Z, France C, Anam M, Shum C. Virtual reality approaches to pain: toward a state of the science. Pain 2021 Feb 01;162(2):325-331. [doi: 10.1097/j.pain.0000000000002060] [Medline: 32868750]
7. Luo H, Cao C, Zhong J, Chen J, Cen Y. Adjunctive virtual reality for procedural pain management of burn patients during dressing change or physical therapy: a systematic review and meta-analysis of randomized controlled trials. Wound Repair Regen 2019 Jan 27;27(1):90-101. [doi: 10.1111/wrr.1] [Medline: 30480854]

8. Chan E, Foster S, Sambell R, Leong P. Clinical efficacy of virtual reality for acute procedural pain management: a systematic review and meta-analysis. PLoS One 2018 Jul 27;13(7):e0200987 [FREE Full text] [doi: 10.1371/journal.pone.0200987] [Medline: 30052655]
9. Gromala D, Tong X, Choo A, Karamnejad M, Shaw C. The virtual meditative walk: virtual reality therapy for chronic pain management. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Presented at: CHI '15: CHI Conference on Human Factors in Computing Systems; Apr 18 - 23, 2015; Seoul Republic of Korea. [doi: 10.1145/2702123.2702344]
10. Jin W, Choo A, Gromala D, Shaw C, Squire P. A virtual reality game for chronic pain management: a randomized, controlled clinical study. Stud Health Technol Inform 2016;220:154-160. [Medline: 27046570]
11. Garrett B, Taverner T, McDade P. Virtual reality as an adjunct home therapy in chronic pain management: an exploratory study. JMIR Med Inform 2017 May 11;5(2):e11 [FREE Full text] [doi: 10.2196/medinform.7271] [Medline: 28495661]
12. Jones T, Moore T, Choo J. The impact of virtual reality on chronic pain. PLoS One 2016 Dec 20;11(12):e0167523 [FREE Full text] [doi: 10.1371/journal.pone.0167523] [Medline: 27997539]
13. Wiederhold BK, Gao K, Sulea C, Wiederhold MD. Virtual reality as a distraction technique in chronic pain patients. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):346-352. [doi: 10.1089/cyber.2014.0207] [Medline: 24892196]
14. Dascal J, Reid M, IsHak W, Spiegel B, Recacho J, Rosen B, et al. Virtual reality and medical inpatients: a systematic review of randomized, controlled trials. Innov Clin Neurosci 2017;14(1-2):14-21. [Medline: 28386517]
15. Cohen SP, Vase L, Hooten WM. Chronic pain: an update on burden, best practices, and new advances. Lancet 2021 May;397(10289):2082-2097. [doi: 10.1016/s0140-6736(21)00393-7]
16. Gupta A, Scott K, Dukewich M. Innovative technology using virtual reality in the treatment of pain: does it reduce pain via distraction, or is there more to it? Pain Med 2018 Jan 01;19(1):151-159. [doi: 10.1093/pm/pnx109] [Medline: 29025113]
17. Villiger M, Bohli D, Kiper D, Pyk P, Spillmann J, Meilick B, et al. Virtual reality-augmented neurorehabilitation improves motor function and reduces neuropathic pain in patients with incomplete spinal cord injury. Neurorehabil Neural Repair 2013 Oct 10;27(8):675-683. [doi: 10.1177/1545968313490999] [Medline: 23757298]
18. Pietrzak E, Pullman S, McGuire A. Using virtual reality and videogames for traumatic brain injury rehabilitation: a structured literature review. Games Health J 2014 Aug;3(4):202-214. [doi: 10.1089/g4h.2014.0013] [Medline: 26192369]
19. Dockx K, Bekkers EM, Van den Bergh V, Ginis P, Rochester L, Hausdorff JM, et al. Virtual reality for rehabilitation in Parkinson's disease. Cochrane Database Syst Rev 2016 Dec 21;12:CD010760. [doi: 10.1002/14651858.CD010760.pub2] [Medline: 28000926]
20. Laver KE, Lange B, George S, Deutsch JE, Saposnik G, Crotty M. Virtual reality for stroke rehabilitation. Cochrane Database Syst Rev 2017 Nov 20;11:CD008349. [doi: 10.1002/14651858.CD008349.pub4] [Medline: 29156493]
21. Laver KE, Lange B, George S, Deutsch JE, Saposnik G, Crotty M. Virtual reality for stroke rehabilitation. Stroke 2018 Apr;49(4):e160-e161. [doi: 10.1161/STROKEAHA.117.020275]
22. Dunn J, Yeo E, Moghaddampour P, Chau B, Humbert S. Virtual and augmented reality in the treatment of phantom limb pain: a literature review. NeuroRehabilitation 2017;40(4):595-601. [doi: 10.3233/NRE-171447] [Medline: 28211829]
23. Thomas JS, France CR, Applegate ME, Leitkam ST, Walkowski S. Feasibility and safety of a virtual reality dodgeball intervention for chronic low back pain: a randomized clinical trial. J Pain 2016 Dec;17(12):1302-1317. [doi: 10.1016/j.jpain.2016.08.011] [Medline: 27616607]
24. France CR, Thomas JS. Virtual immersive gaming to optimize recovery (VIGOR) in low back pain: a phase II randomized controlled trial. Contemp Clin Trials 2018 Jun;69:83-91. [doi: 10.1016/j.cct.2018.05.001] [Medline: 29730393]
25. Rezaei I, Razeghi M, Ebrahimi S, Kayedi S, Rezaeian Zadeh A. A novel virtual reality technique (Cervigame®) compared to conventional proprioceptive training to treat neck pain: a randomized controlled trial. J Biomed Phys Eng 2019;9(3):355-366. [doi: 10.31661/jbpe.v0i0.556]
26. Tejera D, Beltran-Alacreu H, Cano-de-la-Cuerda R, Leon Hernández JV, Martín-Pintado-Zugasti A, Calvo-Lobo C, et al. Effects of virtual reality versus exercise on pain, functional, somatosensory and psychosocial outcomes in patients with non-specific chronic neck pain: a randomized clinical trial. Int J Environ Res Public Health 2020 Aug 16;17(16):5950 [FREE Full text] [doi: 10.3390/ijerph17165950] [Medline: 32824394]
27. Matheve T, Bogaerts K, Timmermans A. Virtual reality distraction induces hypoalgesia in patients with chronic low back pain: a randomized controlled trial. J Neuroeng Rehabil 2020 Apr 22;17(1):55 [FREE Full text] [doi: 10.1186/s12984-020-00688-0] [Medline: 32321516]
28. Tao G, Garrett B, Taverner T, Cordingley E, Sun C. Immersive virtual reality health games: a narrative review of game design. J Neuroeng Rehabil 2021 Feb 11;18(1):31 [FREE Full text] [doi: 10.1186/s12984-020-00801-3] [Medline: 33573684]
29. Ambrose KR, Golightly YM. Physical exercise as non-pharmacological treatment of chronic pain: why and when. Best Pract Res Clin Rheumatol 2015 Feb;29(1):120-130. [doi: 10.1016/j.berh.2015.04.022] [Medline: 26267006]
30. Booth J, Moseley GL, Schiltenwolf M, Cashin A, Davies M, Hübscher M. Exercise for chronic musculoskeletal pain: a biopsychosocial approach. Musculoskeletal Care 2017 Dec 30;15(4):413-421. [doi: 10.1002/msc.1191] [Medline: 28371175]
31. Baranowski T, Buday R, Thompson DI, Baranowski J. Playing for real: video games and stories for health-related behavior change. Am J Prev Med 2008 Jan;34(1):74-82. [doi: 10.1016/j.amepre.2007.09.027] [Medline: 18083454]

32. Bonnechère B, Jansen B, Omelina L, Van Sint Jan S. The use of commercial video games in rehabilitation: a systematic review. Int J Rehabil Res 2016 Dec;39(4):277-290. [doi: 10.1097/MRR.0000000000000190] [Medline: 27508968]
33. Fowler CA, Ballistrea LM, Mazzone KE, Martin AM, Kaplan H, Kip KE, et al. Virtual reality as a therapy adjunct for fear of movement in veterans with chronic pain: single-arm feasibility study. JMIR Form Res 2019 Oct 30;3(4):e11266 [FREE Full text] [doi: 10.2196/11266] [Medline: 31670696]
34. Dobkin BH. Progressive staging of pilot studies to improve phase III trials for motor interventions. Neurorehabil Neural Repair 2009 Oct 30;23(3):197-206. [doi: 10.1177/1545968309331863] [Medline: 19240197]
35. Cleeland CS, Ladinsky JL, Serlin RC, Thuy NC. Multidimensional measurement of cancer pain: comparisons of US and Vietnamese patients. J Pain Symp Manag 1988;3(1):23-27. [doi: 10.1016/0885-3924(88)90134-0]
36. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Ann Acad Med Singapore (1994 Mar) 23(2):129-38. Rehab Oncol 1995;13(1):29-30. [doi: 10.1097/01893697-199513010-00022]
37. Tan G, Jensen MP, Thornby JI, Shanti BF. Validation of the Brief Pain Inventory for chronic nonmalignant pain. J Pain 2004 Mar;5(2):133-137. [doi: 10.1016/j.jpain.2003.12.005] [Medline: 15042521]
38. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, IMMPACT. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005 Jan;113(1-2):9-19. [doi: 10.1016/j.pain.2004.09.012] [Medline: 15621359]
39. Atkinson TM, Mendoza TR, Sit L, Passik S, Scher HI, Cleeland C, et al. The Brief Pain Inventory and its "pain at its worst in the last 24 hours" item: clinical trial endpoint considerations. Pain Med 2010 Mar 01;11(3):337-346. [doi: 10.1111/j.1526-4637.2009.00774.x] [Medline: 20030743]
40. Miller RP, Kori SH, Todd DD. The Tampa Scale. Clin J Pain 1991;7(1):51. [doi: 10.1097/00002508-199103000-00053]
41. Vlaeyen J, Kole-Snijders A, Boeren R, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain 1995;62(3):363-372. [doi: 10.1016/0304-3959(94)00279-N]
42. Roelofs J, Goubert L, Peters M, Vlaeyen J, Crombez G. The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. Eur J Pain 2004 Oct;8(5):495-502. [doi: 10.1016/j.ejpain.2003.11.016] [Medline: 15324781]
43. Goubert L, Crombez G, Van Damme S, Vlaeyen JW, Bijttebier P, Roelofs J. Confirmatory factor analysis of the Tampa Scale for Kinesiophobia: invariant two-factor model across low back pain patients and fibromyalgia patients. Clin J Pain 2004;20(2):103-110. [doi: 10.1097/00002508-200403000-00007] [Medline: 14770050]
44. Monticone M, Ambrosini E, Rocca B, Foti C, Ferrante S. Responsiveness of the Tampa Scale of Kinesiophobia in Italian subjects with chronic low back pain undergoing motor and cognitive rehabilitation. Eur Spine J 2016 Sep 29;25(9):2882-2888. [doi: 10.1007/s00586-016-4682-2] [Medline: 27356516]
45. Ryan C, Grant P, Gray H, Newton M, Granat M. Measuring postural physical activity in people with chronic low back pain. BMR 2008 Apr 29;21(1):43-50. [doi: 10.3233/BMR-2008-21106]
46. Treacy D, Hassett L, Schurr K, Chagpar S, Paul SS, Sherrington C. Validity of different activity monitors to count steps in an inpatient rehabilitation setting. Phys Ther 2017 May 01;97(5):581-588. [doi: 10.1093/ptj/pzx010] [Medline: 28339904]
47. Guy W. ECDEU Assessment Manual for Psychopharmacology. Rockville, MD: U.S. Dept. of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute of Mental Health, Psychopharmacology Research Branch, Division of Extramural Research Programs; 1976.
48. CASP checklist. CASP. URL: https://casp-uk.net/casp-tools-checklists/ [accessed 2021-11-05]
49. Terry G, Hayfield N. Essentials of Thematic Analysis. Washington, D.C., United States: American Psychological Association; 2021.
50. Braun V, Clarke V, Hayfield N, Terry G. Thematic analysis. In: Handbook of Research Methods in Health Social Sciences. Singapore: Springer; 2018.
51. Demeyer H, Burtin C, Hornikx M, Camillo CA, Van Remoortel H, Langer D, et al. The minimal important difference in physical activity in patients with COPD. PLoS One 2016 Apr 28;11(4):e0154587 [FREE Full text] [doi: 10.1371/journal.pone.0154587] [Medline: 27124297]
52. Yilmaz Yelvar GD, Çırak Y, Dalkılınç M, Parlak Demir Y, Guner Z, Boydak A. Is physiotherapy integrated virtual walking effective on pain, function, and kinesiophobia in patients with non-specific low-back pain? Randomised controlled trial. Eur Spine J 2017 Feb 15;26(2):538-545. [doi: 10.1007/s00586-016-4892-7] [Medline: 27981455]
53. Aaron RV, Finan PH, Wegener ST, Keefe FJ, Lumley MA. Emotion regulation as a transdiagnostic factor underlying co-occurring chronic pain and problematic opioid use. Am Psychol 2020 Sep;75(6):796-810. [doi: 10.1037/amp0000678] [Medline: 32915024]
54. Lutz J, Gross RT, Vargovich AM. Difficulties in emotion regulation and chronic pain-related disability and opioid misuse. Addict Behav 2018 Dec;87:200-205. [doi: 10.1016/j.addbeh.2018.07.018] [Medline: 30053706]
55. Herrero R, García-Palacios A, Castilla D, Molinari G, Botella C. Virtual reality for the induction of positive emotions in the treatment of fibromyalgia: a pilot study over acceptability, satisfaction, and the effect of virtual reality on mood. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):379-384. [doi: 10.1089/cyber.2014.0052] [Medline: 24892201]
56. Nicholas M. The pain self-efficacy questionnaire: taking pain into account. Eur J Pain 2007 Feb;11(2):153-163. [doi: 10.1016/j.ejpain.2005.12.008] [Medline: 16446108]

57. Fitzcharles M, Cohen SP, Clauw DJ, Littlejohn G, Usui C, Häuser W. Nociplastic pain: towards an understanding of prevalent pain conditions. Lancet 2021 May;397(10289):2098-2110. [doi: 10.1016/s0140-6736(21)00392-5]
58. Barrett N, Swain I, Gatzidis C, Mecheraoui C. The use and effect of video game design theory in the creation of game-based systems for upper limb stroke rehabilitation. J Rehabil Assist Technol Eng 2016 May 09;3:2055668316643644 [FREE Full text] [doi: 10.1177/2055668316643644] [Medline: 31186903]

Abbreviations

BPI: Brief Pain Inventory
CASP: Critical Appraisal Skills Program
ES: effect size
MCID: minimal clinically important difference
RCT: randomized controlled trial
TARPS: The Auckland Regional Pain Service
TAU: treatment as usual
TSK-13: Tampa Scale of Kinesiophobia-13
VR: virtual reality

Edited by A Mavragani; submitted 29.03.22; peer-reviewed by B Garrett, K Harrington; comments to author 09.05.22; revised version received 23.05.22; accepted 07.06.22; published 13.07.22.

Please cite as:
Tuck N, Pollard C, Good C, Williams C, Lewis G, Hames M, Aamir T, Bean D
Active Virtual Reality for Chronic Primary Pain: Mixed Methods Randomized Pilot Study
JMIR Form Res 2022;6(7):e38366
URL: https://formative.jmir.org/2022/7/e38366
doi:10.2196/38366
PMID:35830224

Original Paper

Evaluation of a Text Message–Based COVID-19 Vaccine Outreach Program Among Older Patients: Cross-sectional Study

Naheed Ahmed[1,2], PhD; Christian Boxley[3], BS; Ram Dixit[3], MS; Seth Krevat[3], MD; Allan Fong[3], MS; Raj M Ratwani[3], PhD; Deliya B Wesley[2,4], PhD

[1]NYU Grossman School of Medicine, New York, NY, United States

[2]MedStar Center for Health Equity Research, Hyattsville, MD, United States

[3]MedStar Health National Center for Human Factors in Healthcare, Hyattsville, MD, United States

[4]Mathematica, Washington, DC, United States

Corresponding Author:
Naheed Ahmed, PhD
NYU Grossman School of Medicine
180 Madison Avenue, 9th Floor
New York, NY, 10016
United States
Email: Naheed.Ahmed@nyulangone.org

Abstract

Background: COVID-19 vaccines are vital tools in the defense against infection and serious disease due to SARS-CoV-2. There are many challenges to implementing mass vaccination campaigns for large, diverse populations from crafting vaccine promotion messages to reaching individuals in a timely and effective manner. During this unprecedented period, with COVID-19 mass vaccination campaigns essential for protecting vulnerable patient populations and attaining herd immunity, health care systems were faced with the dual challenges of vaccine outreach and distribution.

Objective: The aim of this cross-sectional study was to assess the effectiveness of a COVID-19 vaccine text outreach approach for patients aged 65 years and older. Our goal was to determine whether this approach was successful in scheduling patients for COVID-19 vaccine appointments.

Methods: We developed SMS text messages using the Tavoca platform. These messages informed patients of their vaccine eligibility and allowed them to indicate their interest in scheduling an appointment via a specific method (email or phone) or indicate their lack of interest in the vaccine. We tracked the status of these messages and how patients responded. Messages were sent to patients aged 65 years and older (N=30,826) at a nonprofit health care system in Washington, DC. Data were collected and examined from January 14 to May 10, 2021. Data were analyzed using multivariate multinomial and binary logistic regression models in SAS (version 9.4; SAS Institute Inc).

Results: Approximately 57% of text messages were delivered to patients, but many messages received no response from patients (40%). Additionally, 42.1% (12,978/30,826) of messages were not delivered. Of the patients who expressed interest in the vaccine (2938/30,826, 9.5%), Black or African American patients preferred a phone call rather than an email for scheduling their appointment (odds ratio [OR] 1.69, 95% CI 1.29-2.21) compared to White patients. Patients aged 70-74 years were more likely to schedule an appointment (OR 1.38, 95% CI 1.01-1.89) than those aged 65-69 years, and Black or African American patients were more likely to schedule an appointment (OR 2.90, 95% CI 1.72-4.91) than White patients.

Conclusions: This study provides insights into some advantages and challenges of using a text messaging vaccine outreach for patients aged 65 years and older. Lessons learned from this vaccine campaign underscore the importance of using multiple outreach methods and sharing of patient vaccination status between health systems, along with a patient-centered approach to address vaccine hesitancy and access issues.

(JMIR Form Res 2022;6(7):e33260) doi:10.2196/33260

KEYWORDS

vaccine outreach; text messaging; elderly patients; evaluation; smartphone; text message; SMS; appointment; elderly; older adults; vaccine; effectiveness; engagement; cross-sectional

Introduction

Authorization of the emergency use of COVID-19 vaccines, which were developed and tested in record time, was a seminal moment in efforts to control the COVID-19 pandemic [1]. These vaccines are vital tools in the prevention of infection and serious disease from SAS-CoV-2, and COVID-19 mass vaccination campaigns have been essential for protecting vulnerable patient populations and achieving herd immunity. Launching these vaccination campaigns has been challenging for several reasons related to logistical issues of producing, storing, and transporting vaccines, and patient-level factors such as vaccine access and vaccine hesitancy [2-5]. Health care systems were starting points for vaccination campaigns given their direct access to patient populations and having the personnel and resources to store vaccines and vaccinate patients. However, the COVID-19 vaccination effort was unprecedented in its scale, and many health care systems across the United States were faced with challenges of outreach, equity, scheduling, and administration.

With no blueprint for this type of vaccination campaign, health care systems—particularly those with large diverse patient populations—had to quickly design and launch outreach efforts to patients [2]. These efforts were further complicated by ongoing challenges with patient distrust of medical institutions, vaccine hesitancy and access barriers, and interoperability of health record databases [3,6]. Outreach to patients was complicated by out-of-date or incorrect demographic information, which was necessary for determining vaccine eligibility of each patient. Other complications included the absence of coordination between adjacent health care systems, as patients had the opportunity to be vaccinated at other health care sites, but these vaccination records were not shared among health care systems.

Reflecting on these challenges, this cross-sectional study used a not-for-profit health system as a case study to examine the intersection of health information technology and health disparities in vaccine outreach efforts. The focus of this study was on patients aged 65 years and older as they were among the first to be eligible for the COVID-19 vaccine. Additionally, patients in this age group may face additional challenges to vaccination, such as complicated medical conditions, transportation needs, and reliance on caregivers to assist with medical decision-making [7]. Findings from this study generated important lessons learned for ongoing efforts to increase COVID-19 vaccination rates among patients aged 65 years and older.

Methods

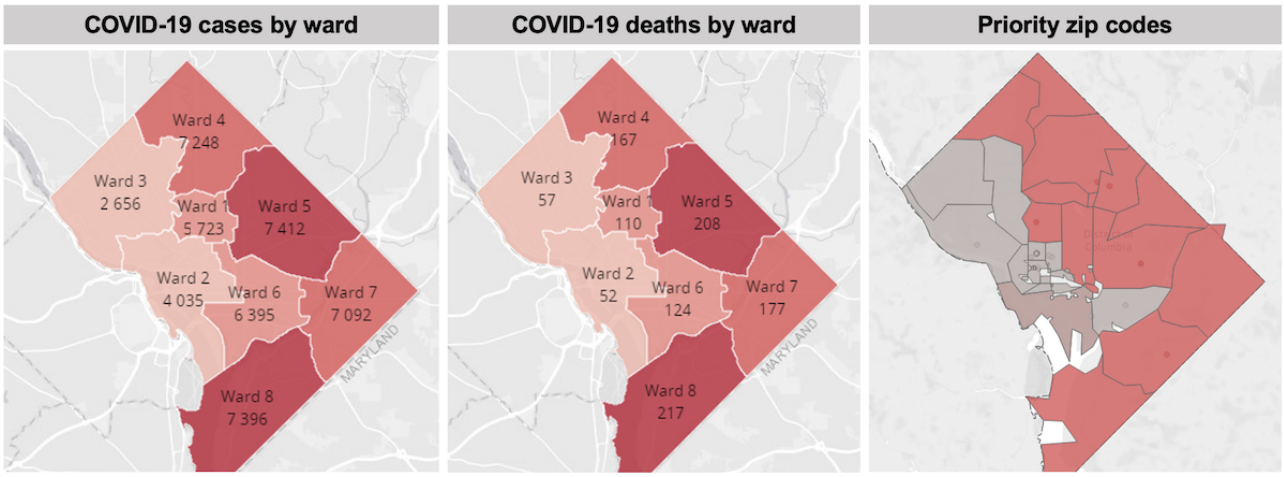
Study Site

This study took place in a not-for-profit health system with 10 hospitals, over 280 outpatient clinics, and nearly 2 million patients in the mid-Atlantic region of the United States.

Target Population

For this study, COVID-19 vaccine outreach to patients in the Washington, DC catchment area was examined. In Washington, DC, the local health department prioritized patients aged 65 years and older in the initial rollout of the vaccine to the public [8,9]. Additionally, certain zip codes were prioritized for outreach efforts based on the disproportionate impact of COVID-19 in terms of morbidity and mortality rates. A side-by-side comparison of priority zip codes with COVID-19 cases and deaths is presented in Figure 1.

Figure 1. Maps of COVID-19 cases and deaths by ward and priority zip codes for vaccination.



Similar to nationwide trends, COVID-19 cases and deaths were concentrated in areas with greater proportions of racial or ethnic minorities and low-socioeconomic-status populations in Washington, DC. Four health systems in the region, the local health department, and local community organizations (eg, faith-based and nonprofit organizations) conducted outreach to residents in priority zip codes, including virtual webinars and door-to-door canvassing for the purpose of answering questions and concerns about the COVID-19 vaccines [10].

Study Population

The health system serves approximately 45,000 Washington, DC residents aged 65 years and older, which required a large-scale outreach operation to contact, screen, and schedule these patients for vaccination. It is important to note that local policy at the time permitted health systems to only vaccinate

patients who had previously been served by the health system. Information on whether a patient was vaccinated at a different facility was unavailable.

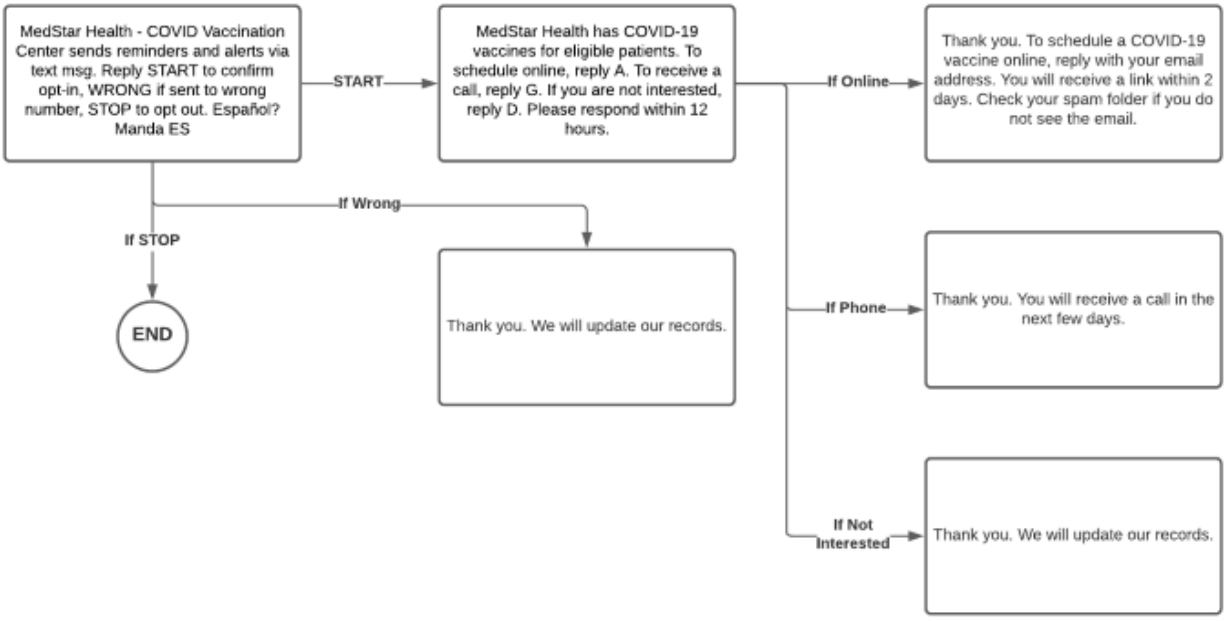
Inclusion and Exclusion Criteria

Eligibility was limited to patients aged 65 years and older with a Washington, DC address per local health department vaccination guidelines. Patients who had been seen at any of the health system facilities in the last 5 years were contacted about the COVID-19 vaccine.

Text Message Vaccine Outreach

During the initial phase of outreach, a dedicated call center contacted patients who met vaccine eligibility criteria. Operators were only able to speak to 41% (1093/2670) of patients contacted, and only 42% (458/1093) of connected calls resulted in scheduled appointments. Thus, only 17% of total calls resulted in a scheduled appointment. Given the relatively slow pace and resource-intensiveness of this approach, a transition was made to an automated text messaging and follow-up strategy, in which patients were sent a text message informing them that they were eligible to receive the vaccine and asking if they were interested in scheduling an appointment. The text messaging platform, Tavoca, was used to create message templates informing patients of their vaccine eligibility and allowing them to indicate their interest in scheduling via a specific method (email or phone) or indicate their lack of interest in the vaccine. Text messages were drafted on the basis of character limits of the Tavoca platform, and the text sequence is shown in Figure 2.

Figure 2. Text messaging flow.



Batches of text messages were sent on the basis of a prioritization algorithm that ranked patients by age and priority zip code, starting with the oldest patients in priority zip codes and then patients in nonpriority zip codes. Patients who indicated they were interested in the vaccine were asked which communication platform they preferred (phone or email), and depending on their response, they were referred to a call center or a web-based scheduling system. There was a high volume of messages not delivered owing to phone numbers no longer being in operation and some being landlines, which do not accept text messages.

Data Collection and Management

Patient information was extracted from the electronic health record and scheduling systems. Data were extracted and analyzed from January 14 to May 10, 2021. During this stage of the vaccine rollout, only the Pfizer and Moderna COVID-19 vaccines were available to patients. For this analysis, patients who shared the same phone number (n=14,419) were excluded, as there was no way to accurately determine which of the multiple patients associated with one phone number responded to the text message (Table 1). Confirmed text messaging interaction data were available for 30,826 patients who had unique phone numbers, which were analyzed for this study (Table 1). The demographic characteristics of the excluded patients were similar to those of the patients included in the study sample.

Table 1. Demographic Information for excluded patients (N=14,419) and study participants (N=30,826).

Characteristics	Excluded patients, n (%)	Study participants, n (%)
Sex		
Female	8224 (57.0)	18,399 (59.7)
Male	6181 (42.8)	12,427 (40.3)
Residence		
Priority zip code	11,408 (79.1)	24,211 (78.5)
Nonpriority zip code	3011 (20.8)	6615 (21.5)
Race or ethnicity		
Black or African American	8939 (61.9)	19,372 (62.8)
Hispanic	477 (3.31)	706 (2.3)
Asian	150 (1.04)	208 (0.67)
White	2885 (20.0)	7315 (23.7)
Other	1968 (13.6)	3225 (10.5)
Age (years)		
65-69	3506 (24.3)	9804 (31.8)
70-74	3186 (22.1)	7770 (25.2)
75-79	2527 (17.5)	5151 (16.7)
80-84	2092 (14.5)	3504 (11.4)
85-90	1505 (10.4)	2273 (7.4)
>90	1603 (11.1)	2324 (7.5)
Dose 1 received		
Yes	744 (5.1)	239 (0.77)
No	13,675 (94.8)	30,587 (99.2)
Dose 2 received		
Yes	643 (4.4)	232 (0.75)
No	13,776 (95.5)	30,594 (99.2)

Independent Variables

A total of 7 independent variables related to patient demographics and health system utilization in this study. Demographic variables included residing in a priority zip code (priority=1 and nonpriority=0), binary variables for each racial or ethnic group (Black or African American, Hispanic, Asian, White, and other), binary variables for each age group (65-69 years, 70-74 years, 75-79 years, 80-84 years, 85-89 years, and >90 years), and sex (male=1, female=0). Health system utilization variables included having a primary care provider (yes=1, no=0) and total visits in the past 5 years (continuous). A series of interaction variables were created for priority zip code and age and for priority zip codes and race or ethnicity.

Outcome Variables

There were 4 primary outcome variables of interest. The first outcome variable was the status of the initial text message sent to patients: 1=delivered and interested in vaccine; 2=delivered and no response; 3=delivered and not interested in the vaccine; and 4=not delivered (number belongs to a landline or is no longer in operation). Among patients who expressed interest in the COVID-19 vaccine, the preferred communication method for scheduling the vaccine appointment (phone=1, email=0) and if an appointment was scheduled (scheduled=1, not scheduled=0) were assessed.

Statistical Analysis

First, descriptive statistics were assessed to determine sample characteristics. Next, multivariate regression models were used to assess the association between patient demographics (priority zip codes, race or ethnicity, age, and sex) and outcomes of interest. Multivariate multinomial regression models were used for the first outcome of interest (text message status) and multivariate logistic regression for the other outcomes (preferred communication method and appointment scheduled). It was not possible to analyze vaccine uptake in this study owing to sample size limitations. Data were analyzed using SAS (version 9.4; SAS Institute Inc).

Ethics Approval

This study received approval from the institutional review board of MedStar Health Research Institute (STUDY00002197).

Results

Sample Characteristics

Demographic information on study patients is detailed in Multimedia Appendix 1. The majority of patients (24,211/30,826, 78.5%) live in a priority zip code, which implies that there is high COVID-19 morbidity and mortality, 59.7% (18,399/30,826) of participants are female, and 62.8% (19,372/30,826) of participants are Black or African American. Further, 57% (17,848/30,826) of text messages were delivered to patients and 42.1% (12,978/30,826) of messages were not delivered (number belongs to a landline or is no longer in operation). Of the messages that were delivered, 40% (12,333/17,848) received no response. A few patients (2,938/17,848, 9.5%) expressed interest in getting the COVID-19 vaccine. Among patients who expressed interest in the vaccine, only 253 scheduled an appointment. A majority of patients who scheduled an appointment completed their first (214/226, 87.5%) and second (207/226, 80%) vaccine doses.

Regression Models

First, differences among the 4 status options for the initial text message sent to patients were examined: interested in the vaccine, not interested in the vaccine, no response, and text message not delivered (Table 2). Notable findings include patients residing in priority zip codes who were more likely to "not respond" (odds ratio [OR] 1.19, 95% CI 1.12-1.27) than for message "not delivered" in comparison with patients living in nonpriority zip codes. In models that included age, among patients aged 70 years and older, the message was more likely to not be delivered than for there to be patient engagement with the message. Black or African American patients were less likely to be interested rather than not having received the message in comparison to White patients. Asian and Hispanic patients were more likely to not respond rather than not having received the message in comparison to White patients. Patients with a primary care provider were more likely to be interested rather than not having received the message in comparison to patients without a primary care provider.

Next, the preferred communication platform for scheduling an appointment among patients interested in the vaccine was assessed (Table 3). Patients living in priority zip codes were more likely to prefer a phone call (OR 1.43, 95% CI 1.15-1.78) than those living in a nonpriority zip code. These results remained the same with the addition of other independent variables in the models. Preference for a phone call was more likely among patients aged 70-74 years (OR 1.33, 95% CI 1.08-1.64) and those aged 80-84 years (OR 1.82, 95% CI 1.17-2.83) than among those aged 65-69 years, and among Black or African American patients (OR 1.69, 95% CI 1.29-2.21) than among White patients. Regarding the interaction effects, patients living in priority zip codes and those aged 75-79 years were significantly less likely to prefer a phone call (OR 0.40, 95% CI 0.20-0.81) than those aged 65-69 years and living in a priority zip code.

Finally, scheduled visits among patients who expressed an interest in being vaccinated were examined (Table 4). Patients aged 70-74 years were more likely to schedule an appointment (OR 1.38, 95% CI 1.01-1.89) than those aged 65-69 years, and Black or African American patients were more likely to schedule an appointment (OR 2.90, 95% CI 1.72-4.91) than White patients. Interaction variables could not be examined in these models given the small sample size.

Table 2. Odds ratios (ORs) of text status (N=30,826).

Text status	Model 1		Model 2		Model 3	
	OR	95% CI	OR	95% CI	OR	95% CI
Living in a priority zip code						
Interested	0.83[a]	0.75-0.91	0.81[a]	0.73-0.89	0.90	0.80-01.01
No response	1.19[a]	1.12-1.27	1.17[a]	1.09-1.25	1.08	1.00-1.17
Not interested	0.48[a]	0.44-0.53	0.48[a]	0.43-0.52	0.97	0.86-1.09
Not delivered	Reference	Reference	Reference	Reference	Reference	Reference
Age groups (years)						
65-69	—[b]	—	Reference	Reference	Reference	Reference
70-74: interested	—	—	0.53[a]	0.48-0.59	0.53[a]	0.48-0.59
70-74: no response	—	—	0.51[a]	0.47-0.55	0.52[a]	0.48-0.56
70-74: not interested	—	—	0.66[a]	0.60-0.74	0.64[a]	0.57-0.72
70-74: not delivered	—	—	Reference	Reference	Reference	Reference
75-79: interested	—	—	0.25[a]	0.22-0.29	0.24[a]	0.21-0.28
75-79: no response	—	—	0.30[a]	0.28-0.32	0.31[a]	0.28-0.33
75-79: not interested	—	—	0.36[a]	0.32-0.41	0.33[a]	0.29-0.38
75-79: not delivered	—	—	Reference	Reference	Reference	Reference
80-84: interested	—	—	0.12[a]	0.10-0.15	0.12[a]	0.10-0.15
80-84: no response	—	—	0.20[a]	0.18-0.22	0.21[a]	0.19-0.23
80-84: not interested	—	—	0.16[a]	0.14-0.20	0.16[a]	0.13-0.19
80-84: not delivered	—	—	Reference	Reference	Reference	Reference
85-89: interested	—	—	0.06[a]	0.05-0.08	0.06[a]	0.05-0.08
85-89: no response	—	—	0.10[a]	0.09-0.12	0.11[a]	0.10-0.12
85-89: not interested	—	—	0.07[a]	0.06-0.10	0.08[a]	0.06-0.10
85-89: not delivered	—	—	Reference	Reference	Reference	Reference
>90: interested	—	—	0.03[a]	0.02-0.04	0.03[a]	0.02-0.05
>90: no response	—	—	0.07[a]	0.06-0.08	0.08[a]	0.07-0.09
>90: not interested	—	—	0.04[a]	0.03-0.06	0.04[a]	0.03-0.06
>90: not delivered	—	—	Reference	Reference	Reference	Reference
Sex						
Interested	—	—	—	—	1.48[a]	1.35-1.61
No response	—	—	—	—	2.11[a]	2.00-2.23
Not interested	—	—	—	—	1.36[a]	1.24-1.49
Not delivered	—	—	—	—	Reference	Reference
Race and ethnicity						
White	—	—	—	—	Reference	Reference
Black or African American: interested	—	—	—	—	0.80[a]	0.71-0.90
Black or African American: no response	—	—	—	—	1.31[a]	1.21-1.42

Text status	Model 1		Model 2		Model 3	
	OR	95% CI	OR	95% CI	OR	95% CI
Black or African American: not interested	—	—	—	—	0.27[a]	0.24-0.31
Black or African American: not delivered	—	—	—	—	Reference	Reference
Asian: interested	—	—	—	—	1.58	0.96-2.61
Asian: no response	—	—	—	—	2.32[a]	1.64-3.38
Asian: not interested	—	—	—	—	0.85	0.50-1.43
Asian: not delivered	—	—	—	—	Reference	Reference
Hispanic or Latino: interested	—	—	—	—	1.10	0.75-1.62
Hispanic or Latino: no response	—	—	—	—	4.56[a]	3.70-5.61
Hispanic or Latino: not interested	—	—	—	—	0.59[a]	0.39-0.88
Hispanic or Latino: not delivered	—	—	—	—	Reference	Reference
Other race: interested	—	—	—	—	1.14	0.97-1.34
Other race: no response	—	—	—	—	1.75[a]	1.57-1.95
Other race: not interested	—	—	—	—	0.59[a]	0.51-0.70
Other race: not delivered	—	—	—	—	Reference	Reference
Total encounters: interested	—	—	—	—	1.00	1.00-1.01
Total encounters: no response	—	—	—	—	1.00	1.00-1.00
Total encounters: no response	—	—	—	—	1.00	1.00-1.00
Total encounters: not delivered	—	—	—	—	Reference	Reference
primary care provider: interested	—	—	—	—	1.17[a]	1.04-1.32
primary care provider: no response	—	—	—	—	0.75[a]	0.70-0.81
primary care provider: not interested	—	—	—	—	1.31[a]	1.16-1.49
primary care provider: not delivered	—	—	—	—	Reference	Reference

[a]Significant at $P<.05$.

[b]Variable not used in model.

Table 3. Odds ratios (ORs) for the preferred communication platform: phone or email (n=2011).

	Model 1		Model 2		Model 3		Model 4		Model 5	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Living in a priority zip code	1.43[a]	1.15-1.78	1.45[a]	1.16-1.81	1.06	0.82-1.38	—[b]	—	—	—
Age groups (years)										
65-69	—	—	Reference	Reference	Reference	Reference	—	—	—	—
70-74	—	—	1.33[a]	1.08-1.64	1.32[a]	1.07-1.63	—	—	—	—
75-79	—	—	1.24	.93-1.64	1.26	.95-1.69	—	—	—	—
80-84	—	—	1.82[a]	1.17-2.83	1.95[a]	1.25-3.04	—	—	—	—
85-89	—	—	0.84	0.45-1.56	0.89	0.48-1.66	—	—	—	—
>90	—	—	1.29	0.64-2.62	1.28	0.63-2.61	—	—	—	—
Sex	—	—	—	—	0.86	0.71-1.04	—	—	—	—
Race and ethnicity										
White	—	—	—	—	Reference	Reference	—	—	—	—
Black or African American	—	—	—	—	1.69[a]	1.29-2.21	—	—	—	—
Asian	—	—	—	—	0.67	0.27-1.68	—	—	—	—
Hispanic or Latino	—	—	—	—	0.94	0.41-2.15	—	—	—	—
Other race	—	—	—	—	1.30	0.91-1.85	—	—	—	—
Total encounters	—	—	—	—	1.00	0.99-1.00	—	—	—	—
Has a primary care provider	—	—	—	—	1.29	0.99-1.70	—	—	—	—
Interaction variables										
Living in a priority zip code and age group 65-69 years	—	—	—	—	—	—	Reference	Reference	—	—
Living in a priority zip code and age group 70-74 years	—	—	—	—	—	—	0.69	0.41-1.15	—	—
Living in a priority zip code and age group 75-79 years	—	—	—	—	—	—	0.40[a]	0.20-0.81	—	—
Living in a priority zip code and age group 80-84 years	—	—	—	—	—	—	0.92	0.34-2.50	—	—
Living in a priority zip code and age group 85-89 years	—	—	—	—	—	—	4.6	0.49-43.93	—	—
Living in a priority zip code and age group >90 years	—	—	—	—	—	—	1.58	0.22-11.39	—	—
Living in a priority zip code and White	—	—	—	—	—	—	—	—	Reference	Reference
Living in a priority zip code and Asian	—	—	—	—	—	—	—	—	0.97	0.15-6.28
Living in a priority zip code and Black or African American	—	—	—	—	—	—	—	—	1.00	0.55-1.82
Living in a priority zip code and Hispanic	—	—	—	—	—	—	—	—	0.85	0.09-7.52
Living in a priority zip code and other race	—	—	—	—	—	—	—	—	1.03	0.50-2.09

[a]Significant at $P<.05$.

[b]Variable not used in model.

Table 4. Odds ratios (ORs) of scheduled visits (n=2011).

	Model 1		Model 2		Model 3	
	OR	95% CI	OR	95% CI	OR	95% CI
Living in a priority zip code	1.30	0.90-1.87	1.29	0.89-1.87	0.72	0.46-1.11
Age group (years)						
65-69	—[a]	—	Reference	Reference	Reference	Reference
70-74	—	—	1.38[b]	1.01-1.89	1.36	0.99-1.86
75-79	—	—	1.15	0.74-1.78	1.20	0.77-1.88
80-84	—	—	0.45	0.18-1.15	0.50	0.19-1.28
85-89	—	—	0.42	0.10-1.78	0.46	0.11-1.98
>90	—	—	2.25	0.95-5.31	2.36	0.98-5.67
Sex	—	—	—	—	1.01	0.75-1.36
Race and ethnicity						
White	—	—	—	—	Reference	Reference
Black or African American	—	—	—	—	2.90[b]	1.72-4.91
Asian	—	—	—	—	0.91	0.11-7.20
Hispanic or Latino	—	—	—	—	0.78	0.10-6.20
Other race	—	—	—	—	1.86	0.93-3.70
Total encounters	—	—	—	—	1.00	1.00-1.00
Has a primary care provider	—	—	—	—	1.65	0.97-2.83

[a]Variable not used in model.

[b]Significant at $P<.05$.

Discussion

Principal Findings

Our findings highlight important considerations for large health systems attempting to use text messaging for effective and efficient vaccination outreach for diverse patient populations. Many text messages were not delivered (12,978/30,826, 42.1%) and of the messages that were delivered, 40% received no response. A small percentage of patients expressed interested in the vaccine (2938/30,826, 9.5%). Among patients who expressed interest in being vaccinated, it was found that patients largely preferred a phone call over email for scheduling their vaccine appointment. Of the patients contacted to schedule an appointment, 70-74–year-old patients were more likely to schedule an appointment than 65-69–year-old patients, and Black or African American patients were more likely to schedule an appointment than White patients.

These results reflect several key challenges. The first challenge was the verification of patient information and ascertaining whether a listed number was a landline or mobile number. Given the large number of text messages not delivered (n=12,978), a mechanism to validate these numbers is needed. In addition, a mechanism to update and clearly designate a number as a mobile number in the electronic health record is critical. Second, based on the selected outreach preference (phone) by those who were successfully contacted, it is critical to leverage nontext outreach methods for effective engagement of some populations. This finding may reflect comfort levels with technology among patients aged 65 years and older related to using a phone or computer to schedule an appointment. As patients had opportunities to be vaccinated elsewhere, it would be imprudent to draw any conclusions about why certain patient groups were more likely to schedule an appointment than others.

Comparison With Other Studies

Results from another study that examined a COVID-19 vaccination text message outreach to older patients found that reminder messages and messages that instilled ownership in patients led to increased scheduling of appointments and vaccination rates [10]. The findings from this study reflect how messaging protocols and how messages are written influence patient responsiveness to text message outreach. A study with older patient populations in Italy found that vaccine passports influenced patient receptiveness to being vaccinated, as the vaccination card allowed them to access to public spaces [11]. Policies around vaccination status may have shaped patient decision-making regarding vaccine uptake in the United States as well.

Implications of Study Findings

There are important lessons learned from this vaccination effort, which most health care systems undertook with no prior knowledge of how to execute this monumental task. The first one is the importance of accurate and updated patient contact information, particularly as it impacts effective outreach modalities. The second one is identifying patient preferences

for communication to increase the likelihood of engagement. The third one is the need for interoperability of patient health records to triangulate patient touch points with other health systems and services. This ensures patient health records are up to date, and finite outreach resources are focused on priority groups. The final lesson is establishing strong relationships with neighboring health systems and government agencies for the purpose of coordinating outreach efforts and sharing pertinent patient data with each other. The COVID-19 vaccination campaign is still underway in the United States, and health systems need to be nimble and flexible in reaching out to patients, including nondigital efforts such as provider-patient communication, and partnering with community-based organizations to reach vulnerable patients.

Limitations

There are several strengths and limitations to this study. A strength is the use of real-time patient data, which allowed capture and analysis of text outreach data linked to patient health records. This study provides critical data on digital outreach efforts as part of the COVID-19 vaccination campaign, which has important implications for local, regional, state, and national vaccination efforts. However, there are some limitations. First, 14,419 patients were excluded owing to shared phone numbers. The excluded patients had a similar demographic background as the study sample (Table 1). Second, it was not possible to verify if the patients received a COVID-19 vaccine outside of the study site.

Conclusions

The United States has made significant strides in vaccinating Americans aged 65 years and older through a combination of digital outreach efforts, community-based vaccine clinics, and at-home visits. This study highlights the benefits and challenges of using text messaging outreach methods, specifically the ability to reach large numbers of patients quickly. While the majority of Americans aged 65 years and older have been fully vaccinated (87.1%-92.7%), younger Americans aged 18-39 years are lagging behind in vaccination rates [12,13]. Emerging data have shown disparities by geographic location with lower vaccination rates in rural areas and differences by patient demographics (eg, education level and political beliefs) [14,15]. Guidelines regarding booster vaccines to protect against new variants will require expanded and ongoing outreach efforts to eligible patients. Reaching these populations will require multiple outreach methods (eg, texting, email, phone calls, and community vaccine clinics) designed to address vaccine hesitancy and access issues. These last 2 points necessitate partnering with trusted figures in the community to encourage vaccination (eg, faith-based and community leaders), providers respectfully and empathetically answering patient questions about the vaccine, health care systems and medical institutions building relationships with community members, and making it easy and simple to get vaccinated through expanded vaccination locations with flexible operating hours [4]. Furthermore, sharing of patients' vaccination statuses among health systems is integral to ensuring that outreach efforts are focused on unvaccinated individuals and individuals eligible for a booster. Coordination among health care systems, partnerships with and input from community leaders and members, and persistence are key elements to increasing COVID-19 vaccine uptake.

Acknowledgments

NA and DBW were previously affiliated with MedStar Center for Health Equity Research at the time of the study. They are currently affiliated with NYU Grossman School of Medicine and with Mathematica, respectively.

Conflicts of Interest

None declared

Multimedia Appendix 1
Sample demographics and interaction with text outreach (N=30,826).
[DOCX File , 20 KB - formative_v6i7e33260_app1.docx]

References

1. Forman R, Shah S, Jeurissen P, Jit M, Mossialos E. COVID-19 vaccine challenges: What have we learned so far and what remains to be done? Health Policy 2021 May;125(5):553-567 [FREE Full text] [doi: 10.1016/j.healthpol.2021.03.013] [Medline: 33820678]
2. Traynor K. COVID-19 vaccination campaigns take shape. Am J Health Syst Pharm 2021 Feb 08;78(4):282-284 [FREE Full text] [doi: 10.1093/ajhp/zxab019] [Medline: 33555342]
3. Daly M, Jones A, Robinson E. Public Trust and Willingness to Vaccinate Against COVID-19 in the US From October 14, 2020, to March 29, 2021. JAMA 2021 Jun 15;325(23):2397-2399 [FREE Full text] [doi: 10.1001/jama.2021.8246] [Medline: 34028495]
4. Quinn SC, Andrasik MP. Addressing Vaccine Hesitancy in BIPOC Communities — Toward Trustworthiness, Partnership, and Reciprocity. N Engl J Med 2021 Jul 08;385(2):97-100. [doi: 10.1056/nejmp2103104]
5. Tewarson H, Greene K, Fraser MR. State Strategies for Addressing Barriers During the Early US COVID-19 Vaccination Campaign. Am J Public Health 2021 Jun;111(6):1073-1077. [doi: 10.2105/ajph.2021.306241]

6. Schaffer DeRoo S, Pudalov NJ, Fu LY. Planning for a COVID-19 Vaccination Program. JAMA 2020 Jun 23;323(24):2458-2459. [doi: 10.1001/jama.2020.8711] [Medline: 32421155]
7. Prescott GM, Prescott WA. Health information technology utilization and impact on COVID-19 vaccination. J Am Pharm Assoc (2003) 2021 Jul;61(4):e230-e232 [FREE Full text] [doi: 10.1016/j.japh.2021.03.020] [Medline: 33858807]
8. Coronavirus (COVID-19) Vaccine. Government of the District of Columbia. URL: https://coronavirus.dc.gov/vaccine [accessed 2021-06-03]
9. Gold JA, Rossen LM, Ahmad FB, Sutton P, Li Z, Salvatore PP, et al. Race, Ethnicity, and Age Trends in Persons Who Died from COVID-19 - United States, May-August 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 23;69(42):1517-1521 [FREE Full text] [doi: 10.15585/mmwr.mm6942e1] [Medline: 33090984]
10. Dai H, Saccardo S, Han MA, Roh L, Raja N, Vangala S, et al. Behavioural nudges increase COVID-19 vaccinations. Nature 2021 Sep 02;597(7876):404-409 [FREE Full text] [doi: 10.1038/s41586-021-03843-2] [Medline: 34340242]
11. Gallè F, Sabella EA, Roma P, Da Molin G, Diella G, Montagna MT, et al. Acceptance of COVID-19 Vaccination in the Elderly: A Cross-Sectional Study in Southern Italy. Vaccines (Basel) 2021 Oct 21;9(11):1222 [FREE Full text] [doi: 10.3390/vaccines9111222] [Medline: 34835152]
12. Peterson M, Benhe F, Denget B, Nowotny K, Brinkley-Rubinstein L. Uneven Rollout Of COVID-19 Vaccinations In United States Prisons. Health Affairs. 2021 Apr 15. URL: https://www.healthaffairs.org/do/10.1377/forefront.20210413.559579/full/ [accessed 2021-10-12]
13. COVID-19 Vaccination and Case Trends by Age Group, United States. Centers for Disease Control and Prevention. URL: https://data.cdc.gov/Vaccinations/COVID-19-Vaccination-and-Case-Trends-by-Age-Group-/gxj9-t96f [accessed 2022-06-21]
14. Saelee R, Zell E, Murthy BP, Castro-Roman P, Fast H, Meng L, et al. Disparities in COVID-19 Vaccination Coverage Between Urban and Rural Counties - United States, December 14, 2020-January 31, 2022. MMWR Morb Mortal Wkly Rep 2022 Mar 04;71(9):335-340 [FREE Full text] [doi: 10.15585/mmwr.mm7109a2] [Medline: 35239636]
15. Agarwal R, Dugas M, Ramaprasad J, Luo J, Li G, Gao G. Socioeconomic privilege and political ideology are associated with racial disparity in COVID-19 vaccination. Proc Natl Acad Sci U S A 2021 Aug 17;118(33) [FREE Full text] [doi: 10.1073/pnas.2107873118] [Medline: 34326130]

Abbreviations

OR: odds ratio

Edited by A Mavragani; submitted 30.08.21; peer-reviewed by M Alfelali, S Badawy, D Kerr; comments to author 10.12.21; revised version received 06.01.22; accepted 16.06.22; published 18.07.22.

Please cite as:
Ahmed N, Boxley C, Dixit R, Krevat S, Fong A, Ratwani RM, Wesley DB
Evaluation of a Text Message–Based COVID-19 Vaccine Outreach Program Among Older Patients: Cross-sectional Study
JMIR Form Res 2022;6(7):e33260
URL: https://formative.jmir.org/2022/7/e33260
doi:10.2196/33260
PMID:35724339

Original Paper

Exploring the Use of a Web-Based Menu Planning Tool in Childcare Services: Qualitative Cross-sectional Survey Study

Jessica V Kempler[1], BBSc, BNutrSc (Hons), MHumNutr; Penelope Love[1,2], BScDiet (Hons), GCHIED, PhD; Kristy A Bolton[1,2], BAS (Hons), PhD; Margaret Rozman[3], B.AS, MNutDiet; Alison C Spence[1,2], BNutrDiet (Hons), GCHIED, PhD

[1]School of Exercise and Nutrition Sciences, Deakin University, Geelong, Australia

[2]Institute for Physical Activity and Nutrition (IPAN), Deakin University, Geelong, Australia

[3]Healthy Eating Advisory Service, Nutrition Australia Victorian Division, Docklands, Australia

Corresponding Author:
Jessica V Kempler, BBSc, BNutrSc (Hons), MHumNutr
School of Exercise and Nutrition Sciences
Deakin University
75 Pigdons Road
Geelong, 3216
Australia
Phone: 61 3 9246 8651
Email: jkempler@deakin.edu.au

Abstract

Background: Early childhood is a critical period for supporting the development of healthy eating habits, which may affect lifelong health. Childcare services are important settings for promoting early childhood nutrition; however, food provision in childcare frequently does not align with dietary guidelines. Web-based menu planning tools are well suited to support healthy food provision in childcare, although little is known about their use. Research is needed to understand how web-based menu planning tools are used in the childcare setting and how they can effectively support healthy menu planning and food provision for children in childcare.

Objective: We aimed to explore the use of a web-based menu planning tool called *FoodChecker*, which is available to childcare services in Victoria, Australia. We also aimed to gain insights and perspectives from childcare staff involved in menu planning about their use of the tool to plan healthy menus and guide healthy food provision for children.

Methods: We conducted a qualitative descriptive study using a cross-sectional web-based survey completed by the staff involved in menu planning in childcare services. Thematic analysis was performed using NVivo software. Emergent themes were mapped against constructs of the Technology Acceptance Model regarding perceived usefulness, perceived ease of use, and external variables influencing perceptions and use.

Results: The participants included 30 cooks and 34 directors from 53 childcare services. Participants perceived the web-based menu planning tool as useful for supporting child nutrition and health, improving organizational processes, and aiding the menu planner role. Perceptions regarding ease of use were mixed. External variables influencing perceptions and use included awareness of the tool, perceived need, time, resources, organizational support, and the food budget. Participants made recommendations to improve the tool, particularly the need to integrate functionality to make it easier and faster to use or to include more links to resources to support healthy menu planning.

Conclusions: The web-based menu planning tool was perceived as useful for cooks and directors in childcare services. Areas for improvement were identified; for example, the need for integrated digital features to make the tool easier and faster to use. As the first qualitative study to explore childcare staff experiences with a web-based menu planning tool, these findings inform future research and development of such tools to aid scalable and sustainable support for healthier food provision in the childcare sector.

(JMIR Form Res 2022;6(7):e35553) doi:10.2196/35553

KEYWORDS

child care; preschool; early childhood services; child nutrition; menu planning; healthy eating; web-based systems; web-based tool; internet-based intervention; user experience

Introduction

Background

Early childhood (commonly defined as 0-5 years) is a critical period for supporting the development of healthy eating habits that may track into later life [1]. An unhealthy diet during childhood is associated with both undernutrition and overweight and obesity [2]. As such, an unhealthy diet is a risk factor for nutrient deficiencies, impaired growth and development, and adverse chronic disease outcomes that can influence lifelong health [2-4]. Internationally, health [5] and government [6,7] authorities have established dietary guidelines outlining the types and amounts of foods children and adults are recommended to eat to support good health. However, population surveys demonstrate that globally, compliance with dietary guidelines is low across age groups and that children's diets are suboptimal (eg, diets low in vegetables and high in energy-dense, nutrient-poor foods are prevalent) [8-11].

Setting-based health promotion, where health is created and lived by people within the settings of their everyday life [12], is widely advocated as evidence-based best practice [13]. Early childhood education and care settings have been identified in systematic reviews as opportune places to promote early childhood nutrition [14,15]. Long day care or center-based care (herein referred to as childcare) is the most common form of early childhood education and care setting in Australia, with almost 800,000 children attending for an average of 30.5 hours (approximately 3 days) per week [16]. Similarly high patterns of childcare attendance are observed across other high-income countries, for example, in European countries [17], the United States [18], and the United Kingdom [19], reflecting changes in family workforce patterns, including increased female participation and shared caring responsibilities [20].

In Australia, half of childcare services operate as private, for-profit organizations, while 35% are private and not-for-profit. The remainder are managed by state or local governments (11%) or nongovernment schools (4%) [21]. Families accessing childcare services are supported by a means-tested national government subsidy, whereby families with lower income are eligible for a greater subsidy amount [22]. However, this does not guarantee affordability for everyone, and access to childcare is unequally distributed across Australia, with regional, remote, and disadvantaged areas more likely to experience low provision or absence of childcare [23].

Childcare services commonly provide meals and snacks for attending children, contributing up to two-thirds of their daily food intake [24]. As such, they have an important opportunity to support early childhood nutrition, and fundamental to this is planning a healthy childcare menu [25]. Recognizing this, leading health [26-28] and childcare [29,30] authorities around the world have established recommendations for healthy menu planning and food provision in the childcare setting. For example, the World Health Organization Commission on Ending Childhood Obesity advocates mandatory childcare nutrition standards [26], and the Australian National Quality Standard [29] requires childcare services to ensure "healthy eating... [is] promoted and appropriate for each child" (Element 2.1.3). However, a broad, international evidence base indicates that childcare menus do not meet dietary guidelines and are suboptimal for both food [25,31-33] and nutrient [34,35] provision.

Several barriers to healthy food provision in childcare have been identified in the literature, including insufficient menu planning tools and support resources, lack of time, and limited nutrition and dietary guideline knowledge [36]. Although limited in number, small randomized controlled trials (RCTs) have shown that some intervention strategies, including menu auditing and feedback [37,38], provision of menu planning resources [37-39], and expert implementation support [37,39], can improve childcare food provision. However, implementation models have traditionally relied on in-person support [40] and ongoing resourcing [41], limiting the scalability and sustainability of intervention strategies to date.

Web-based menu planning tools are emerging as a novel strategy for improving childcare food provision [42,43]. Given that almost all childcare services have access to computers and the internet [44], web-based tools may provide a mechanism for delivering scalable and sustainable menu planning support across the childcare sector, including in geographically dispersed locations. Such tools can be embedded into existing web-based childcare management systems [43] and completed at a time, location, and pace convenient for end users, with modest financial and staff resourcing requirements compared with other mechanisms [15]. Using digitalized systems, web-based menu planning tools can integrate user-engagement features, such as automated calculations of food groups on menus, comparisons with dietary guidelines, provision of instant feedback reports, and direct links to easily accessible and relevant web-based support resources [42,43].

To our knowledge, only two published RCTs have considered the impact of childcare programs that incorporate web-based tools to support healthy menu planning and food provision: (1) a pilot RCT (n=31) of the *Go-NAPSACC* (Nutrition and Physical Activity Self-Assessment for Child Care) program in the United States [42] and (2) an Australian RCT (n=54) of the *feedAustralia* menu planning tool in the state of New South Wales [43]. Although both tools were shown to improve healthy food choices on childcare menus, neither resulted in significant increases in menu compliance with sector food provision guidelines [42,43]. Authors from both studies called for future research to identify factors that influence the implementation of web-based menu planning tools in the childcare setting and exploration of strategies to inform their widespread use across the sector [42,43].

The implementation and effectiveness of web-based health promotion tools in achieving public health impact is largely

determined by end-user engagement [45]. In the childcare setting, users of web-based menu planning tools are most likely to be staff members who plan, prepare, and provide food for children—namely childcare cooks and directors [25,46]. However, little is known about how these users engage with web-based menu planning tools to support healthy food provision. From the limited evidence base, user acceptance of such programs is reportedly high, although studies to date have considered feedback from directors only [42,43] or captured only quantitative data [43]. This indicates that existing evidence may not reflect the nuanced perspectives of all stakeholders, including cooks and directors, who are likely to use web-based menu planning programs in practice.

In the Australian state of Victoria, the Victorian government has invested in the development and implementation of a web-based menu planning tool called *FoodChecker* [47]. Delivered by Nutrition Australia Victorian Division (NAV), *FoodChecker* is freely available to all Victorian childcare services to support menu alignment with sector dietary recommendations [48]. *FoodChecker* has been used by a third of Victorian childcare services since its inception in 2017 (NAV Program Manager, personal communication, December 14, 2021). The flow of the *FoodChecker* website, including the home page, menu data input template, and a sample automated report of menu alignment with dietary guidelines, is shown in Figures 1-3 [47]. The rollout of *FoodChecker* provides an opportunity to explore the use of web-based menu planning tools for providing equitable, scalable, and sustainable menu planning support in the childcare sector.

Figure 1. Screenshot of the *FoodChecker* homepage showing available services [47].

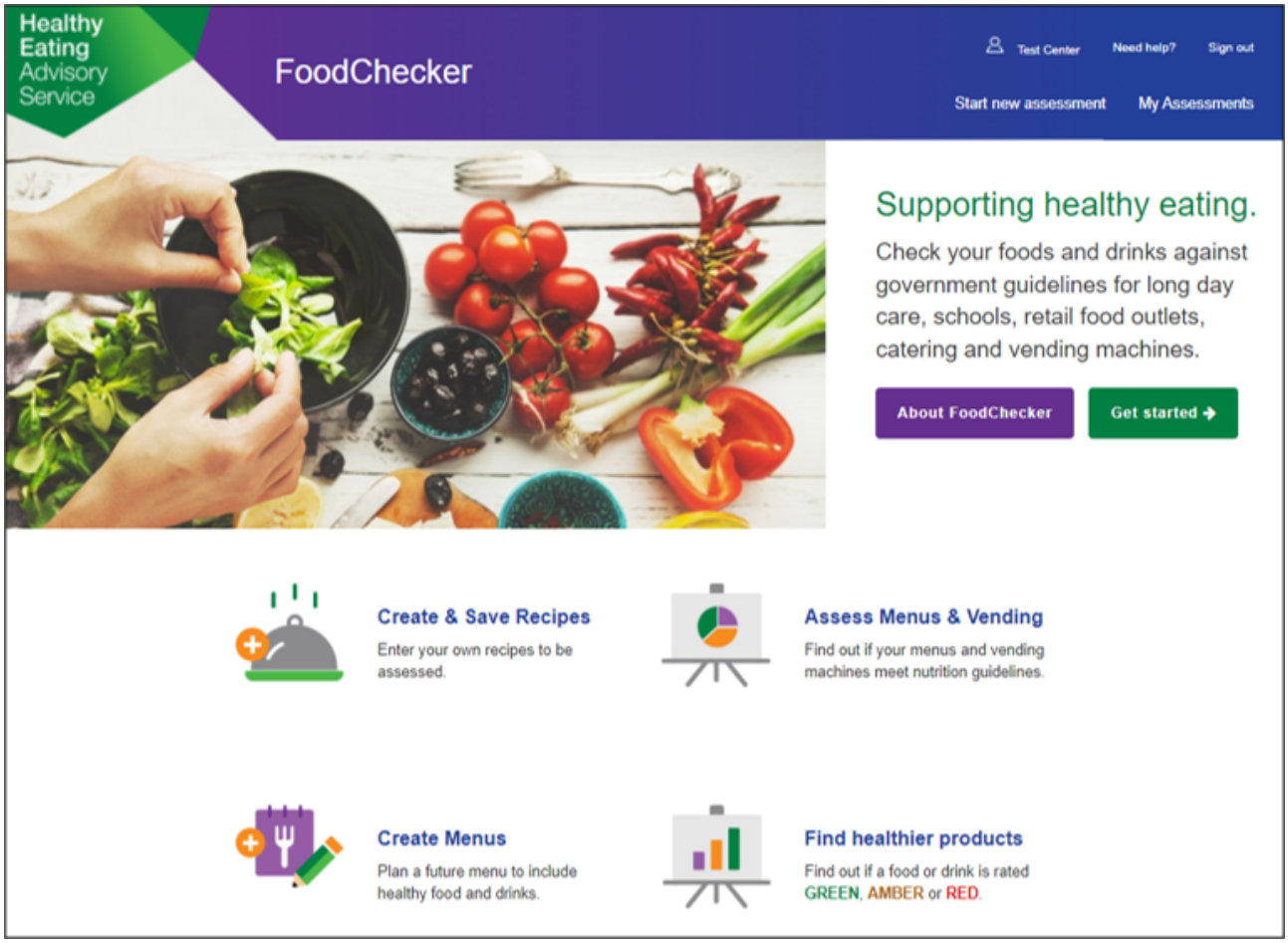


Figure 2. Screenshot of the *FoodChecker* menu data input template [47].

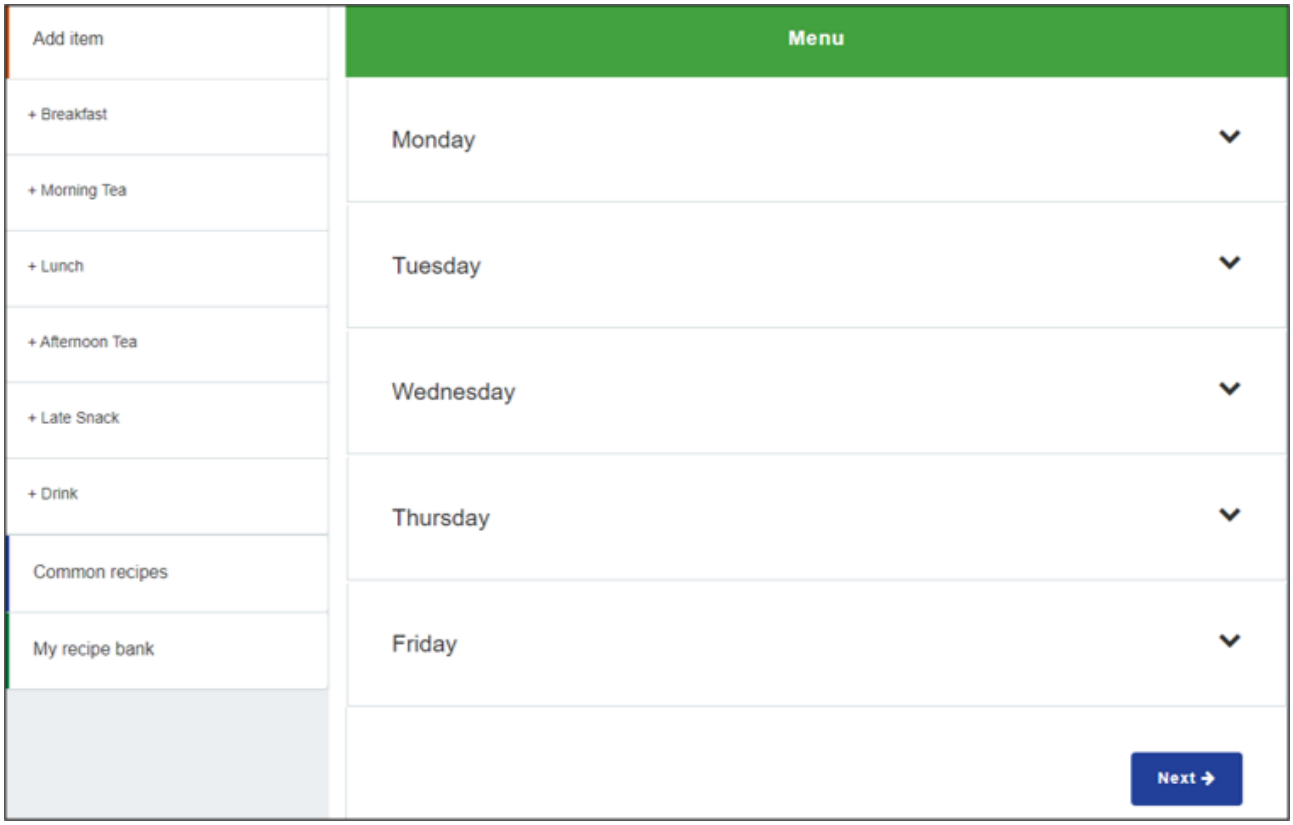
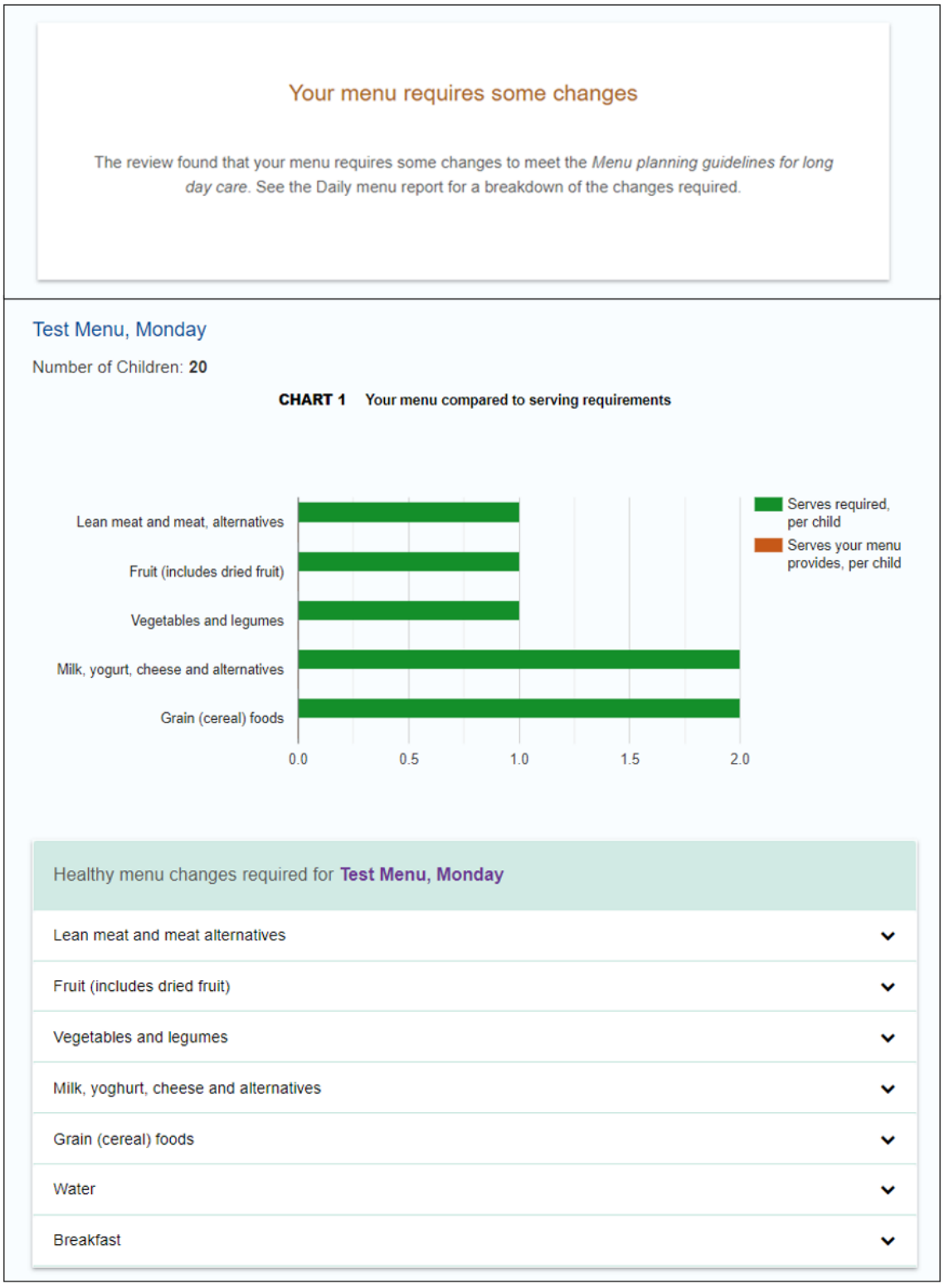


Figure 3. Screenshot of a sample automated report of menu alignment with dietary guidelines [47].



Objectives

Despite the potential of web-based menu planning tools to improve childcare food provision, to date, there are no published qualitative studies on the ways menu planning staff have engaged with these tools. The primary aim of this study was to explore the use of the web-based menu planning tool *FoodChecker* in Victorian childcare services. In particular, we aimed to gain insights and perspectives from menu planning staff members, in their own words, about their use of the web-based tool to plan healthy menus and guide healthy food provision for children in childcare.

Methods

Ethics Approval

This study was part of a broader research project on healthy eating and physical activity in childcare, with ethics approval

from the Deakin University Human Ethics Advisory Group (HEAG-H91_2021). All participants provided voluntary and informed consent to participate and received an Aus $20 (US $13.94) gift card in appreciation of their time.

Design and Setting

A qualitative descriptive study was conducted to explore participants' perspectives about their use of a web-based menu planning tool in the childcare setting. Although all research team members held nutrition qualifications, they sought to learn from the experience and expertise of the childcare staff. The researchers held an ontological position that embraced subjectivity, focusing on participants' personal experiences, insights, and opinions as opposed to seeking an *absolute truth*. The methods and results of this study were reported in accordance with the Standards for Reporting Qualitative Research checklist [49].

A cross-sectional survey of childcare staff members involved in planning childcare menus was conducted between July and September 2021 in Victoria, Australia. The survey explored the use of the web-based menu planning tool *FoodChecker*. Participant perceptions of *FoodChecker* were captured using a web-based qualitative survey, a data collection method recognized as beneficial for harnessing nuanced accounts of participant experiences within the qualitative descriptive paradigm [50].

Participants

Childcare services were identified from the Australian Children's Education and Care Quality Authority National Register [51] in July 2021. Eligible services were required to (1) be located in Victoria, Australia; (2) be open for at least 8 hours each weekday; (3) operate for at least 48 weeks annually; and (4) prepare and provide lunch, morning tea, and afternoon tea for attending children on each weekday. Services that did not provide food for children (eg, where meals were provided by parents) were ineligible because of differing meal planning requirements and because these represent a minority of childcare services in Victoria [52]. As childcare cooks and directors frequently share menu planning responsibilities [25,46], data were collected from both staff groups. A target sample size was not predetermined because of the inductive nature of the investigation and the desire to capture as broad a range of responses as possible from those with experiential expertise in childcare menu planning. Given that there are no published data on the proportion of childcare services in Victoria that provide food to children, the size of the target population was unknown. As such, data collection continued until no further responses were received.

Recruitment

An email invitation was sent to directors of all Victorian childcare services on the Australian Children's Education and Care Quality Authority National Register in July 2021 (N=1726) with a link to a voluntary self-administered *director survey* on the secure REDCap (Research Electronic Data Capture; Vanderbilt University) platform [53]. Directors providing consent responded to a screening question (within the *director survey*) regarding *FoodChecker* use (yes, no, or unsure). Directors who nominated that *FoodChecker* was or may have been used at their service were sent a link to a *FoodChecker survey*, which included individual consent. This could be forwarded to the cook responsible for planning the service's menu or completed by the director if they were involved in menu planning. One reminder email was sent to directors who did not respond to the initial recruitment email after 2 weeks. To maximize cooks' participation in the *FoodChecker survey*, the study was advertised to cooks in September 2021 via a post on a social media webpage commonly accessed by the target population.

Data Collection and Measures

Childcare Service and Participant Characteristics

Similar to previous research within the Australian childcare setting [43], participants reported their childcare service postcode and type of management (private or community), as well as their role in the service, years of employment, educational attainment, and whether they had received nutrition training. Participants also reported whether their service had ever used *FoodChecker* for menu planning (yes, no, or unsure).

FoodChecker Survey Design

As this is the first study, to the best of our knowledge, focusing on the qualitative exploration of a web-based childcare menu planning tool, a set of questions about *FoodChecker* use was purpose-designed by the research team. Previous international studies on user experiences with digital health tools have identified the need to capture information in the domains of user attitudes, experiences and expectations, and resultant changes in confidence, learning, and behavior [54,55]. To ensure that the study needs were addressed, additional domains were included to capture information about the frequency and purpose of use of *FoodChecker* and barriers and enablers influencing use (Table 1).

Topic-based qualitative questions were designed to be open and as succinct, clear, and unambiguous as possible, using the guidance for designing qualitative survey questions provided by Braun et al [50]. To contribute to internal generalizability and to support the interpretation of findings within the qualitative analysis [56], 5 quantitative questions using a nominal (yes or no) scale were added. Questions were then tested for face validity by 7 researchers (including JVK, ACS, PL, and KAB) with expertise in early childhood nutrition, 3 NAV staff members (including MR), 1 user experience design professional, and 1 previous childcare cook, with feedback incorporated into the final survey questions. Readability scores for the final set of *FoodChecker* questions (n=16 questions; Table 1) were 69.8 on the Flesch Reading Ease Test (desirable range 60-70) and 5.9 on the Flesch-Kincaid Grade Level Test, indicating that the content could likely be understood by a person approaching sixth grade in the United States [57].

The *FoodChecker survey* included the complete set of *FoodChecker* questions. Four of these questions were included in the *director survey*. This approach preempted the expectation that most participants responding to the *FoodChecker survey* would be cooks but that it was also important to seek insights from directors who often play a role in menu planning [25,46].

For both surveys, the number of questions included was within the range of 4 to 16, which is commonly observed in the literature for qualitative survey analyses focusing on lived experiences [50].

Table 1. Domains and questions in the *FoodChecker* question set[a].

Domain	Question
Frequency of use	• Q1. How often do you use *FoodChecker*?
Purpose of use	• Q2a. Please briefly state why you use or have used *FoodChecker*.[b] or • Q2b. Please explain why your center does not use *FoodChecker* for menu planning.[b]
User experiences	• Q3. What is the first thing that comes to mind about your experience with using *FoodChecker*?
User attitudes	• Q4. What do you like the most about using *FoodChecker*? • Q5. What do you like the least about using *FoodChecker*? • Q6. Do you think that online menu planning tools like *FoodChecker* are useful for your role? (yes/no) Please tell us why/why not. • Q7. Do you think that online menu planning tools like *FoodChecker* are useful for childcare centers? (yes/no) Please tell us why/why not.[b]
Enablers to use	• Q8. Have you accessed any support to help you use *FoodChecker*? • Q9. What organizational support do you receive (if any) to use *FoodChecker*?
Barriers to use	• Q10. What challenges do you face (if any) regarding the use of *FoodChecker*?
Changes in confidence, learning and behavior	• Q11. What do you think has changed for you or your center as a result of using *FoodChecker*?[b] • Q12. As a result of using *FoodChecker*, has your confidence about planning healthy menus improved? (yes/no) • Q13. As a result of using *FoodChecker*, have you learnt something? (yes/no) What have you learnt? • Q14. As a result of using *FoodChecker*, has your center's menu changed? (yes/no) What has changed?
User expectations	• Q15. If *FoodChecker* was being updated, is there anything that you would like to see in an "ideal" online menu planning tool?
Other	• Q16. Is there anything else that you think is important for us to know about *FoodChecker*?

[a]All questions were included in the *FoodChecker survey*.

[b]Questions included in *director survey* with the additional question, "Who has used *FoodChecker* in your center?"

Data Analysis

Survey response files were downloaded from REDCap, deidentified, and uploaded to NVivo (version 20; QSR International), a secure web-based data analysis platform [58].

Statistical Analysis

Postcode to remoteness area matching using the Australian Statistical Geography Standard [59] was used to classify each childcare service's geographic location as metropolitan or regional, based on proximity to a major city. Area-level socioeconomic position (SEP) for each childcare service was determined using the Index of Relative Socioeconomic Advantage and Disadvantage [60]. Each service was allocated a decile score based on its postcode to determine the relative level of advantage and disadvantage (1=greatest disadvantage and 10=greatest advantage) in the local area. Descriptive statistics were generated for the following: (1) childcare service characteristics including location (metropolitan or regional), type of management (private or community), area-level low SEP (score of 1-3), middle SEP (score of 4-7), or high SEP (score of 8-10) and *FoodChecker* use (yes, no, or unsure); (2) participant characteristics including role (director or cook), years of employment, educational attainment, and nutrition training; and (3) nominal (yes or no) data about *FoodChecker* usefulness and changes in confidence, learning, and the childcare menu.

Thematic Analysis

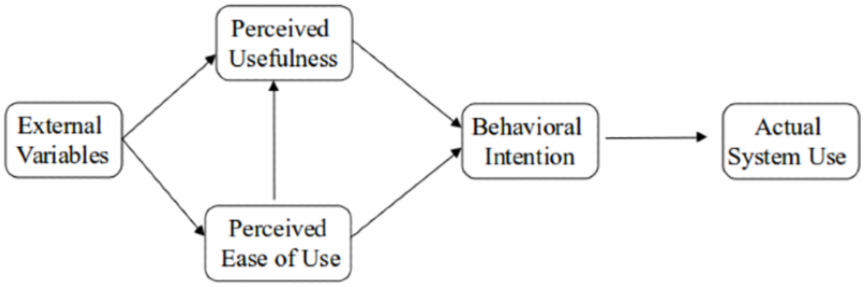
The Braun and Clarke [61] approach to inductive thematic analysis was used to explore menu planner perceptions of *FoodChecker*, as is consistent with the qualitative descriptive methodology [62]. Through an iterative process, an open coding technique was used to assign previously undefined codes to raw data extracts using NVivo. To minimize the risk of bias, a 10% sample of survey response files (n=6) was independently analyzed by 2 researchers (JVK and ACS), each of whom developed a preliminary coding framework. Differences in the coding frameworks were discussed until a consensus was reached. This verification process has been previously used in inductive thematic analysis of qualitative descriptive research [63]. Data extracted from the remaining survey response files were coded by 1 researcher (JVK). The codes were systematically categorized to determine common themes and their trends, patterns, and relationships. Through ongoing iteration and analysis, themes were reviewed and discussed with the research team, refined and named, and then rechecked to

ensure that they accurately reflected coded extracts and raw data.

Application of Theory

Constructs of the Technology Acceptance Model (TAM) were used to report themes identified from the inductive analysis related to the degree to which participants perceived *FoodChecker* would be useful and easy to use, and external variables specific to the individual or organization influencing perceptions and use (Figure 4 [64]). The TAM is a validated, widely used, and highly predictive model of information technology use [65], which posits that acceptance of a technology is directed by the degree to which users perceive the system to be useful (or enhance their job performance) and easy to use (or free from effort) [66]. It has been significantly associated with the intention to use a digital menu planning tool in the childcare sector [44].

Figure 4. Technology Acceptance Model (Venkatesh and Davis [64]).

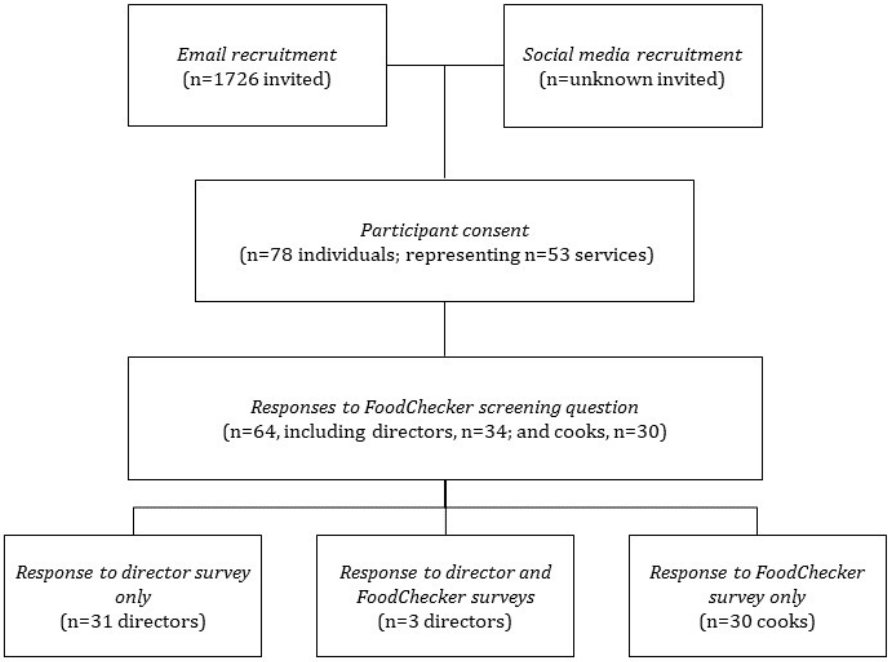


Results

Participants

A total of 64 participants (comprising 34 directors and 30 cooks) from 53 childcare services participated in this study (Figure 5); 52% (33/64; n=30 cooks and 3 directors) responded to the *FoodChecker survey*. The remaining participants (31/64, 48% directors) responded to the *director survey* only.

Figure 5. Childcare service and participant recruitment and survey respondents.



Childcare Service and Participant Characteristics

Childcare service and participant characteristics are presented in Table 2. Most services were located in metropolitan areas of Victoria (47/53, 89%) and privately owned (40/53, 75%). Half of the services (n=27) were located in regions classified as high SEP. Furthermore, 60% (32/53) of services reported using *FoodChecker* for menu planning. Most commonly, *FoodChecker* was used monthly or when menus were updated. Most participants had a minimum of certificate- or diploma-level qualification (55/64, 86%). Fewer than half (26/64, 41%) of the participants reported receiving nutrition training.

Table 2. Childcare service and participant characteristics.

Characteristics	Values, n (%)
Childcare service (n=53)	
Geographic location	
Metropolitan	47 (89)
Regional	6 (11)
Management	
Private	40 (75)
Community	13 (25)
Area-level SEP[a]	
Low SEP (scores 1-3)	10 (19)
Middle SEP (scores 4-7)	16 (30)
High SEP (scores 8-10)	27 (51)
***FoodChecker* use**	
No	21 (40)
Yes	32 (60)
Frequency of *FoodChecker* use (n=32)	
Once-off	3 (9)
Monthly	8 (25)
Every 3 months	4 (13)
Every 6 months	1 (3)
Whenever I update my menu	8 (25)
Other	1 (3)
No response	7 (22)
Participant (n=64)	
Role	
Director	34 (53)
Cook	30 (47)
Years of employment	
<1 year	7 (11)
1-2 years	6 (9)
2-3 years	10 (16)
3-4 years	3 (5)
>4 years	38 (59)
Educational attainment	
≤Grade 12	3 (5)
Trade, apprenticeship, diploma, or certificate	34 (53)
University degree	21 (33)
No response	6 (9)
Nutrition training	
Yes	26 (41)
No	32 (50)
No response	6 (9)

[a]SEP: socioeconomic position.

FoodChecker Usefulness and Changes in Confidence, Learning, and Menus

Participant responses to quantitative questions about *FoodChecker* are presented in Table 3. A total of 79% (26/33) of participants responded to quantitative questions about *FoodChecker*. Of these, the majority agreed *FoodChecker* was useful for their role (22/24, 92%) or childcare services (23/25, 92%). Furthermore, most participants agreed that due to using *FoodChecker*, their confidence about planning healthy menus had improved (20/26, 77%), they had learned something (17/24, 71%) and that their service's menu had changed (17/25, 68%).

Table 3. Participant reports of *FoodChecker* usefulness and changes in confidence, learning, and menus (n=33)[a].

Question	Response rate, n (%)	Yes, n (%)	No, n (%)
Do you think that online menu planning tools like *FoodChecker* are useful for childcare centers?	25 (76)	23 (92)	2 (8)
Do you think that online menu planning tools like *FoodChecker* are useful for your role?	24 (73)	22 (92)	2 (8)
As a result of using *FoodChecker*, has your confidence about planning healthy menus improved?	26 (79)	20 (77)	6 (23)
As a result of using *FoodChecker*, have you learnt something?	24 (73)	17 (71)	7 (29)
As a result of using *FoodChecker*, has your center's menu changed?	25 (76)	17 (68)	8 (32)

[a]Participant responses to the *FoodCheckersurvey* (30 cooks and 3 directors).

FoodChecker Themes

Overview

From the thematic analysis, 10 common themes were constructed about *FoodChecker* use, which were reported according to TAM constructs (Table 4).

Table 4. Overview of 10 themes constructed from thematic analysis and organized under constructs of the Technology Acceptance Model (TAM).

TAM construct	Theme

Perceived usefulness	• Theme 1: Supporting child nutrition and health
	• Theme 2. Improving organizational processes
	  • Quality improvement and accountability
	  • Meeting food provision recommendations and standards
	  • Improving menu planning processes
	  • Improving menu quality
	  • Engaging families
	• Theme 3. Aiding the menu planner
	  • Increasing confidence and learning
	  • Reducing workload
	• Theme 4. Ways to improve usefulness

Perceived ease of use	• Theme 5: Mixed perceptions about ease of use • Theme 6: Ways to improve ease of use
External variables	• Theme 7: Awareness and perceived need • Theme 8: Time and resources • Theme 9: Organizational support • Theme 10: Food budget

TAM Construct: Perceived Usefulness

Almost all participants, including directors and cooks, described *FoodChecker* as useful in their role or for the childcare sector, particularly for supporting child nutrition and health, improving organizational processes, and aiding the menu planner role.

Theme 1: Supporting Child Nutrition and Health

Child health was a priority for directors and cooks, who commonly discussed health motivations when describing the usefulness of *FoodChecker* in their role. This was demonstrated by one director who explained that the tool "changed the way we think about health" (director-34). Several cooks drew links between using *FoodChecker* and supporting children's health. One explained that it helped their service provide "nutritious food to reduce illness" (cook-6) and another specified a "healthy diet has direct benefits to students' mental health" (cook-13). Another drew a further link to children's learning, stating that the web-based tool was useful "especially for us: Strong Foundations! Healthy eating and healthy bodies = learning!" (cook-3). Participants acknowledged that food provided at their

service influenced children's health and discussed using *FoodChecker* to ensure that menus met children's nutrition and dietary needs. One cook stated, "it's a part of the program we use to ensure our children have a balanced, nutritional diet" (cook-8). A director emphasized this was particularly important given "many children receive most of [their] meals/snacks at the service" (director-25).

Theme 2: Improving Organizational Processes

The overview of this theme has been described as follows:

1. *Quality improvement and accountability*: quality improvement was important for participants, as evidenced by a director who explained "anything that assists us in continuous improvement has great value" (director-20). Cooks and directors alike described the need for childcare services and staff to be accountable for food provision practices. For example, a director described using *FoodChecker* to "ensure menu planning is on track and give accountability to the chef" (director-23), while a cook discussed using it "to keep services accountable for what they feed their children. There are still far to[o] many services with terrible budgets and menus that are not well balanced" (cook-24).
2. *Meeting food provision recommendations and standards*: *FoodChecker* was considered useful for supporting services to meet food provision recommendations. For example, a cook described using the web-based system to ensure "best-practice nutritional guidelines for children are being met in daily menus...and make sure we are meeting all healthy eating standards for children" (cook-3). Another referred to national childcare standards, stating that meeting children's daily nutrition needs is "required for rating and assessment" (cook-24). In some services, *FoodChecker* was used to meet food provision benchmarks within government-funded and endorsed health promotion initiatives. Indeed, one director stated, "As a part of the Achievement Program [67], we use *FoodChecker* to make sure our menus meet the daily intake of required foods and is healthy for all" (director-17).
3. *Improving menu planning processes*: several cooks expressed *FoodChecker* was valuable for improving menu planning processes. For example, they explained that *FoodChecker* "makes it easier and quicker to plan the menu" (cook-29) or "helps in the planning of meals, as it serves as a reference for appropriate quantities, number of serves, variety of food from the five food groups, and portion size... [it] also helps in wastage control, budgeting and ordering" (cook-5). *FoodChecker* was reported to offer further menu planning guidance through links to web-based resources including "healthy ingredient swaps and shopping tips" (cook-15) and ways "to deal with challenges such as allergies and budgeting" (cook-12).
4. *Improving menu quality*: most participants (17/25, 68%) who responded to the question about whether their childcare menu had changed because of using *FoodChecker* reported that menu changes had occurred owing to using the web-based program (Table 3). When describing these changes, some provided general information, explaining that they felt their menus were better, healthier, or more varied. For example, a director explained "we have a season[al] menu now where before it was a fortnightly menu which didn't change" (director-17). Others described specific changes, such as adapting portion sizes or food provision (eg, providing more dairy, vegetables or grains, or less fatty, salty, or sweet foods). One cook explained that they had rearranged the menu so it was now "designed and implemented based on the recommendations available on the system" (cook-5). Some respondents also thought that menu changes had resulted in dietary changes, as one cook stated, "we eat and enjoy more nutritional foods" (cook-8).
5. *Engaging families*: for some participants, *FoodChecker* provided a platform for accessing recipes, information, and guidelines to share with families and support their engagement in menu planning. Indeed, a director explained "children have more input and... parents are asking for a copy of recipes... parents are using the menus at home" (director-17).

Theme 3: Aiding the Menu Planner

The overview of this theme is described as follows:

1. *Increasing confidence and learning*: most participants who responded to questions about whether *FoodChecker* had impacted their confidence and learning reported that using the web-based tool had helped them build confidence in menu planning (20/26, 77%) and learn information relevant to their role (17/24, 71%; Table 3). Several cooks explained that from the web-based system they "learned how to cook" (cook-21) or "how to plan food for the kids" (cook-20). One stated that they "learned a lot of nutrition knowledge, which is very useful" (cook-15).
2. *Reducing workload*: several cooks reported *FoodChecker* made it "easier and quicker" (cook-8) to plan menus and that using the web-based tool reduced their workload. Some explained "it lightened a lot of work and made me more relaxed" (cook-14) or "it lightens my workload. I'm very satisfied that I can do other things" (cook-16).

Theme 4: Ways to Improve Usefulness

Although there was agreement about the usefulness of *FoodChecker*, the participants described updates that would further improve its value. For example, a cook recommended "Keep improving the tool... Provide us with accreditation evidence (a tick) so that families can see that our menus meet *FoodChecker* standards" (cook-3). Others suggested that the tool should provide information about food suppliers. Further recommendations were provided regarding additional resources that could be made available through the platform, such as sample menus, ingredient substitutions, nutrition information, allergy resources, and a greater variety of recipes. One cook noted the need for the digital system to ensure confidentiality, stating "I also need to know my recipes are private and won't be used in any way without my authorization" (cook-24).

TAM Construct: Perceived Ease of Use

Theme 5: Mixed Perceptions About Ease of Use

Perceptions about the ease of use of *FoodChecker* were mixed. Some participants reported that the tool was "easy to use" (director-23) or noted specific elements such as an "easy checklist to ensure a balanced weekly menu" (cook-3). Conversely, some described the web-based functionality as "not very easy to navigate" (director-13) or "a little too complicated in some ways" (cook-21). One cook stated it was "...too confusing, I needed help...computers are tricky for me... I don't like to use [it]" (cook-7).

Theme 6: Ways to Improve Ease of Use

Some participants stated that updates to navigation and functionality within the tool would make it easier to use. One cook, who described *FoodChecker* to be useful in their role, also stated "I just hope they work on the navigation of the site" (cook-17). Others discussed the need to integrate strategies to reduce data input time, explaining that "data input is quite time-consuming" (cook-5) or suggested the need for functionality to easily fix errors in data input, for example, "not having to start the entire menu over for small incorrect servings" (cook-4).

TAM Construct: External Variables

Participants identified several variables that influenced their perceptions of *FoodChecker* and their use of the tool.

Theme 7: Awareness and Perceived Need

Some participants had not used *FoodChecker* because they were unaware of the tool. One cook explained "I didn't know about it until now. I have registered and will look at it from now on" (cook-2). Others who did not use *FoodChecker* perceived their service as having adequate processes in place to ensure healthy menu planning. This included the presence of cooks and directors believed to be adequately skilled in healthy menu planning, as well as input from staff, parents, and children. One director explained "we are a small private center with a self-managed system in place that works well" (director-21). Others accessed support from external consultants or used "the alternative [menu planning tool] from *feedAustralia*" (director-5).

Theme 8: Time and Resources

Time was described as an important factor related to *FoodChecker* use. Participants commonly stated that inputting menu data into the program was "very time consuming" (director-13). The short turnaround times for planning new menu cycles and the need to reassess menus with each change presented challenges. For example, participants explained that they "normally allow two weeks to complete [a] new season menu" (director-17) and "each time we change a menu, we need to food check again" (cook-3).

Lack of time was a reported barrier to *FoodChecker* use, as a director explained, "the cook is aware of *FoodChecker* but is limited on time to use this service" (director-14). For one cook, lack of time was exacerbated by a lack of technological resources, as they explained, "there is no computer or iPad in the kitchen" (cook-24). They further described the challenge of competing priorities within their role, stating they:

> *Get menu planning time, [b]ut it is also documentation time, cleaning, and food safety plan time. As well as newsletters and posting on story park. So [I] need to prioritize the work and FoodChecker is sometimes last.* [Cook-24]

Others reported that dedicated "paid time for menu planning" (cook-4) within their role enabled them to use *FoodChecker*.

Theme 9: Organizational Support

Cooks reported that management support and leadership facilitated their use of *FoodChecker*. One cook expressed they received "encouragement, time and practical support from management to use *FoodChecker*" (cook-3) and another explained, "the director and teachers of our center are very satisfied and give us the greatest support" (cook-15). In some services, using the web-based system was perceived to be a directive from management. Cooks discussed using *FoodChecker* at the "request of the business to ensure that we meet (and exceed) the nutritional requirements of the children" (cook-5) or that it was "part of our policy...to use *FoodChecker*" (cook-29).

Theme 10: Food Budget

The food budget was an important factor for cooks, as reported by one participant who stated, "cost control is our biggest headache" (cook-15). While some reported *FoodChecker* provided links to web-based resources that supported their service with budgeting, others explained that recommendations made by the *FoodChecker* system presented challenges for the food budget. For example, they explained that when using *FoodChecker* they were "unable to control costs" (cook-12) or that it was "easy to exceed our budget and buy food materials" (cook-16).

Discussion

Principal Findings

In this novel study, we aimed to explore the use of a web-based menu planning tool for childcare services. Among the first of its kind, the study sought insights and perspectives from childcare cooks and directors, in their own words, about their use of a web-based tool to plan menus and guide food provision for children in childcare. The study found that cooks and directors alike considered the web-based tool to be useful in their roles, although use was influenced by a variety of factors including awareness, perceived need, time, resources, organizational support, and budgetary considerations. Participants made recommendations to improve the web-based tool, including the need to update navigation and functionality, integrate strategies to reduce data input time, and provide more links to relevant web-based resources to support healthy menu planning.

Comparison With Prior Work

To the best of our knowledge, this is the first study to consider insights from both cooks and directors about their use of a web-based childcare menu planning tool, with limited previous

analyses focusing on acceptance by childcare directors only [42,43]. As such, this study offers an end-user perspective most likely to represent insights from both staff groups involved in menu planning. Participants described a variety of motivations important for menu planning, particularly the need to support children's health and nutrition, improve food provision and menu planning processes, and aid the menu planner role. Similar motivations have been reported in previous analyses of childcare menu planners [68,69]. However, this is the first study to document the usefulness of a web-based menu planning tool for integrating these motivations into practice.

In this study, directors and cooks emphasized the value of a web-based menu planning tool for both the childcare sector and within their specific role. Comparatively, in limited prior evidence, reports on the perceived value of web-based menu planning tools in childcare settings have been mixed. For instance, while directors have previously reported high intentions to use a web-based menu planning tool and high levels of computer access [44], cooks have reported not using or requiring web-based menu planning tools and having limited computer literacy and access [68]. This discordance could indicate differing needs, levels of computer access, or perceived levels of computer literacy between cooks and directors in childcare settings. Indeed, in this study, one cook reported low levels of digital literacy, and another indicated that a lack of technology in the kitchen was a barrier to using a web-based menu planning tool. These challenges were not reported by the directors.

In this study, the web-based menu planning tool was considered valuable for supporting engagement with families, particularly for sharing menus and recipes. This is important, given that family engagement is widely recommended to increase the impact of childcare-based healthy eating interventions [70]. Most childcare services use web-based platforms to communicate with families [44], indicating the potential to extend the use and reach of web-based menu planning tools to the family and home environment. Indeed, this is demonstrated in the *feedAustralia* intervention where parents can use a mobile app to view daily food offerings and access sample menus and recipes that they can recreate at home [71].

Despite the consensus about the usefulness of the web-based menu planning tool in this study, there were mixed reports from both directors and cooks about how easy it was to use the tool in practice. It stands to reason that user training may improve ease of use, as reported in the *Go-NAPSACC* trial [42]. However, the participants in this study did not discuss the need for *FoodChecker* training but rather the need to update navigation and functionality within the web-based tool to make it easier to use. Given that up to 80% of health technologies have limited success owing to a lack of end-user adoption or sustained use [45], understanding and integrating user preferences within web-based systems is necessary to increase their use and impact [72]. Future developments of web-based childcare menu planning tools should therefore consider strategies to improve user engagement (such as simpler site navigation or functionality to facilitate faster data input), to amplify their adoption, for sustained use over time, and for public health impact. This indicates the importance of directing funding toward the ongoing development of web-based menu planning tools for the childcare sector to meet user expectations, particularly in the current era of rapid technological advancement.

The limited evidence available has demonstrated that, even when the acceptability of a web-based menu planning program in the childcare sector is high, use may still be variable [42,43]. This indicates that factors external to the system itself (eg, individual or organizational factors) may influence use. Exploration of such factors was novel to this study, with participants reporting that awareness, perceived need, time, management support, and the food budget were important variables relevant to their uptake of the web-based menu planning tool.

Several participants were unaware of the tool and, as such, had not used it. Others who were aware of the tool but did not use it perceived that childcare staff had adequate nutrition knowledge and skills to plan a healthy menu. Given the scope of this study did not include a menu assessment component, it was not possible to triangulate staff perceptions about their knowledge and skills in healthy menu planning and the degree to which childcare menus complied with food provision guidelines. This is an important direction for future research. Despite this, there were multiple instances in which both cooks and directors demonstrated good knowledge of the scientific evidence related to their roles. For example, participants discussed the need to provide optimum nutrition to children who received most of their meals and snacks in childcare or links between healthier eating and children's health outcomes, as evidenced in the literature [1,24]. Although these examples of accurate nutrition knowledge are reassuring, it cannot be inferred that this knowledge is consistent or accurately translated into menu planning practices, indicating the need for healthy menu planning support in the sector.

Although support in the form of nutrition training has been shown to improve menu quality, such training is not routine [25]. Indeed, less than half of the participants in this study reported having received nutrition training. This is similar to previous studies where childcare staff reported low levels of nutrition training despite their responsibility to plan a healthy menu that meets food provision guidelines [25]. This indicates that there is scope for the design and implementation of interventions, such as web-based menu planning tools that offer accessible, evidence-based guidance that can be integrated into menu planning practices, even in the absence of formal staff nutrition training.

Some participants who stated that they did not use the web-based menu planning tool reported using an external consultant to support healthy menu planning in their service. Such consultancy may increase the financial burden of healthy menu planning, and may not be viable for all childcare services. However, it may also present a sensible approach to ensure accuracy in menu compliance with dietary guideline recommendations, if such support is available at a feasible cost.

Time was an important factor in this study, with mixed perceptions among the participants. Staff members who received dedicated and paid time within their role to use the web-based menu planning tool reported that this facilitated its use.

Moreover, some participants who used the tool reported that it was time saving and reduced their workload. However, as in previous studies [73], most participants did not report receiving paid time for menu planning within their roles. Furthermore, the use of the tool was commonly perceived to be time consuming, even by participants who had never used it before. As such, perceptions about time were a deterrent to initial adoption. This indicates that there is scope to (1) integrate user-engagement features within childcare menu planning systems to improve time efficiency and (2) establish strategies to shift user perceptions about such systems from being *time consuming* to being *time saving*. These are important areas for future development, particularly given that childcare staff members are known to be a time-poor population group [36].

In addition to dedicated time, service management support and leadership were reported to facilitate the uptake of the web-based menu planning tool. Interestingly, one cook who did not want to use the program had used it on instruction from their manager to ensure menu compliance with dietary guidelines. This indicates that despite the presence of management support, menu planner resistance may exist and could be a potential barrier to ongoing use of the web-based menu planning tool.

The food budget and cost control were important considerations for participants who described budgeting resources linked to the web-based menu planning tool as useful. However, several cooks reported that specific food recommendations generated by the web-based system exceeded their allocated food budget. This is a novel finding, indicating the need for dietary guidelines and recommendations embedded within such systems to consider food budgets and financial constraints. It may also suggest the need to investigate the capacity of childcare food budgets to adequately provide optimal nutrition for children. There is scope for web-based childcare menu planning systems to integrate strategies to support services in establishing and managing food budgets, which is an area for future research and development.

Although the services in this study were located in areas of varying levels of advantage and disadvantage, half were located in regions classified by postcode as high SEP. This is consistent with data indicating that, in Victoria, there is a higher provision of childcare services in areas experiencing greater levels of advantage [23]. However, as the study achieved a modest sample size, it was not possible to determine whether area-level SEP was associated with the acceptance of the web-based menu planning tool or the degree to which the tool supported healthy food provision in childcare services. Given that area-level disadvantage is associated with poorer diet quality [74], and given that services in low SEP areas were least represented in this study, there is a need for future research to better understand the use of web-based childcare menu planning tools in lower SEP areas.

Strengths and Limitations

The strengths of this study warrant discussion. As this study captured perspectives from cooks and directors, the findings are likely to represent the perspectives of both staff groups involved in menu planning. The study's qualitative descriptive underpinning provided scope to capture participant-generated data on menu planner experiences in their own words. Insights generated from spontaneous reporting were likely to reflect motivations and perspectives that are most important to users when compared with quantitative analyses using predefined, researcher-generated items and response scales [62]. The use of the validated TAM contributed to the study's underlying theoretical foundation. In line with emerging methodology-focused evidence, a web-based qualitative survey facilitated data collection from a time-poor and varied, dispersed, and geographically heterogeneous population with the potential to reduce social desirability bias [50]. A further strength of this study was that it analyzed the use of a real-world web-based tool in current practice, to which real-time refinements are possible. As such, the exploration brought together research, reflection, and practical solutions as part of an action research approach.

Given that the *FoodChecker* tool is only freely available in Victoria, the scope of this study was purposely delimited to Victorian childcare services. A modest study sample was achieved, as observed in previous studies on food provision in childcare settings [68]. Furthermore, given that childcare staff are known to be a time-poor population group, the perspectives of those who have particularly limited time to engage with a web-based menu planning tool, and as such, to participate in this research, may not have been captured. This may impact the ability to generalize the findings to the broader childcare sector. However, as the analysis explored participant perceptions and insights that are subjective by nature, the findings may be indicative of childcare menu planner experiences more broadly. This should be further investigated in larger studies exploring the use of web-based menu planning tools to support healthy food provision in childcare. It was beyond the scope of this study to include a quantitative analysis of the impact of the web-based menu planning tool on menu compliance with dietary guidelines, food provision, or children's dietary intake. These are important areas for future research.

Conclusions

This novel qualitative descriptive study demonstrates the usefulness of a web-based tool to support healthy menu planning in childcare services. Use of the tool was impacted by its internal functionality as well as external organizational factors. Recommendations were made to improve the web-based menu planning system.

Further research is needed to better understand how web-based menu planning tools can improve food provision and children's consumption in the childcare setting. In particular, studies should investigate and evaluate strategies to improve user engagement with web-based menu planning tools in childcare to increase their adoption, use, and public health impact.

Authors' Contributions

JVK, ACS, PL, and KAB contributed to the conception and design of the study and data acquisition. JVK, ACS, and PL contributed to data analysis and interpretation, and JVK conducted the data analysis. PL led the acquisition of data. MR oversaw the development of the *FoodChecker* intervention. JVK drafted the manuscript. All authors took part in revising the manuscript, gave approval for this version to be published, and agreed to be accountable for all aspects of the work.

Conflicts of Interest

JVK is a previous employee of NAV, which delivers *FoodChecker*. JVK currently holds no contractual or financial relationship with NAV. MR is an employee of NAV. Individual members of the research team had no prior relationship with the participants. Participants may have interacted with NAV as an organization.

References

1. Craigie AM, Lake AA, Kelly SA, Adamson AJ, Mathers JC. Tracking of obesity-related behaviours from childhood to adulthood: A systematic review. Maturitas 2011 Nov;70(3):266-284. [doi: 10.1016/j.maturitas.2011.08.005] [Medline: 21920682]
2. The State of the World's Children 2019 - Children, food and nutrition: Growing well in a changing world. UNICEF. URL: https://www.unicef.org/reports/state-of-worlds-children-2019 [accessed 2022-05-03]
3. Shrestha R, Copenhaver M. Long-term effects of childhood risk factors on cardiovascular health during adulthood. Clin Med Rev Vasc Health 2015 Aug 12;7:1-5 [FREE Full text] [doi: 10.4137/CMRVH.S29964] [Medline: 26312015]
4. Silink M. Childhood diabetes: A global perspective. Horm Res 2002 Nov 17;57 Suppl 1(1):1-5. [doi: 10.1159/000053304] [Medline: 11979014]
5. Healthy diet. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/healthy-diet [accessed 2022-02-06]
6. Dietary Guidelines for Americans, 2020-2025, 9th Edition. U.S. Department of Agriculture and U.S. Department of Health and Human Services. 2020. URL: https://www.dietaryguidelines.gov [accessed 2021-11-11]
7. Australian Dietary Guidelines. National Health and Medical Research Council. 2013. URL: https://www.eatforhealth.gov.au/guidelines [accessed 2021-11-08]
8. 2011-12 Australian Health Survey: Nutrition first results - foods and nutrients. Australian Bureau of Statistics. URL: https://www.abs.gov.au/statistics/health/health-conditions-and-risks/australian-health-survey-nutrition-first-results-foods-and-nutrients/latest-release [accessed 2021-11-08]
9. National Diet and Nutrition Survey rolling programme: Results from years 9 to 11 (2016 to 2017 and 2018 to 2019). Public Health England. URL: https://www.gov.uk/government/statistics/ndns-results-from-years-9-to-11-2016-to-2017-and-2018-to-2019 [accessed 2021-11-12]
10. Liu J, Rehm CD, Onopa J, Mozaffarian D. Trends in diet quality among youth in the United States, 1999-2016. JAMA 2020 Mar 24;323(12):1161-1174 [FREE Full text] [doi: 10.1001/jama.2020.0878] [Medline: 32207798]
11. Williams J, Buoncristiano M, Nardone P, Rito AI, Spinelli A, Hejgaard T, et al. A snapshot of European children's eating habits: Results from the fourth round of the WHO European Childhood Obesity Surveillance Initiative (COSI). Nutrients 2020 Aug 17;12(8):2481 [FREE Full text] [doi: 10.3390/nu12082481] [Medline: 32824588]
12. The Ottawa Charter for Health Promotion. World Health Organization. 1986. URL: https://apps.who.int/iris/bitstream/handle/10665/59557/WHO_HPR_HEP_95.1.pdf;jsessionid=1CF18F1C35D715E750435B1FCB396795?sequence=1 [accessed 2022-02-06]
13. Creating health promoting settings. World Health Organization. URL: https://www.who.int/westernpacific/activities/creating-health-promoting-settings [accessed 2021-11-08]
14. Mikkelsen MV, Husby S, Skov LR, Perez-Cueto FJ. A systematic review of types of healthy eating interventions in preschools. Nutr J 2014 Jun 06;13(1):56 [FREE Full text] [doi: 10.1186/1475-2891-13-56] [Medline: 24906305]
15. Matwiejczyk L, Mehta K, Scott J, Tonkin E, Coveney J. Characteristics of effective interventions promoting healthy eating for pre-schoolers in childcare settings: An umbrella review. Nutrients 2018 Mar 01;10(3):293 [FREE Full text] [doi: 10.3390/nu10030293] [Medline: 29494537]
16. Child care in Australia report March quarter 2020. Australian Government Department of Education, Skills and Employment. URL: https://www.dese.gov.au/key-official-documents-about-early-childhood/early-childhood-and-child-care-reports/child-care-australia/child-care-australia-report-march-quarter-2020 [accessed 2021-11-08]
17. European Commission's Expert Group on Gender and Employment Issues (EGGE). The provision of childcare services: A comparative review of 30 European countries. Belgium: Publications Office of the European Union; 2009.
18. Who's minding the kids? Child care arrangements: Spring 2011. United States Census Bureau. URL: https://www.census.gov/library/publications/2013/demo/p70-135.html [accessed 2021-11-08]
19. Childcare and early years providers survey: 2019. UK Department for Education. URL: https://www.gov.uk/government/statistics/childcare-and-early-years-providers-survey-2019#history [accessed 2021-11-08]

20. Mehta K, Booth S, Coveney J, Strazdins L. Feeding the Australian family: Challenges for mothers, nutrition and equity. Health Promot Int 2020 Aug 01;35(4):771-778. [doi: 10.1093/heapro/daz061] [Medline: 31326984]
21. NQF Snapshot Q4 2021: A quarterly report from the Australian children's education and care quality authority. Australian Children's Education and Care Quality Authority. URL: https://www.acecqa.gov.au/nqf/snapshots [accessed 2022-05-03]
22. Child care subsidy. Australian Government Department of Education, Skills and Employment. URL: https://www.dese.gov.au/child-care-subsidy [accessed 2022-05-03]
23. Deserts and oases: How accessible is childcare in Australia? Victoria University. 2022. URL: https://www.vu.edu.au/mitchell-institute/early-learning/childcare-deserts-oases-how-accessible-is-childcare-in-australia?msclkid=e15dd1adcb3c11ec96d494f108fe4450 [accessed 2022-05-04]
24. Pollard C, Lewis J, Miller M. Food service in long day care centres - An opportunity for public health intervention. Aust N Z J Public Health 1999 Dec;23(6):606-610. [doi: 10.1111/j.1467-842x.1999.tb01545.x] [Medline: 10641351]
25. Grady A, Stacey F, Seward K, Finch M, Jones J, Yoong SL. Menu planning practices in early childhood education and care - factors associated with menu compliance with sector dietary guidelines. Health Promot J Austr 2020 Apr;31(2):216-223. [doi: 10.1002/hpja.286] [Medline: 31397031]
26. Report of the commission on ending childhood obesity. World Health Organization. URL: https://www.who.int/publications/i/item/9789241510066 [accessed 2021-11-08]
27. Buscemi J, Kanwischer K, Becker AB, Ward DS, Fitzgibbon ML, on behalf of the Society of Behavioral Medicine Health Policy Committee. Society of Behavioral Medicine position statement: Early care and education (ECE) policies can impact obesity prevention among preschool-aged children. Transl Behav Med 2015 Mar 29;5(1):122-125 [FREE Full text] [doi: 10.1007/s13142-014-0297-5] [Medline: 25729461]
28. Benjamin Neelon SE, Briley ME, on behalf of the American Dietetic Association. Position of the American Dietetic Association: Benchmarks for nutrition in child care. J Am Diet Assoc 2011 Apr;111(4):607-615. [doi: 10.1016/j.jada.2011.02.016] [Medline: 21443997]
29. Guide to the National Quality Framework. Australian Children's Education and Care Quality Authority. URL: https://www.acecqa.gov.au/nqf/about/guide [accessed 2021-11-08]
30. Child care licensing: Nutrition. Ontario Ministry of Education. URL: http://www.edu.gov.on.ca/earlyyears/nutrition.html [accessed 2022-06-25]
31. Benjamin Neelon SE, Burgoine T, Hesketh KR, Monsivais P. Nutrition practices of nurseries in England. Comparison with national guidelines. Appetite 2015 Feb;85:22-29 [FREE Full text] [doi: 10.1016/j.appet.2014.11.002] [Medline: 25450898]
32. Gerritsen S, Dean B, Morton SM, Wall CR. Do childcare menus meet nutrition guidelines? Quantity, variety and quality of food provided in New Zealand early childhood education services. Aust N Z J Public Health 2017 Aug 14;41(4):345-351. [doi: 10.1111/1753-6405.12667] [Medline: 28616873]
33. Wallace R, Costello L, Devine A. Over-provision of discretionary foods at childcare dilutes the nutritional quality of diets for children. Aust N Z J Public Health 2017 Aug 28;41(4):447. [doi: 10.1111/1753-6405.12658] [Medline: 28245524]
34. Myszkowska-Ryciak J, Harton A. Implementation of dietary reference intake standards in preschool menus in Poland. Nutrients 2018 May 10;10(5):592 [FREE Full text] [doi: 10.3390/nu10050592] [Medline: 29748511]
35. Frampton AM, Sisson SB, Horm D, Campbell JE, Lora K, Ladner JL. What's for lunch? An analysis of lunch menus in 83 urban and rural Oklahoma child-care centers providing all-day care to preschool children. J Acad Nutr Diet 2014 Sep;114(9):1367-1374. [doi: 10.1016/j.jand.2013.09.025] [Medline: 24332085]
36. Seward K, Finch M, Yoong SL, Wyse R, Jones J, Grady A, et al. Factors that influence the implementation of dietary guidelines regarding food provision in centre based childcare services: A systematic review. Prev Med 2017 Dec;105:197-205 [FREE Full text] [doi: 10.1016/j.ypmed.2017.09.024] [Medline: 28965755]
37. Seward K, Wolfenden L, Finch M, Wiggers J, Wyse R, Jones J, et al. Improving the implementation of nutrition guidelines in childcare centres improves child dietary intake: Findings of a randomised trial of an implementation intervention. Public Health Nutr 2017 Nov 27;21(03):607-617. [doi: 10.1017/s1368980017003366]
38. Finch M, Seward K, Wedesweiler T, Stacey F, Grady A, Jones J, et al. Challenges of increasing childcare center compliance with nutrition guidelines: A randomized controlled trial of an intervention providing training, written menu feedback, and printed resources. Am J Health Promot 2019 Mar 13;33(3):399-411. [doi: 10.1177/0890117118786859] [Medline: 30004247]
39. Leis A, Ward S, Vatanparast H, Humbert ML, Chow AF, Muhajarine N, et al. Effectiveness of the Healthy Start-Départ Santé approach on physical activity, healthy eating and fundamental movement skills of preschoolers attending childcare centres: A randomized controlled trial. BMC Public Health 2020 Apr 19;20(1):523 [FREE Full text] [doi: 10.1186/s12889-020-08621-9] [Medline: 32306943]
40. Knowledge into practice: NACCRRA's survey of child care resource and referral on-site technical assistance. National Association of Child Care Resource and Referral Agencies. 2007. URL: https://www.yumpu.com/en/document/read/52514523/knowledge-into-practice [accessed 2022-02-06]
41. Sari N, Muhajarine N, Froehlich Chow A. The Saskatchewan/New Brunswick Healthy Start-Départ Santé intervention: Implementation cost estimates of a physical activity and healthy eating intervention in early learning centers. BMC Health Serv Res 2017 Jan 19;17(1):57 [FREE Full text] [doi: 10.1186/s12913-017-1978-9] [Medline: 28103861]

42. Ward DS, Vaughn AE, Mazzucca S, Burney R. Translating a child care based intervention for online delivery: Development and randomized pilot study of Go NAPSACC. BMC Public Health 2017 Nov 21;17(1):891 [FREE Full text] [doi: 10.1186/s12889-017-4898-z] [Medline: 29162057]
43. Grady A, Wolfenden L, Wiggers J, Rissel C, Finch M, Flood V, et al. Effectiveness of a web-based menu-planning intervention to improve childcare service compliance with dietary guidelines: Randomized controlled trial. J Med Internet Res 2020 Feb 04;22(2):e13401 [FREE Full text] [doi: 10.2196/13401] [Medline: 32014843]
44. Yoong SL, Williams CM, Finch M, Wyse R, Jones J, Freund M, et al. Childcare service centers' preferences and intentions to use a web-based program to implement healthy eating and physical activity policies and practices: A cross-sectional study. J Med Internet Res 2015 Apr 30;17(5):e108 [FREE Full text] [doi: 10.2196/jmir.3639] [Medline: 25931430]
45. 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]
46. Mann L, Power D, MacLellan V. Development of menu planning resources for child care centres: A collaborative approach. J Child Stud 2016 Feb 17;38(2):34-40. [doi: 10.18357/jcs.v38i2.15449]
47. FoodChecker homepage. Healthy Eating Advisory Service. URL: https://foodchecker.heas.health.vic.gov.au [accessed 2021-11-08]
48. Menu planning guidelines for long day care. Healthy Eating Advisory Service. URL: https://heas.health.vic.gov.au/early-childhood-services/menu-planning/long-day-care/guidelines [accessed 2021-11-10]
49. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: A synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
50. Braun V, Clarke V, Boulton E, Davey L, McEvoy C. The online survey as a qualitative research tool. Int J Social Res Methodol 2020 Aug 16;24(6):641-654. [doi: 10.1080/13645579.2020.1805550]
51. National registers. Australian Children's Education and Care Quality Authority. URL: https://www.acecqa.gov.au/resources/national-registers [accessed 2021-11-28]
52. Spence A, Love P, Byrne R, Wakem A, Matwiejczyk L, Devine A, et al. Childcare food provision recommendations vary across Australia: Jurisdictional comparison and nutrition expert perspectives. Int J Environ Res Public Health 2020 Sep 17;17(18):6793 [FREE Full text] [doi: 10.3390/ijerph17186793] [Medline: 32957687]
53. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
54. Mayer G, Gronewold N, Alvarez S, Bruns B, Hilbel T, Schultz J. Acceptance and expectations of medical experts, students, and patients toward electronic mental health apps: Cross-sectional quantitative and qualitative survey study. JMIR Ment Health 2019 Nov 25;6(11):e14018 [FREE Full text] [doi: 10.2196/14018] [Medline: 31763990]
55. Roberts AE, Davenport TA, Wong T, Moon H, Hickie IB, LaMonica HM. Evaluating the quality and safety of health-related apps and e-tools: Adapting the Mobile App Rating Scale and developing a quality assurance protocol. Internet Interv 2021 Apr;24:100379 [FREE Full text] [doi: 10.1016/j.invent.2021.100379] [Medline: 33777705]
56. Maxwell JA. Using numbers in qualitative research. Qual Inquiry 2010 Apr 15;16(6):475-482. [doi: 10.1177/1077800410364740]
57. Flesch R. A new readability yardstick. J Appl Psychol 1948 Jun;32(3):221-233. [doi: 10.1037/h0057532] [Medline: 18867058]
58. NVivo. QSR International. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home [accessed 2021-11-11]
59. Australian Statistical Geography Standard (ASGS): Volume 5 - Remoteness Structure, July 2016. Data cube 2017 postcode to 2016 remoteness area. Australian Bureau of Statistics. URL: https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/1270.0.55.005July%202016?OpenDocument [accessed 2021-11-11]
60. Census of population and housing: Socio-Economic Indexes for Areas (SEIFA), Australia, 2016. Australian Bureau of Statistics. URL: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2033.0.55.001~2016~Main%20Features~IRSAD%20Interactive%20Map~16 [accessed 2022-05-03]
61. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
62. 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]
63. Wingrove K, Lawrence MA, Russell C, McNaughton SA. Evidence use in the development of the Australian Dietary Guidelines: A qualitative study. Nutrients 2021 Oct 23;13(11):3748 [FREE Full text] [doi: 10.3390/nu13113748] [Medline: 34836004]
64. Venkatesh V, Davis FD. A model of the antecedents of perceived ease of use: Development and test. Decision Sci 1996 Sep;27(3):451-481. [doi: 10.1111/j.1540-5915.1996.tb00860.x]

65. King WR, He J. A meta-analysis of the technology acceptance model. Inf Manag 2006 Sep;43(6):740-755. [doi: 10.1016/j.im.2006.05.003]
66. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
67. Achievement Program. Cancer Council Victoria. URL: https://www.achievementprogram.health.vic.gov.au [accessed 2021-11-28]
68. Matwiejczyk L, Mehta K, Coveney J. Factors influencing food service provision decisions in centre-based early childhood education and care services: Cooks' perspective. Health Promot J Austr 2021 Jan 29;32(1):107-116. [doi: 10.1002/hpja.308] [Medline: 31724778]
69. Otten JJ, Hirsch T, Lim C. Factors influencing the food purchases of early care and education providers. J Acad Nutr Diet 2017 May;117(5):725-734. [doi: 10.1016/j.jand.2016.10.029] [Medline: 28139425]
70. Reynolds MA, Jackson Cotwright C, Polhamus B, Gertel-Rosenberg A, Chang D. Obesity prevention in the early care and education setting: Successful initiatives across a spectrum of opportunities. J Law Med Ethics 2013 Jan 01;41 Suppl 2(S2):8-18. [doi: 10.1111/jlme.12104] [Medline: 24446993]
71. feedAustralia. Australian Government Department of Health. URL: https://feedaustralia.org.au [accessed 2022-02-04]
72. Grady A, Yoong S, Sutherland R, Lee H, Nathan N, Wolfenden L. Improving the public health impact of eHealth and mHealth interventions. Aust N Z J Public Health 2018 Apr 31;42(2):118-119. [doi: 10.1111/1753-6405.12771] [Medline: 29384248]
73. Grady A, Seward K, Finch M, Fielding A, Stacey F, Jones J, et al. Barriers and enablers to implementation of dietary guidelines in early childhood education centers in Australia: Application of the theoretical domains framework. J Nutr Educ Behav 2018 Mar;50(3):229-37.e1. [doi: 10.1016/j.jneb.2017.09.023] [Medline: 29170057]
74. Livingstone K, Olstad D, Leech R, Ball K, Meertens B, Potter J, et al. Socioeconomic inequities in diet quality and nutrient intakes among Australian adults: Findings from a nationally representative cross-sectional study. Nutrients 2017 Oct 04;9(10):1092 [FREE Full text] [doi: 10.3390/nu9101092] [Medline: 28976927]

Abbreviations

NAPSACC: Nutrition and Physical Activity Self-Assessment for Child Care
NAV: Nutrition Australia Victorian Division
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
SEP: socioeconomic position
TAM: Technology Acceptance Model

Edited by A Mavragani; submitted 13.02.22; peer-reviewed by E Hohman, L Tully; comments to author 14.04.22; revised version received 05.05.22; accepted 06.05.22; published 18.07.22.

Please cite as:
Kempler JV, Love P, Bolton KA, Rozman M, Spence AC
Exploring the Use of a Web-Based Menu Planning Tool in Childcare Services: Qualitative Cross-sectional Survey Study
JMIR Form Res 2022;6(7):e35553
URL: https://formative.jmir.org/2022/7/e35553
doi:10.2196/35553
PMID:35849438

Original Paper

Identifying Family and Unpaid Caregivers in Electronic Health Records: Descriptive Analysis

Jessica E Ma[1,2], MD; Janet Grubber[3], MSPH; Cynthia J Coffman[3,4], PhD; Virginia Wang[2,3,5,6], PhD; S Nicole Hastings[1,3,5,7,8], MHS, MD; Kelli D Allen[3,9,10], PhD; Megan Shepherd-Banigan[3,5], PhD; Kasey Decosimo[3], MPH; Joshua Dadolf[3], MSW, LICSW; Caitlin Sullivan[3], MS; Nina R Sperber[3,5], PhD; Courtney H Van Houtven[3,5,6], PhD

[1]Geriatric Research, Education, and Clinical Center, Durham Veterans Affairs Health Care System, Durham, NC, United States

[2]Division of General Internal Medicine, Department of Medicine, Duke University School of Medicine, Durham, NC, United States

[3]Center of Innovation to Accelerate Discovery and Practice Transformation, Durham Veterans Affairs Health Care System, Durham, NC, United States

[4]Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC, United States

[5]Department of Population Health Sciences, Duke University, Durham, NC, United States

[6]Duke-Margolis Center for Health Policy, Duke University, Durham, NC, United States

[7]Center for the Study of Aging, Duke University School of Medicine, Durham, NC, United States

[8]Division of Geriatrics, Department of Medicine, Duke University School of Medicine, Durham, NC, United States

[9]Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[10]Thurston Arthritis Research Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
Jessica E Ma, MD
Geriatric Research, Education, and Clinical Center
Durham Veterans Affairs Health Care System
508 Fulton St (152)
Durham, NC, 27705
United States
Phone: 1 9192860411
Email: jessica.ma@duke.edu

Abstract

Background: Most efforts to identify caregivers for research use passive approaches such as self-nomination. We describe an approach in which electronic health records (EHRs) can help identify, recruit, and increase diverse representations of family and other unpaid caregivers.

Objective: Few health systems have implemented systematic processes for identifying caregivers. This study aimed to develop and evaluate an EHR-driven process for identifying veterans likely to have unpaid caregivers in a caregiver survey study. We additionally examined whether there were EHR-derived veteran characteristics associated with veterans having unpaid caregivers.

Methods: We selected EHR home- and community-based referrals suggestive of veterans' need for supportive care from friends or family. We identified veterans with these referrals across the 8 US Department of Veteran Affairs medical centers enrolled in our study. Phone calls to a subset of these veterans confirmed whether they had a caregiver, specifically an unpaid caregiver. We calculated the screening contact rate for unpaid caregivers of veterans using attempted phone screening and for those who completed phone screening. The veteran characteristics from the EHR were compared across referral and screening groups using descriptive statistics, and logistic regression was used to compare the likelihood of having an unpaid caregiver among veterans who completed phone screening.

Results: During the study period, our EHR-driven process identified 12,212 veterans with home- and community-based referrals; 2134 (17.47%) veteran households were called for phone screening. Among the 2134 veterans called, 1367 (64.06%) answered the call, and 813 (38.1%) veterans had a caregiver based on self-report of the veteran, their caregiver, or another person in the household. The unpaid caregiver identification rate was 38.1% and 59.5% among those with an attempted phone screening and completed phone screening, respectively. Veterans had increased odds of having an unpaid caregiver if they were married (adjusted odds ratio [OR] 2.69, 95% CI 1.68-4.34), had respite care (adjusted OR 2.17, 95% CI 1.41-3.41), or had adult day health care (adjusted OR 3.69, 95% CI 1.60-10.00). Veterans with a dementia diagnosis (adjusted OR 1.37, 95% CI 1.00-1.89) or

veteran-directed care referral (adjusted OR 1.95, 95% CI 0.97-4.20) were also suggestive of an association with having an unpaid caregiver.

Conclusions: The EHR-driven process to identify veterans likely to have unpaid caregivers is systematic and resource intensive. Approximately 60% (813/1367) of veterans who were successfully screened had unpaid caregivers. In the absence of discrete fields in the EHR, our EHR-driven process can be used to identify unpaid caregivers; however, incorporating caregiver identification fields into the EHR would support a more efficient and systematic identification of caregivers.

Trial Registration: ClincalTrials.gov NCT03474380; https://clinicaltrials.gov/ct2/show/NCT03474380

(JMIR Form Res 2022;6(7):e35623) doi:10.2196/35623

KEYWORDS

veterans; caregivers; electronic health record

Introduction

In the United States, approximately 26.4 million people provide unpaid care to adults aged >50 years [1]. Family and other unpaid caregivers (hereafter called unpaid caregivers) are people who provide volunteer care to a loved one [2]. The care provided by unpaid caregivers is associated with reduced hospital readmissions and increased time at home for patients [3-5]. At the same time, unpaid caregivers are at risk of negative impacts on their physical, social, emotional, and financial well-being. Most unpaid caregivers provide care with little or no training or support [6-11]. Furthermore, unless present for appointments or enrolled in caregiver programs, caregivers are not easily identified within the health systems.

To address the unmet needs of unpaid caregivers, outreach is a critical step. However, there is no systematic method of identifying caregivers for supportive services or research interventions. Prior caregiver research studies relied on caregivers to respond to advertisements or through existing caregiver programs [12-16]. Past approaches may not be ideal for overburdened caregivers, especially for those who are not already involved in these programs, and could result in underwhelming caregiver response and participation. Moreover, passive approaches can decrease representativeness and bias toward caregivers already engaged in these interventions or those already empowered to actively seek help or services [17].

Given the need to improve the representation and systematic identification of unpaid caregivers for caregiving research and supportive services, a systematic, reliable, and proactive process to identify caregivers is needed to increase engagement, expand participation, and reduce sample bias. The electronic health record (EHR) provides an important standardized method of identifying and reaching potential caregivers for research and interventions. The aim of this paper was to evaluate an EHR-driven process used to identify unpaid caregivers for a caregiver survey study in the US Department of Veteran Affairs (VA) cluster-randomized, multisite, stepped-wedge, pragmatic trial—implementation of Helping Invested Families Improve Veteran Experiences Study (iHI-FIVES) [18]. We describe the patient-focused EHR-driven process, its caregiver identification capabilities, and veteran characteristics across the initial EHR and screening groups.

Methods

This study specifically analyzed caregiver identification through patient EHRs for a caregiver survey from the overarching iHI-FIVES study [19].

Setting

iHI-FIVES was conducted at 8 VA medical centers from April 2018 to October 2020 using a type III hybrid implementation effectiveness stepped-wedge, cluster-randomized trial design as part of the Optimizing Function and Independence Quality Enhancement Research Initiative program. The study evaluated the implementation of an unpaid caregiver program designed to promote the function and independence of veterans through caregiver group training aimed at improving caregiver coping, support seeking, and health system navigation skills [18,20]. Eligible caregivers were friends or family members who assisted a patient at home because of ongoing health problems (eg, helping them get around the house, bathe, or pay bills) [19]. This study focused on the identification of caregivers for a caregiver survey that was administered at all sites to assess secondary outcomes (caregiver burden, depression, and satisfaction with VA care).

Ethics Approval

The iHI-FIVES study was approved by the Durham VA Health Care System institutional review board (02040) and registered at ClinicalTrials.gov (NCT03474380).

EHR-Driven Process

As there is no standardized method in EHRs for identifying patients with unpaid caregivers, we evaluated an EHR-driven process to identify unpaid caregivers and patient characteristics (eg, age, race, ethnicity, sex, and comorbidities) associated with having an unpaid caregiver.

We identified 5 relevant VA home and community-based services as these 5 referral types represent an increased need for care in the veteran's home and are likely to indicate the presence of caregivers [21]. These 5 referral types included homemaker home health care, home-based primary care, adult day health care, respite care, and veteran-directed care. Homemaker home health care provides assistance in daily activities (eg, eating, getting dressed, grooming, bathing, and going to the bathroom) through an aide from a home health care agency. Adult day health care is a full- or half-day program for

veterans to be supervised for daily activities during the weekday. Home-based primary care provides veterans who have difficulty traveling to clinic visits to their illness with primary care, social work, and rehabilitation visits at home. Respite care provides home or nursing home care when a caregiver is unavailable. Veteran-directed care assists veterans in connecting to community care services for daily activities [21].

There is no uniform referral naming convention for these 5 categories of home and community-based services in the VA; therefore, we purposively selected related keywords in the text of EHR referral (eg, *%home%*, *%respit%*, *%grec%*, *%adult%*, *%adhc%*, *%hbpc%*, *%vd hcbs%*, and *%veteran directed%*). The study team reviewed referral titles identified by the keywords and then applied exclusionary conditions to the programmed code to weed out inappropriate referrals garnered by the broad search. The culled list of referral titles was also confirmed by clinical experts (physicians and caregiver support coordinators) and VA EHR programmers (Multimedia Appendix 1). Some enrolled sites requested expansion beyond these 5 referral types because of the high likelihood of having an active unpaid caregiver present (eg, skilled home health referrals). We subsequently included all these referrals, as well as one additional (respiratory therapy for chronic obstructive pulmonary disease inpatient home transition program) referral in our EHR-driven process, classifying them as *other referrals*.

Screening

We used the EHR-driven process to identify all veterans in our study window who had these qualifying referral types. The study team identified veteran patients with a qualifying referral by using VA EHR data stored in the VA Corporate Data Warehouse. Patients with referrals to hospice care were excluded [19].

Owing to the limited contact with veterans and an option to opt out, waivers of informed consent and Health Insurance Portability and Accountability Act documentation were approved for research contact with patients in this study. As the study focused on caregivers and institutional review board approval, the study team could directly confirm and seek nonveterans' participation in the screening process and study. In addition, veterans' treating clinicians were not involved in the identification process.

Identified veterans were stratified by site, sorted by the earliest qualifying referral date, and selected for recruitment over a 30-month study period. Those with the earliest dates were mailed a recruitment letter with an opt-out telephone number. The letter described that if a family member or friend helps the veteran with their ongoing health problem, they may qualify for a caregiver study. As the goal of the iHI-FIVES caregiver survey study was to enroll and collect baseline surveys for approximately 450 caregiver surveys, the study staff conducted phone screening until caregiver recruitment goals were met at each site.

For the telephone screening, the study staff used the veteran's phone number and assessed the presence of a caregiver based on information provided by the veteran, the caregiver, or another person(s) who answered the phone if neither the veteran nor the caregiver was available at the time of the call. As it is not always the case that veterans or caregivers recognize the term *caregiver* or do not identify as having or being a caregiver, the screener was written to identify assistance. Specifically, veterans were asked, "Do you have a family member or friend who helps care for you because of your ongoing health problems (for example, helping you get around the house, bathe, or pay bills?)." If nonveterans answered the phone, a similar screening question was asked, starting with "Does the Veteran..." The screening script distinguished between paid, formally trained caregivers such as home health aides and caregivers who were not paid typically and for whom the veteran had an established personal relationship (eg, friends, family). The script also asked if the veteran did not have a caregiver and whether they needed help with their care.

The study team used 2 call attempts to contact veterans after sending the letter to determine if they had a caregiver. For phone messages, the study team used a general Health Insurance Portability and Accountability Act–compliant script describing a VA research study on friends or family who help care for a veteran and a follow-up phone call attempt.

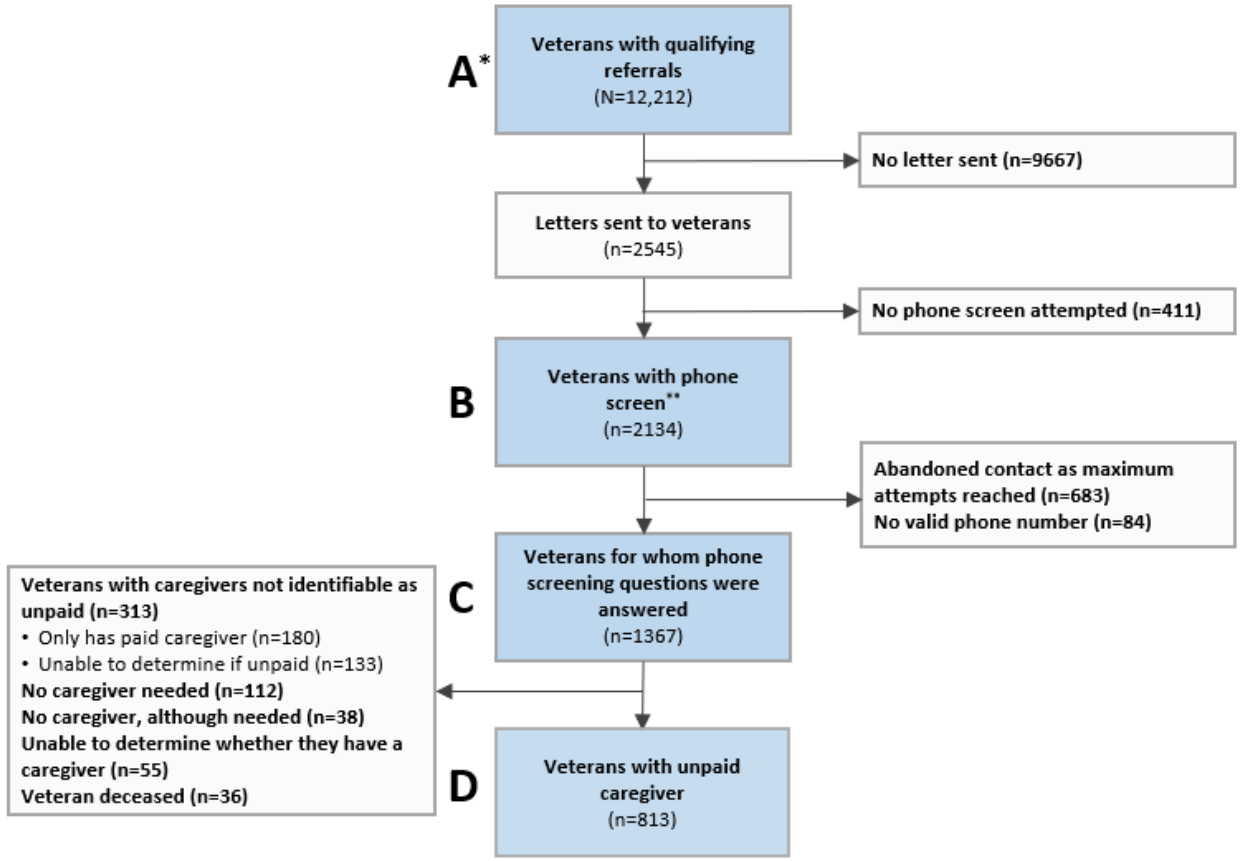
The study team used DatStat Illume (version 6.1) [22] to administer and store phone screens and baseline and 3-month survey data. In the event of a highly distressed caregiver respondent, the study staff was trained in administering a harm protocol.

EHR Data and Measures

This study also evaluated whether particular veteran characteristics were associated with veterans having an unpaid caregiver. To examine these questions, we compared four patient groups: (1) patients with a referral to home- and community-based services, (2) patients contacted for a phone screen, (3) patients with a completed phone screen, and (4) patients confirmed through screening to have an unpaid caregiver (we label these groups as boxes A-D in the CONSORT [Consolidated Standards of Reporting Trials] diagram, Figure 1).

The study identified the following patient data from the VA Corporate Data Warehouse: demographics (age, race, ethnicity, marital status, and rural residence), VA health care eligibility (measured by service connection), outpatient International Classification of Diseases–10th Revision (ICD-10) codes in the prior year, chronic health condition risk score measuring expected health care costs compared with the average patient (Nosos [23]), and home- and community-based care referrals. We used ICD-10 codes to group patient comorbidities by Quan Deyo diagnosis categories [24] to identify patients with dementia and calculate the Charlson Comorbidity Index (CCI) [24,25]. We only included dementia, as patients with dementia have known increased caregiver needs [26]. Nosos risk scores were centered around 1, representing the national expected average cost of VA patients. A risk score >1 represents a higher than expected cost for the patient, whereas a score of 3 represents a patient with an expected cost 3 times higher than the average patient [23,27]. VA service connection indicates a medical condition associated with a veteran's military service, for which the VA completely subsidizes all health care costs [28].

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram for the iHI-FIVES (implementation of Helping Invested Families Improve Veteran Experiences Study). *The blue callouts (letters A to D) indicate the columns listed in Table 1. **The phone screen may have involved speaking to only the veteran, only the caregiver, or both the veteran and the caregiver, which may have occurred during only one call or across multiple calls.



Statistical Analyses

We calculated the screening contact rate based on the study by Slattery et al [29]. In addition, we calculated the unpaid caregiver identification rates. The screening contact rate is the percentage of veterans for whom our study staff was able to talk to a person (veteran, their caregiver, or another person(s)) when we called the veterans' phone number of record, out of all phone screening calls attempted by research staff (see box B in the CONSORT diagram in Figure 1). The unpaid caregiver identification rate is the number of veterans positively identified to have an unpaid caregiver divided by the denominator of interest (phone screen attempted [box B] and phone screen questions completed [box C]; see the CONSORT diagram in Figure 1).

Descriptive statistics for the veteran characteristics are presented for the 4 subgroups described previously. Descriptive statistics were calculated for dichotomous (sex, ethnicity, rural residence, service connectedness, and dementia by ICD-10 codes), categorical (race, marital status, rural residence, highest level of education, and qualifying referral type), and continuous (age, CCI, and Nosos score) variables. Percentages were calculated for dichotomous and categorical data and means and SDs or medians and IQRs for continuous variables.

Simple logistic regression models were fit to estimate the association (odds ratio [OR] and 95% CI) between having an unpaid caregiver (vs not having one; among veterans with a completed phone screen) for each demographic and clinical variable defined previously. A multivariable logistic model was then fit with the same dependent variable but with all demographic and clinical variables examined independently in the simple models into one model to examine adjusted associations. Complete EHR data were used for the veteran factors in the analyses. The reference groups for the analyses were age (<55 years), sex (female), race (White), ethnicity (not Hispanic), marital status (never married), rural residence (urban), service connected (not service connected), dementia (absence of dementia diagnosis), CCI (score of ≤2), Nosos (score of ≤1), and referral qualification (homemaker home health care). The purpose of these models was to determine whether any veteran demographics, disease burden, health care cost risk scores, or referral types were associated with having an unpaid character with the idea that if such associations exist, they might inform attempts to recruit caregivers into training or research studies.

Analyses were performed using R (version 4.0.3) [30].

Results

Descriptives of Veteran Groups

A total of 12,212 veterans (box A in Figure 1) were identified through the EHR-driven process of 5 home- and community-based service referrals. A subset of veterans (2545/12,212, 20.84%) was sent a letter describing the study,

followed by attempts to call in order of earliest qualifying referral. Only one-fifth were sent letters and screened by phone because of recruitment goals and resource limitations. Of the 2545 veterans who were sent a letter, 2134 (83.94%; box B in Figure 1) had an attempted phone screening. Of the 2545 veterans, no phone call was attempted for 411 (16.15%); 34 of those opted out of the study prior to a phone call attempt. Of the 2134 veterans contacted for phone screening, 1367 (64.06%; box C in Figure 1) veterans were reached, and 767 (35.94%) were unable to be reached. Therefore, the screening contact rate for initial veteran screening for unpaid caregivers was 64.06% (1367/2134).

Of the 1367 veterans who completed the screening questions, 1126 (82.37%) had a caregiver; 813 (59.47%; box D in Figure 1) had an unpaid caregiver, and 180 (13.17%) were paid. We were unable to determine whether 11.81% (133/1126) of caregivers were unpaid or paid. The identification rates of unpaid caregivers were 38.10% (813/2134) and 59.47% (813/1367) among veterans with attempted phone screens and veterans with completed phone screens, respectively.

Descriptive Comparisons Between Veteran Groups

The veteran characteristics were similar among the 4 groups, including age, sex, race, urban residence, CCI, and Nosos score. The most common type of referral for veterans across all subgroups of interest was a homemaker home health care referral (Table 1). In addition, as of July 7, 2021, 38 months after the start of the study, one-third (4089/12,212, 33.48%) of the identified veterans had died.

Table 1. Electronic health record–based demographic characteristics of veterans by study milestone[a] (N=12,212).

Veteran characteristic	Veterans with qualifying referral (box A[b]; N=12,212)	Veterans with attempted phone screen (box B; n=2134)	Veterans with completed phone screen (box C; n=1367)	Veterans with unpaid caregiver (box D; n=813)
Age (years), mean (SD)	74.85 (11.95)	74.95 (11.85)	75.52 (11.57)	76.0 (11.9)
Sex (male), n (%)	11,443 (93.70)	2000 (93.72)	1285 (94.00)	772 (95.0)
Missing	1 (0.01)	0 (0.00)	0 (0.00)	0 (0.0)
Race, n (%)				
White	8723 (71.42)	1529 (71.65)	986 (72.13)	578 (71.1)
Black	2080 (17.03)	377 (17.67)	245 (17.92)	151 (18.6)
Multiple races or other	624 (5.11)	107 (5.01)	63 (4.61)	38 (4.7)
Missing	785 (6.43)	121 (5.67)	73 (5.34)	46 (5.7)
Ethnicity, n (%)				
Hispanic	400 (3.28)	74 (3.47)	36 (2.63)	25 (3.1)
Missing	447 (3.66)	71 (3.33)	46 (3.37)	23 (2.8)
Marital status, n (%)				
Never married	994 (8.14)	176 (8.25)	103 (7.53)	44 (5.4)
Divorced or separated	3297 (27.00)	553 (25.91)	332 (24.29)	157 (19.3)
Married	6155 (50.40)	1105 (51.78)	741 (54.21)	509 (62.6)
Widow	1691 (13.85)	290 (13.59)	185 (13.53)	99 (12.2)
Missing	75 (0.61)	10 (0.47)	6 (0.44)	4 (0.5)
Urban residence, n (%)	7752 (63.48)	1367 (64.06)	864 (63.20)	512 (63.0)
Missing	3 (0.02)	1 (0.05)	1 (0.07)	0 (0.0)
Service connected, n (%)	7050 (57.73)	1253 (58.72)	814 (59.55)	501 (61.6)
Missing	1 (0.01)	0 (0.00)	0 (0.00)	0 (0.0)
Dementia, n (%)	2597 (21.27)	485 (22.73)	320 (23.41)	227 (27.9)
Missing	26 (0.21)	3 (0.14)	2 (0.15)	1 (0.1)
CCI,[c] median (IQR 25th-75th percentile)	2.00 (1.00-4.00)	2.00 (1.00-4.00)	2.00 (1.00-4.00)	2.0 (1.0-4.0)
Missing, n (%)	26 (0.21)	3 (0.14)	2 (0.15)	1 (0.1)
Nosos, median (IQR 25th-75th percentile)	1.88 (0.89-3.87)	2.06 (0.97-3.99)	2.06 (0.96-3.89)	2.1 (0.9-3.9)
Missing, n (%)	129 (1.06)	10 (0.47)	7 (0.51)	5 (0.6)
Referral, n (%)				
Homemaker home health care	7699 (63.04)	1368 (64.10)	873 (63.86)	494 (60.8)
Adult day health care	440 (3.60)	87 (4.08)	51 (3.73)	45 (5.5)
Home-based primary care	746 (6.11)	111 (5.20)	71 (5.19)	37 (4.6)
Respite care	929 (7.61)	221 (10.36)	159 (11.63)	122 (15.0)
Veteran-directed care	358 (2.93)	77 (3.61)	46 (3.37)	35 (4.3)
Other[d]	2040 (16.70)	270 (12.65)	167 (12.22)	80 (9.8)

[a]Veterans were identified by first home or community referral placed during the study time frame.

[b]Boxes on the CONSORT (Consolidated Standards of Reporting Trials) diagram (Figure 1).

[c]CCI: Charlson Comorbidity Index.

[d]Some enrolled sites requested the addition of specific referral types (eg, skilled home health, nursing home, and specialty home care) because of the high probability of an unpaid caregiver being present when these referrals occur. We subsequently included all these referrals, as well as one additional (respiratory therapy for chronic obstructive pulmonary disease inpatient home transition program) referral, in our electronic health record–driven process, classifying them as *other*.

Certain characteristics were different among the groups. Compared with veterans with a qualifying referral (box A in Figure 1) and those contacted for phone screen (box B in Figure 1), veterans who answered phone screening (box C in Figure 1) had higher proportions of veterans who were aged 75 to 84 years (396/1367, 28.97%), married (741/1367, 54.21%), and referred for respite care (159/1367, 11.63%). Veterans with unpaid caregivers (813/12,212, 6.66%; box D in Figure 1) had higher proportions of initial referrals for respite (122/813, 15%), adult day care (45/813, 5.5%), and veteran-directed care (35/813, 4.3%) than those of the other groups.

Modeled Associations Between Veteran Characteristics and Presence of an Unpaid Caregiver

Among veterans for whom screening questions were answered and were without missing data (1261/12,212, 10.22%), the odds of having an unpaid caregiver were higher among married veterans (unadjusted OR 2.90, 95% CI 1.88-4.50) than among unmarried veterans and veterans with dementia (unadjusted OR 2.06, 95% CI 1.55-2.75) than those without dementia. Similarly, the odds of having an unpaid caregiver were higher among veterans referred for respite services (unadjusted OR 2.64, 95% CI 1.76-4.06), adult day health care (unadjusted OR 4.82, 95% CI 2.17-12.80), and veteran-directed care (unadjusted OR 2.07, 95% CI 1.05-4.38) than for veterans referred to homemaker home health care. In addition to dementia diagnosis and veteran-directed care, adjusted ORs for the associations between having an unpaid caregiver and marital status, as well as health care referral type, were similar to the unadjusted results (Table 2).

Table 2. ORs[a] and 95% CIs from simple logistic regression models (unadjusted) and multiple logistic regression models (adjusted) for the association between veterans having unpaid caregivers among veterans with a completed phone screen and without missing data (N=1261)[b] and veteran electronic health record characteristics.

Veteran characteristic	Unadjusted[c]		Adjusted[d]	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Age (years)				
<55	1.00 (reference)	N/A[e]	1.00 (reference)	N/A
55-64	0.78 (0.41-1.47)	.45	0.71 (0.36-1.37)	.31
65-74	0.71 (0.40-1.24)	.24	0.56 (0.30-1.01)	.06
75-84	0.96 (0.54-1.69)	.89	0.66 (0.34-1.23)	.20
>84	1.17 (0.65-2.08)	.60	0.80 (0.41-1.53)	.50
Sex				
Female	1.00 (reference)	N/A	1.00 (reference)	N/A
Male	1.49 (0.94-2.38)	.09	1.13 (0.68-1.88)	.63
Ethnicity				
Non-Hispanic	1.00 (reference)	N/A	1.00 (reference)	N/A
Hispanic	1.38 (0.68-2.99)	.39	0.87 (0.40-1.97)	.73
Race				
White	1.00 (reference)	N/A	1.00 (reference)	N/A
Black	1.14 (0.85-1.53)	.38	1.33 (0.97-1.84)	.08
Multiple races or other	1.17 (0.69-2.02)	.56	1.18 (0.67-2.10)	.57
Marital status				
Never married	1.00 (reference)	N/A	1.00 (reference)	N/A
Divorced or separated	1.17 (0.74-1.86)	.50	1.27 (0.79-2.06)	.33
Married	2.90 (1.88-4.50)	<.001	2.69 (1.68-4.34)	<.001
Widow	1.50 (0.91-2.49)	.11	1.43 (0.83-2.49)	.20
Rural residence				
Urban	1.00 (reference)	N/A	1.00 (reference)	N/A
Rural	1.04 (0.83-1.32)	.72	1.01 (0.78-1.30)	>.95
Service connection				
Not service connected	1.00 (reference)	N/A	1.00 (reference)	N/A
Service connected	1.20 (0.95-1.50)	.13	1.17 (0.90-1.51)	.24
Dementia				
No diagnosis of dementia	1.00 (reference)	N/A	1.00 (reference)	N/A
Dementia diagnosis	2.06 (1.55-2.75)	<.001	1.37 (1.00-1.89)	.06
Charlson comorbidity index				
≤2	1.00 (reference)	N/A	1.00 (reference)	N/A
>2	1.08 (0.86-1.35)	.52	1.13 (0.87-1.47)	.35
Nosos				
≤1	1.00 (reference)	N/A	1.00 (reference)	N/A
>1 and <4	0.87 (0.66-1.14)	.32	0.86 (0.63-1.17)	.33
≥4	0.91 (0.66-1.25)	.56	1.00 (0.68-1.46)	>.98
Referral				
Homemaker home health care	1.00 (reference)	N/A	1.00 (reference)	N/A

Veteran characteristic	Unadjusted[c]		Adjusted[d]	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Adult day health care	4.82 (2.17-12.80)	<.001	3.69 (1.60-10.00)	.005
Home-based primary care	0.83 (0.50-1.39)	.48	0.87 (0.51-1.47)	.60
Respite care	2.64 (1.76-4.06)	<.001	2.17 (1.41-3.41)	<.001
Veteran-directed care	2.07 (1.05-4.38)	.04	1.95 (0.97-4.20)	.07
Other	0.69 (0.49-0.97)	.03	0.77 (0.54-1.10)	0.15

[a]OR: odds ratio.

[b]A complete case analysis was used, and veterans with missing data (n=106) were excluded from this analysis.

[c]Unadjusted ORs from a simple logistic regression model, including a single electronic health record characteristic as an independent variable.

[d]Adjusted ORs from multiple logistic regression models, including all electronic health record characteristics.

[e]N/A: not applicable.

Discussion

Principal Findings

We describe an approach to systematically identify family and other unpaid caregivers in a caregiver study using the VA EHR. This EHR process was effective in identifying caregivers through EHR home- and community-based referrals.

We identified >12,000 veterans with relevant home- and community-based referrals. We called >2000 veterans for screening. Among the 1367 veterans who answered the phone screening, 813 (59.47%) had a family or other unpaid caregivers. Through the EHR-driven process, we found that veteran characteristics were similar among all 4 identified groups (veterans with qualifying referrals, attempted phone screening, completed phone screen, and unpaid caregivers); some notable exceptions included marital status, dementia, and referral type. Veterans with unpaid caregivers were more likely to be married and have a referral for adult day health care or respite care. Although not statistically significant in the adjusted analysis, dementia diagnosis and veteran-directed care referral were also suggestive of an association with having an unpaid caregiver.

Comparison With Prior Work

This study is unique in that there are currently few replicable or standardized EHR methods for identifying veterans with unpaid caregivers. Using the EHR to identify veterans with unpaid caregivers can uniquely position the VA to identify caregivers interested in participating in research studies and increase the representation of caregivers that may not be identified through traditional avenues [31-34]. However, although EHRs can be useful for identifying patient participants through standardized fields for patient demographic, laboratory, medication, and diagnostic data, caregiver identification through EHRs can be difficult because it is rare to have a discrete caregiver field in EHRs [35].

As there is no discrete caregiver field in the VA EHR, we used selected home- and community-based referrals to define a patient population that may have an unpaid caregiver. Although this EHR-driven process was able to identify caregivers independent of their participation or nonparticipation in existing VA caregiver support and services, the process was resource intensive. To meet the goal of 450 caregiver surveys, 4 research staff members would spend approximately 8 to 16 hours every week during active data collection screening veterans or caregivers; screening approximately one-fifth of the identified EHR referral population resulted in the reaching of survey goals.

Therefore, although the EHR-driven process is a helpful approach to identifying caregivers not already directly linked to services in a health care system (and to be clear, most health systems do not have large-scale programs of caregiver support such as seen in VA), potential future improvements to the EHR should consider a discrete caregiver field to improve efficiency in identifying caregivers for research studies and interventions. VA is currently making investments to link caregiver records to veteran EHR records, beginning with those caregivers already engaged in the VA Caregiver Support Program support and services. This will allow the examination of program impacts on the outcomes of the enrolled caregivers and veterans over time.

Importantly, these caregiver EHR fields should be assessed using an activity-based brief screening question based on the types of assistance received in the home because of a patient's health problems rather than whether a patient *has* a caregiver. The term caregiver is problematic for many older adults; many patients do not recognize that they have a caregiver, and many family members and friends do not identify themselves as caregivers [36].

This study identified several veteran characteristics associated with having an unpaid caregiver. First, we found that being married was positively associated with a veteran having an unpaid caregiver. Although marital status is rarely a variable in other EHRs or insurance claims (ie, Medicaid or Medicare) data, marital status could be a useful target for the identification of caregivers within the VA and potentially in other health systems, with certain caveats. Specifically, marital status alone may not be sufficient to determine the presence of a caregiver as the marital status may be out of date in the EHR and dependent on individual sites and personnel to update this discrete field.

Second, in the unadjusted analysis, veterans with dementia were more likely to have unpaid caregivers. The OR and 95% CI for this association were 1.37 (1.00-1.89) in the adjusted analysis,

suggesting a moderate and likely clinically meaningful association. A dementia diagnosis is more likely to be in the medical record and may identify additional unpaid caregivers in the absence of information about the home care network in the EHR.

Third, focusing on homemaker home health care, respite care referrals, and veteran-directed care may also identify unpaid caregivers. Although most (494/813, 60.8%) of veterans with an unpaid caregiver had a homemaker home health care referral, veterans with respite, adult day health care, and veteran-directed care referrals were more likely to have an unpaid caregiver (122/813, 15%; 45/813, 5.5%; and 35/813, 4.3%, respectively) than veterans identified by the initial EHR pool (929/12,212, 7.61%; 440/12,212, 3.60%; and 358/12,212, 2.93%, respectively). This is likely because respite care, adult day health care, and veteran-directed care programs are specifically designed to assist unpaid caregivers who need time separate from the veterans [37,38]. As such, depending on the purpose of the study, future VA EHR-driven approaches may want to narrow the types of home- and community-based services from which to recruit unpaid caregivers, focusing on respite care, adult day health care, and veteran-directed care.

Strengths and Limitations

A strength of this study is the creation of a systematic approach to screening a large number of veterans across multiple VA sites in the United States. Compared with a prior study at a VA medical center where 89% of patients referred to home- and community-based services reported having an unpaid caregiver, the systematic approach taken here at multiple sites with more geographic representation found that 59.47% (813/1367) of veterans had an unpaid caregiver [39]. Moreover, the percentage of unpaid caregivers may be larger, as we were unable to determine the caregiver status for 16.39% (224/1367) of the patients screened. Using geographically diverse sites, this study provides a more generalizable estimate using EHR-based, home- and community-based referrals to identify unpaid caregivers.

There are several limitations to consider. First, this analysis was limited to examining the associations between veteran characteristics and the presence of unpaid caregivers. We did not have caregiver information for veterans who were not screened or who had unpaid caregiver characteristics.

Second, the generalizability of this study outside the VA health system may be limited, as this study uses VA home- and community-based referrals that may not be available in non-VA settings. Despite this limitation, this was a multisite study, and veterans identified in the EHR were from VAs of various sizes in both urban and rural settings across many regions of the United States. Although the application of this process to identify caregivers may be limited to VA, the process described in this study may be applicable to many VA settings across the country looking to identify caregivers for research studies or for caregiver interventions. Other health care systems in the private sector with a unified EHR system that uses home- and community-based referrals could also use our approach to identify unpaid caregivers.

Finally, because of the resource-intensive process, we screened one-fifth of the veterans identified in the EHR-driven process. Even so, >1000 veterans answered phone screenings for unpaid caregivers. Even if the exact count of unpaid caregivers among all veterans identified is unattainable, a large number of veterans were successfully screened, and there was similarity in veteran characteristics across veterans with a qualifying referral and phone screen; therefore, our results may suggest a similar percentage of unpaid caregivers among veterans who were not contacted for the study.

Conclusions

We present a systematic process for the identification of unpaid caregivers using veteran referrals and additional EHR factors for a caregiver survey study. Using this EHR-driven process, we were able to positively identify the presence of unpaid caregivers for approximately 60% (813/1367) of veterans whose households responded to phone screening calls. In the setting of a resource-intensive process, additional research is needed to increase the efficiency and effectiveness of caregiver identification efforts in EHRs for caregiver interventions and research. Potential future steps may include the incorporation of a discrete caregiver field into the EHR. This type of field would not only improve the efficiency of identifying caregivers within the EHR but also the clinical care for patients in recognizing and including the primary caregiver in the health care team.

Acknowledgments

This work was funded by the US Department of Veterans Affairs (VA) Quality Enhancement Research Initiative (QUE-16-170) and by the Center of Innovation to Accelerate Discovery and Practice Transformation (CIN 13-410) of the Durham VA Health Care System. The efforts of JEM were supported by the US Department of VA Office of Academic Affiliations postdoctoral fellowship (TPH 21-000) and the Durham VA Center of Innovation to Accelerate Discovery and Practice Transformation (CIN 13-410). MSB was supported by the US Department of VA, Veterans Health Administration, Office of Research and Development, and United States Department of VA Health Services Research and Development (HSR&D) Career Development Award (CDA 17-006). The funding agencies had no role in the design or conduct of the study; collection, analysis, or interpretation of the data; or the preparation, review, or approval of the manuscript. The contents do not represent the views of the US Department of VA or the US Government. iHI-FIVES (implementation of Helping Invested Families Improve Veteran Experiences Study) was implemented at the participating sites as a clinical program; evaluation was approved as human subjects research by the Durham VA institutional review board (protocol #02040). The authors thank the caregivers who participated in the study. The authors

would also like to thank additional members of the research team who supported the iHI-FIVES administration and data collection: Elizabeth Mahanna, Katina Morris, Nadya Majette Elliott, Matthew Tucker, Cassie Meyer, and Rebecca Bruening.

Authors' Contributions

JEM, CHVH, and JG were involved in the study conceptualization. SNH, KDA, CHVH, and VW were involved in the funding acquisition. KD, JD, and CS were involved in project coordination and data collection. JEM, JG, CJC, CHVH, VW, and KD wrote the original article. JEM, JG, CJC, CHVH, VW, NRS, MSB, KD, CS, and JD reviewed and edited the article. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
SQL search term algorithm for the 5 qualifying home- and community-based services.
[PDF File (Adobe PDF File), 64 KB - formative_v6i7e35623_app1.pdf]

References

1. Freedman VA, Wolff JL. The changing landscape of family caregiving in the United States. In: Sawhill IV, Stevenson B, Aaronson SR, Freedman VA, Goldin C, Jacobs E, et al, editors. Paid Leave for Caregiving: Issues and Answers. Washington, DC, USA: Brookings Institution; Nov 2020:11-30.
2. Schulz R, Tompkins CA, National Research Council (US) Committee on the Role of Human Factors in Home Health Care. Informal caregivers in the United States: prevalence, caregiver characteristics, and ability to provide care. In: The Role of Human Factors in Home Health Care: Workshop Summary. Washington, DC, USA: National Academies Press (US); 2010:7.
3. Charles KK, Sevak P. Can family caregiving substitute for nursing home care? J Health Econ 2005 Nov;24(6):1174-1190. [doi: 10.1016/j.jhealeco.2005.05.001] [Medline: 15979742]
4. Van Houtven CH, Norton EC. Informal care and health care use of older adults. J Health Econ 2004 Nov;23(6):1159-1180. [doi: 10.1016/j.jhealeco.2004.04.008] [Medline: 15556241]
5. Rodakowski J, Rocco PB, Ortiz M, Folb B, Schulz R, Morton SC, et al. Caregiver integration during discharge planning for older adults to reduce resource use: a metaanalysis. J Am Geriatr Soc 2017 Aug;65(8):1748-1755 [FREE Full text] [doi: 10.1111/jgs.14873] [Medline: 28369687]
6. Bevans M, Sternberg EM. Caregiving burden, stress, and health effects among family caregivers of adult cancer patients. JAMA 2012 Jan 25;307(4):398-403 [FREE Full text] [doi: 10.1001/jama.2012.29] [Medline: 22274687]
7. Van Houtven CH, Coe NB, Skira MM. The effect of informal care on work and wages. J Health Econ 2013 Jan;32(1):240-252. [doi: 10.1016/j.jhealeco.2012.10.006] [Medline: 23220459]
8. Van Houtven CH, Oddone EZ, Weinberger M. Informal and formal care infrastructure and perceived need for caregiver training for frail US veterans referred to home and community-based services. Chronic Illn 2010 Mar;6(1):57-66 [FREE Full text] [doi: 10.1177/1742395309352694] [Medline: 20308351]
9. Kalra L, Evans A, Perez I, Melbourn A, Patel A, Knapp M, et al. Training carers of stroke patients: randomised controlled trial. BMJ 2004 May 08;328(7448):1099 [FREE Full text] [doi: 10.1136/bmj.328.7448.1099] [Medline: 15130977]
10. Hepburn KW, Tornatore J, Center B, Ostwald SW. Dementia family caregiver training: affecting beliefs about caregiving and caregiver outcomes. J Am Geriatr Soc 2001 Apr;49(4):450-457. [doi: 10.1046/j.1532-5415.2001.49090.x] [Medline: 11347790]
11. Burgdorf J, Roth DL, Riffin C, Wolff JL. Factors associated with receipt of training among caregivers of older adults. JAMA Intern Med 2019 Jun 01;179(6):833-835 [FREE Full text] [doi: 10.1001/jamainternmed.2018.8694] [Medline: 30958503]
12. Leslie M, Khayatzadeh-Mahani A, MacKean G. Recruitment of caregivers into health services research: lessons from a user-centred design study. Res Involv Engagem 2019 May 20;5:17 [FREE Full text] [doi: 10.1186/s40900-019-0150-6] [Medline: 31139432]
13. Bruening R, Sperber N, Miller K, Andrews S, Steinhauser K, Wieland GD, et al. Connecting caregivers to support: lessons learned from the VA caregiver support program. J Appl Gerontol 2020 Apr;39(4):368-376. [doi: 10.1177/0733464818825050] [Medline: 30658547]
14. Preissner K, Finlayson M, Henkel C. Recruiting for caregiver education research: perspectives of caregivers of people with multiple sclerosis. Int J MS Care 2012;14(4):188-196 [FREE Full text] [doi: 10.7224/1537-2073-14.4.188] [Medline: 24453751]
15. Hansen D, Petrinec A, Hebeshy M, Sheehan D, Drew BL. Advancing the science of recruitment for family caregivers: focus group and Delphi methods. JMIR Nurs 2019 Jul 22;2(1):e13862 [FREE Full text] [doi: 10.2196/13862] [Medline: 34345769]

16. Morrison K, Winter L, Gitlin LN. Recruiting community-based dementia patients and caregivers in a nonpharmacologic randomized trial: what works and how much does it cost? J Appl Gerontol 2016 Jul;35(7):788-800 [FREE Full text] [doi: 10.1177/0733464814532012] [Medline: 24799354]
17. Dura JR, Kiecolt-Glaser JK. Sample bias in caregiving research. J Gerontol 1990 Sep;45(5):P200-P204. [doi: 10.1093/geronj/45.5.p200] [Medline: 2394916]
18. Shepherd-Banigan M, Kaufman BG, Decosimo K, Dadolf J, Boucher NA, Mahanna EP, et al. Adaptation and implementation of a family caregiver skills training program: from single site RCT to multisite pragmatic intervention. J Nurs Scholarsh 2020 Jan;52(1):23-33 [FREE Full text] [doi: 10.1111/jnu.12511] [Medline: 31497935]
19. Wang V, Allen K, Van Houtven CH, Coffman C, Sperber N, Mahanna EP, et al. Supporting teams to optimize function and independence in veterans: a multi-study program and mixed methods protocol. Implement Sci 2018 Apr 20;13(1):58 [FREE Full text] [doi: 10.1186/s13012-018-0748-3] [Medline: 29678137]
20. Van Houtven CH, Oddone EZ, Hastings SN, Hendrix C, Olsen M, Neelon B, et al. Helping Invested Families Improve Veterans' Experiences Study (HI-FIVES): study design and methodology. Contemp Clin Trials 2014 Jul;38(2):260-269 [FREE Full text] [doi: 10.1016/j.cct.2014.05.003] [Medline: 24837544]
21. Geriatrics and Extended Care: Home and Community Based Services. U.S. Department of Veterans Affairs. 2021. URL: https://www.va.gov/GERIATRICS/pages/Home_and_Community_Based_Services.asp?utm_source=geriatrics_home_page [accessed 2022-03-07]
22. DatStat Illume, Version 6.1. DatStat. 2016. URL: https://www.datstat.com/products/illume-next [accessed 2018-04-01]
23. Wagner T, Stefos T, Moran E, Cashy J, Shen ML, Gehlert E, et al. Technical Report 30: Risk Adjustment: Guide to the V21 and Nosos Risk Score Programs. U.S. Department of Veterans Affairs. Menlo Park, CA, USA: Health Economics Resource Center; 2016. URL: https://www.herc.research.va.gov/include/page.asp?id=technical-report-risk-adjustment [accessed 2020-09-28]
24. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005 Nov;43(11):1130-1139. [doi: 10.1097/01.mlr.0000182534.19832.83] [Medline: 16224307]
25. Glasheen WP, Cordier T, Gumpina R, Haugh G, Davis J, Renda A. Charlson Comorbidity Index: ICD-9 update and ICD-10 translation. Am Health Drug Benefits 2019;12(4):188-197 [FREE Full text] [Medline: 31428236]
26. Wolff JL, Spillman BC, Freedman VA, Kasper JD. A national profile of family and unpaid caregivers who assist older adults with health care activities. JAMA Intern Med 2016 Mar;176(3):372-379 [FREE Full text] [doi: 10.1001/jamainternmed.2015.7664] [Medline: 26882031]
27. Gehlert E. HERC: Nosos Risk Adjustment. U.S. Department of Veterans Affairs. URL: https://www.herc.research.va.gov/include/page.asp?id=risk-adjustment [accessed 2021-09-12]
28. Your health care costs. U.S. Department of Veterans Affairs. 2020. URL: https://www.va.gov/health-care/about-va-health-benefits/cost-of-care/ [accessed 2021-09-12]
29. Slattery ML, Edwards SL, Caan BJ, Kerber RA, Potter JD. Response rates among control subjects in case-control studies. Ann Epidemiol 1995 May;5(3):245-249. [doi: 10.1016/1047-2797(94)00113-8] [Medline: 7606315]
30. R Core Team. R. Version 4.0.3. The R Project for Statistical Computing. 2017. URL: https://www.r-project.org/ [accessed 2020-09-18]
31. Coorevits P, Sundgren M, Klein GO, Bahr A, Claerhout B, Daniel C, et al. Electronic health records: new opportunities for clinical research. J Intern Med 2013 Dec;274(6):547-560 [FREE Full text] [doi: 10.1111/joim.12119] [Medline: 23952476]
32. Beskow LM, Brelsford KM, Hammack CM. Patient perspectives on use of electronic health records for research recruitment. BMC Med Res Methodol 2019 Feb 26;19(1):42 [FREE Full text] [doi: 10.1186/s12874-019-0686-z] [Medline: 30808279]
33. Thadani SR, Weng C, Bigger JT, Ennever JF, Wajngurt D. Electronic screening improves efficiency in clinical trial recruitment. J Am Med Inform Assoc 2009;16(6):869-873 [FREE Full text] [doi: 10.1197/jamia.M3119] [Medline: 19717797]
34. Weng C, Appelbaum P, Hripcsak G, Kronish I, Busacca L, Davidson KW, et al. Using EHRs to integrate research with patient care: promises and challenges. J Am Med Inform Assoc 2012;19(5):684-687 [FREE Full text] [doi: 10.1136/amiajnl-2012-000878] [Medline: 22542813]
35. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. Annu Rev Public Health 2016;37:61-81 [FREE Full text] [doi: 10.1146/annurev-publhealth-032315-021353] [Medline: 26667605]
36. Committee on Family Caregiving for Older Adults, Board on Health Care Services, Health and Medicine Division, National Academies of Sciences, Engineering, and Medicine. Family Caregiving Roles and Impacts. In: Schulz R, Eden J, editors. Families Caring for an Aging America. Washington, DC: National Academies Press (US); 2016:73-122.
37. Geriatrics and Extended Care: Respite Care. U.S. Department of Veterans Affairs. URL: https://www.va.gov/GERIATRICS/pages/Respite_Care.asp [accessed 2021-09-12]
38. Geriatrics and Extended Care: Adult Day Health Care. U.S. Department of Veterans Affairs. URL: https://www.va.gov/GERIATRICS/pages/Adult_Day_Health_Care.asp [accessed 2021-09-12]

39. Van Houtven CH, Smith VA, Lindquist JH, Chapman JG, Hendrix C, Hastings SN, et al. Family Caregiver Skills Training to Improve Experiences of Care: a Randomized Clinical Trial. J Gen Intern Med 2019 Oct;34(10):2114-2122. [doi: 10.1007/s11606-019-05209-x] [Medline: 31388914]

Abbreviations

CCI: Charlson Comorbidity Index
CONSORT: Consolidated Standards of Reporting Trials
EHR: electronic health record
HSR&D: Health Services Research and Development
ICD-10: International Classification of Diseases–10th Revision
iHI-FIVES: implementation of Helping Invested Families Improve Veteran Experiences Study
OR: odds ratio
VA: Veterans Affairs

Edited by T Leung, A Mavragani; submitted 10.12.21; peer-reviewed by L Nichols, L Beskow; comments to author 11.01.22; revised version received 08.03.22; accepted 22.04.22; published 18.07.22.

Please cite as:
Ma JE, Grubber J, Coffman CJ, Wang V, Hastings SN, Allen KD, Shepherd-Banigan M, Decosimo K, Dadolf J, Sullivan C, Sperber NR, Van Houtven CH
Identifying Family and Unpaid Caregivers in Electronic Health Records: Descriptive Analysis
JMIR Form Res 2022;6(7):e35623
URL: https://formative.jmir.org/2022/7/e35623
doi:10.2196/35623
PMID:35849430

Original Paper

Machine Learning Prediction of Hypoglycemia and Hyperglycemia From Electronic Health Records: Algorithm Development and Validation

Harald Witte[1], PhD; Christos Nakas[1,2], PhD; Lia Bally[3], MD; Alexander Benedikt Leichtle[1,4], MD

[1]University Institute of Clinical Chemistry, Inselspital - Bern University Hospital and University of Bern, Bern, Switzerland

[2]Laboratory of Biometry, University of Thessaly, Volos, Greece

[3]Department of Diabetes, Endocrinology, Nutritional Medicine and Metabolism, Inselspital - Bern University Hospital and University of Bern, Bern, Switzerland

[4]Center of Artificial Intelligence in Medicine, University of Bern, Bern, Switzerland

Corresponding Author:
Alexander Benedikt Leichtle, MD
University Institute of Clinical Chemistry
Inselspital - Bern University Hospital and University of Bern
Freiburgstrasse 10
Bern, 3010
Switzerland
Phone: 41 316328330
Email: Alexander.Leichtle@insel.ch

Abstract

Background: Acute blood glucose (BG) decompensations (hypoglycemia and hyperglycemia) represent a frequent and significant risk for inpatients and adversely affect patient outcomes and safety. The increasing need for BG management in inpatients poses a high demand on clinical staff and health care systems in addition.

Objective: This study aimed to generate a broadly applicable multiclass classification model for predicting BG decompensation events from patients' electronic health records to indicate where adjustments in patient monitoring and therapeutic interventions are required. This should allow for taking proactive measures before BG levels are derailed.

Methods: A retrospective cohort study was conducted on patients who were hospitalized at a tertiary hospital in Bern, Switzerland. Using patient details and routine data from electronic health records, a multiclass prediction model for BG decompensation events (<3.9 mmol/L [hypoglycemia] or >10, >13.9, or >16.7 mmol/L [representing different degrees of hyperglycemia]) was generated based on a second-level ensemble of gradient-boosted binary trees.

Results: A total of 63,579 hospital admissions of 38,250 patients were included in this study. The multiclass prediction model reached specificities of 93.7%, 98.9%, and 93.9% and sensitivities of 67.1%, 59%, and 63.6% for the main categories of interest, which were nondecompensated cases, hypoglycemia, or hyperglycemia, respectively. The median prediction horizon was 7 hours and 4 hours for hypoglycemia and hyperglycemia, respectively.

Conclusions: Electronic health records have the potential to reliably predict all types of BG decompensation. Readily available patient details and routine laboratory data can support the decisions for proactive interventions and thus help to reduce the detrimental health effects of hypoglycemia and hyperglycemia.

(JMIR Form Res 2022;6(7):e36176) doi:10.2196/36176

KEYWORDS

diabetes; blood glucose decompensation; multiclass prediction model; dysglycemia; hyperglycemia; hypoglycemia

Introduction

Blood Glucose Decompensations Are Associated With Poor Outcomes in Inpatients

Diabetes is one of the most common lifestyle diseases worldwide (affecting approximately 537 million people in 2021), particularly among older adults (aged >65 years), with numbers on the rise [1,2]. Consequently, the number of inpatients with diabetes is also increasing, with currently approximately every sixth hospital bed being used by patients with diabetes [3,4].

The hallmark of diabetes is loss of control over blood glucose (BG) levels. Failure to maintain BG levels within the ranges normally set by functional glucose homeostasis in nondiabetic people manifests as hyper- and hypoglycemia, where BG levels decompensate, that is, exceed or fall below a critical threshold, respectively.

Both hypoglycemia and hyperglycemia have been associated with numerous complications in inpatients. This includes an increased length of stay [5], diabetic ketoacidosis or hyperglycemic hyperosmolar state in the case of hyperglycemia [6], increased risk of infection [7,8], admission to intensive care units (ICUs) [9], and an overall increase in mortality [10-12]. The association between hypoglycemia and hyperglycemia and poor outcomes in inpatients who are critically and noncritically ill calls for a rigorous inpatient management approach toward reducing dysglycemia [13,14].

Challenges of Inpatient Dysglycemia Management

The environment of inpatients (ie, a hospital setting) is usually well-controlled; nevertheless, the maintenance of BG levels in a normoglycemic range is demanding. It is complicated by the fragile health status of the inpatient, stress (including postoperative stress), prolonged fasting, changes in diet and meal timings, and inadequate dosing of or changes in the type of antidiabetic drugs administered to name a few main troubles [14,15]. Standard diabetes therapy using adjusted subcutaneous insulin injections (sliding-scale insulin) combined with insufficient glucose monitoring is another issue and a potential risk factor for hypoglycemia [16], aggravated by the frequent shortage of nursing staff in hospitals.

The combination of continuous glucose monitoring (CGM) systems and subcutaneous insulin pumps with automated insulin delivery (closed-loop) systems is a recent promising development [17-19], which may reduce the workload of clinical staff in addition to benefiting patients.

Prediction of BG Decompensation

An alternative approach to continuous BG measurements is the prevention of BG decompensation by the detection of early signs or patterns associated with it, thereby avoiding the immediate and long-term adverse effects of hypoglycemia and hyperglycemia [20]. To date, studies have primarily assessed the prediction of hypoglycemia [20] based on, for instance, laboratory data [21,22] or data from electrocardiograms [23], or they required subcutaneous glucose readings from CGM systems [24].

Aim of This Study

This study investigated whether readily available standard laboratory results and patient information could be used to reliably predict both hypoglycemia and hyperglycemia in inpatients with a clinically relevant prediction horizon.

Methods

Patient Cohort

The data set contained anonymized hospital admission data collected during the routine management of adult (aged ≥18 years) inpatients of the 6 hospitals of the Insel-Gruppe (Bern, Switzerland) from January 1, 2014, to December 5, 2019. Data were retrieved as is; that is, no extra data were collected for this retrospective cohort study. Data of patients whose BG levels had been assessed at least once were included if they met ≥1 of the following inclusion criteria:

Diagnosis of diabetes or diabetes-related syndromes (codes as specified in the 10th revision of the International Statistical Classification of Diseases and Related Health Problems of the World Health Organization [25]), including E10 to E14, E16, E66 to E68, G59, G63, H28, H36, K77.8, K85, M14.2, N08.3, O24, R73, and R81

Administration of an antidiabetic drug falling into code category A10 (or subcategories thereof) of the Anatomical Therapeutic Chemical Classification System

Extreme BG levels regardless of any formal diagnosis of diabetes, including a BG level of <4.0 or ≥11.1 mmol/L measured at any time, a fasting venous BG level of ≥7.0 mmol/L, a 2-hour value of ≥11.1 mmol/L during an oral glucose tolerance test, or an HbA_{1c} value of ≥48.0 mmol/mol (International Federation of Clinical Chemistry) or ≥6.5% (Diabetes Control and Complications Trial and National Glycohemoglobin Standardization Program)

The rationale behind these 3 inclusion criteria was to include a broad range of patients with potential indications of dysglycemia. The fact that BG level tests belong to the routine panel of laboratory tests in inpatients further helped reduce potential cohort bias. The unfiltered inclusion of all patients' BG measurement data was not possible because of the limitations set by the Swiss Human Research Act.

In total, the patient cohort comprised 38,250 patients (n=16,842, 44.03% women and n=21,408, 55.97% men) who had undergone 63,579 hospital admissions (cases), during which at least 1 laboratory analysis of 52 parameters was performed (Multimedia Appendix 1).

Definition of Dysglycemia

The types of BG decompensation are defined by BG levels of <3.9 mmol/L (<70 mg/dL; hypoglycemia) [26] or >10, >13.9, or >16.7 mmol/L (representing different degrees of hyperglycemia; >180, >250, and >300 mg/dL, respectively) [26,27]. A second level of hypoglycemia (BG <3.0 mmol/L [<54 mg/dL]) [26] was not considered specifically as both BG levels of <3.0 or <3.9 mmol/L may lead to an "altered mental and/or physical status requiring assistance" [26]; that is, a BG

level that is <3.9 mmol/L may already pose a significant risk for a patient.

BG level intervals were assigned category numbers of 0 to 4, with category 0 representing cases not showing BG decompensation (Table 1). These categories were selected on the basis of clinical relevance, intuition in interpretation, and clinicians' decision-making and were aligned with clinical practice guidelines [26,27].

Table 1. Category assignment of blood glucose decompensation types.

Blood glucose level interval (mmol/L)	Category	Decompensation type
≥3.9 to ≤10	0	Nondecompensated
<3.9	1	Hypoglycemia (level 1)
>10 to ≤13.9	2	Mild hyperglycemia
>13.9 to ≤16.7	3	Moderate hyperglycemia
>16.7	4	Severe hyperglycemia

Data Preprocessing

Before use, the data sets were cleaned and preprocessed to remove erroneous entries and unreasonable, unlikely, or even impossible values.

For patient variables, values outside the following (reasonable) limits were set to "N/A" (whereas all other values of the respective patient were retained) as they are most likely the result of errors during data input: ages 18 to 130 years, body height 100 to 250 cm, and weight 25 to 400 kg. For laboratory measurements, negative values were removed (apart from the measurements of base excess, which allowed negative results). Values generally *incompatible with human life* were published only for a relatively small number of analytes [28], and measurements outside the following limits were excluded on the basis of such published limits for chloride (<65 or >138 mmol/L) [29], plasma pH (<6.8 or >7.8) [30,31], potassium (<1.3 or >9.0 mmol/L) [29], and sodium (<100 or >191 mmol/L) [29].

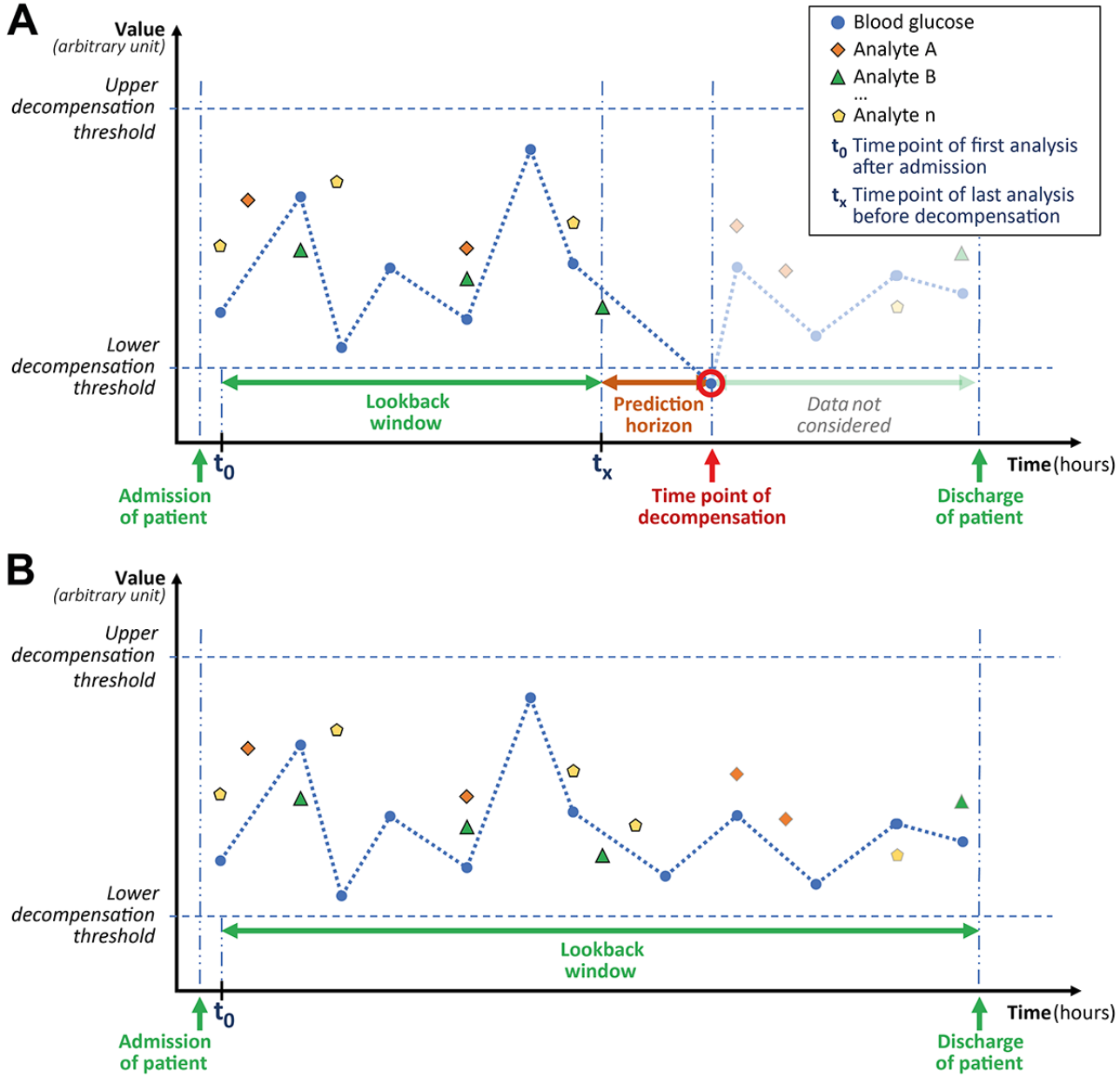
Other putative outliers were identified using the *Isolation Forest* algorithm [32] (implemented in the Python *sklearn.ensemble. IsolationForest* module; version 0.21.3) and were flagged. These binary outlier flags were used as additional variables to label potentially severe cases.

Predictors and Outcomes

Decompensation events corresponding to the different decompensation categories specified in Table 1 were identified in the data for each admission case. If an event corresponding to a decompensation category was present, all data collected before the decompensation event (look-back window; see Figure 1 for a fictive case with [Figure 1A] and without [Figure 1B] a decompensation event) were used to form derived variables for each laboratory analyte (descriptive statistics: mean, SD, IQR, total range, recent trend, most extreme and most recent value, and analysis count). These derived variables are supposed to reflect the current overall status of a patient, mimicking a physician's assessment of the information present in a patient's health record. Only the first decompensation event was considered for each category. If a decompensation event was detected with the first measurements of a case, it was not considered at all in the corresponding decompensation categories. If no decompensation event of a specific category occurred within a case, the data collected at all its time points were used to create the derived variables (Figure 1B). The time span from the last piece of data of at least 1 predictor variable in the electronic health record (EHR) before the decompensation event of interest was considered the prediction horizon for the respective category (Figure 1A).

Analyte statistics were subsequently combined with patient demographics (age, sex, weight, height, language, and civil status), information on the previous and current administration of antidiabetic drugs, previous incidents of BG decompensation, history of diagnoses, and stay in ICU (complete variable list is provided in Multimedia Appendix 2). The data points were hashed, duplicates were removed, and a final data table was generated for the development and evaluation of the prediction models.

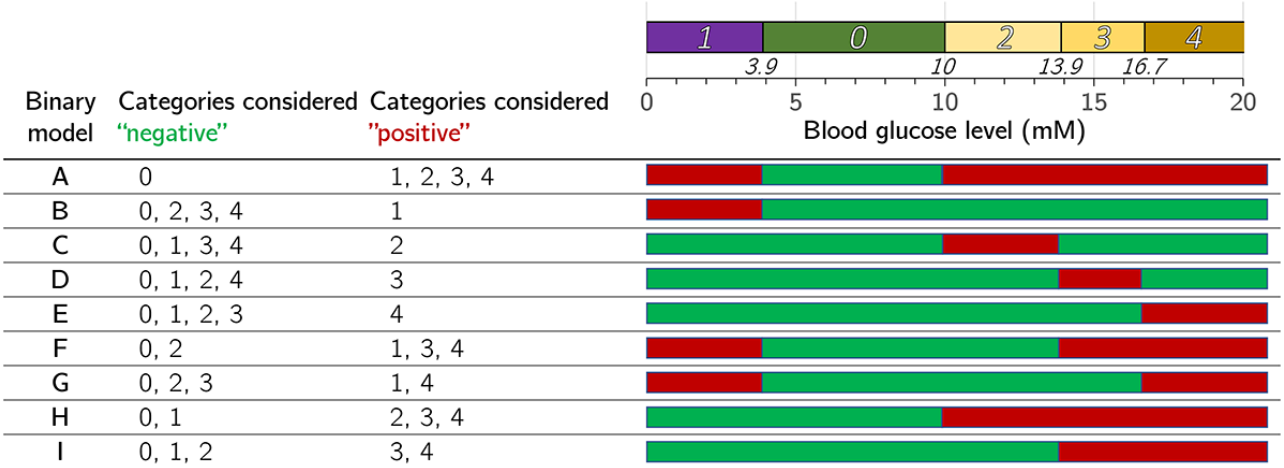
Figure 1. Look-back window and prediction horizon (A: decompensated case; B: nondecompensated case).



Prediction Models and Second-Level Ensemble

Binary prediction models were set up to distinguish clinically relevant combinations of decompensation categories (Figure 2). To this end, decision tree–based classification models using extreme gradient boosting (XGBoost; XGBoost package [33]; version 1.0.0.2) were trained. XGBoost models do not make any assumptions about the distribution of the data and can deal with incomplete data sets. This is particularly useful for clinical data where the time points and types of analyses are specific for each individual patient. XGBoost uses a sparsity-aware split-finding algorithm that assigns a default direction to each node [33]. If a feature value is missing, the default direction along the corresponding node is taken [33,34].

Figure 2. Binary models detecting clinically relevant combinations of blood glucose decompensation types.



Binary model	Categories considered "negative"	Categories considered "positive"
A	0	1, 2, 3, 4
B	0, 2, 3, 4	1
C	0, 1, 3, 4	2
D	0, 1, 2, 4	3
E	0, 1, 2, 3	4
F	0, 2	1, 3, 4
G	0, 2, 3	1, 4
H	0, 1	2, 3, 4
I	0, 1, 2	3, 4

The data of cases falling into several categories were used to train each of the relevant prediction models. For each binary model, cases that did not meet the respective decompensation criterion were used as controls. The resulting categorizations may be puzzling for the human observer at first glance; for example, a binary hypoglycemia model using the data of a hyperglycemic case as *control*.

The model parameters were optimized over a wide parameter space using a stochastic grid search approach covering 2000 to 4000 parameter combinations. Data were split 70/30; that is, 70% of the cases were used for iterative training and the withheld 30% for model testing.

To counter the imbalance in the data sets, each sample was assigned a weight corresponding to the inverse frequency of its class during model training.

The best models with respect to precision (positive predictive value), sensitivity (recall), area under the curve of the receiver operating characteristic curve (AUC ROC), and informedness were retained. The *importance* of each variable (feature importance), which is a score indicating how influential each variable was during the construction of the tree ensemble, was determined during model training [33]. The selection of variables in advance was not required for the XGBoost models. Feature importance was calculated for each separate decision tree and node as the amount by which splitting on that particular node reduces the loss function, weighted by the number of leaf assignments to which the node contributes. Therefore, features affecting more decisions and decisions with higher significance will have higher relative importance. For the final assignment, the importance was averaged across the entire ensemble for each variable [33,35].

Subsequently, multiple binary models were assembled into a second-level ensemble to build a multiclass classifier. The composition of this second-level ensemble was optimized for average precision using a genetic algorithm [36,37], as the number of possibilities for combining binary models into an ensemble grows exponentially.

Each multiclass is described by a (theoretical) combination of predictions of all binary classifiers (class-specific bit strings, ie, *ideal results*). For multiclass classification, a case-specific bit string for the votes of all binary classifiers of the second-level ensemble was created. The label of the *ideal* class-specific bit string with the minimal Hamming distance to the predicted case-specific bit string was chosen as the predicted case label.

In the case of a tie between ≥2 classes, they were prioritized according to clinical relevance (severity): hypoglycemia>severe hyperglycemia>moderate hyperglycemia>nondecompensated case>mild hyperglycemia. Mild hyperglycemia was assigned the lowest priority to reduce false alarms.

Ethics Approval

This study was approved by the Bernese cantonal ethics committee (Kantonale Ethikkommission, KEK) and registered with the Business Administration System for Ethics Committees of the canton of Bern (KEK/NZE file number: Req-2018-00335).

Results

Patient Cohort Characteristics

For this study, we analyzed clinical data from 38,250 inpatients who had undergone 63,579 hospital admissions between January 1, 2014, and December 5, 2019, during which at least 1 laboratory analysis was performed. Of the 38,250 patients, 16,842 (44.03%) were women (age: mean 62.6, SD 19.6 years), and 21,408 (55.97%) were men (age: mean 65.9, SD 14.4 years). Specific details for the different categories of dysglycemia (based on clinical relevance; compare with the *Methods* section) are summarized in Table 2.

Table 2. Patient cohort characteristics at baseline.

	Category				
	0: nondecompensated	1: hypoglycemia	2: mild hyperglycemia	3: moderate hyperglycemia	4: severe hyperglycemia
Characteristics					
Number of cases[a]	24,330	8164	26,788	10,419	8484
Prevalence per admission, %	38.3	12.8	42.1	16.4	13.3
Gender, n (%)					
Female	10,820 (44.5)	3911 (47.9)	9981 (36.9)	3687 (35.4)	2989 (35.2)
Male	13,510 (55.5)	4253 (52.1)	16,897 63.1)	6732 (64.6)	5495 (64.8)
Age (years), median (IQR)	69 (56-78)	65 (48-75)	69 (60-77)	68 (60-77)	68 (58-76)
Height (cm), median (IQR)	169 (163-175)	168 (162-175)	170 (163-176)	170 (163-176)	170 (163-176)
Weight (kg), median (IQR)	81.9 (69.5-96.5)	71.0 (60.3-83.4)	79.1 (67.9-91.9)	79.6 (68.1-92.4)	78.8 (67.0-91.0)
BMI, median (IQR)	28.4 (24.7-33.4)	25.0 (21.7-29.0)	27.3 (24.0-31.4)	27.5 (24.1-31.7)	27.3 (23.8-31.5)
Cases in the ICU[b], n (%)	1381 (5.7)	2062 (25.3)	7694 (28.7)	3171 (30.4)	2523 (29.7)
Length of stay (days), median (IQR)	3.8 (1.9-7.6)	7.6 (3.6-14.8)	6.8 (3.0-12.1)	7.1 (3.2-13.1)	6.9 (3.1-13.1)
Previous diagnosis of diabetes, n (%)	9529 (39.2)	2449 (30)	9860 (36.8)	4568 (43.8)	3985 (47)
Antidiabetic drugs (administered before hospital admission), n (%)	8046 (33.1)	1975 (24.2)	8396 (31.3)	3829 (36.8)	3291 (38.8)
Previous BG[c] decompensation, n (%)	6801 (28)	4238 (51.9)	10,820 (40.4)	7885 (75.7)	6415 (75.6)
Decompensation level (mM), median (IQR)	N/A[d]	3.6 (3.2-3.7)	11.4 (10.6-12.3)	14.9 (14.4-15.7)	18.6 (17.4-20.8)
Time point of decompensation (hours; time after hospital admission), median (IQR)	N/A	34 (4-115)	16 (1-43)	25 (3-70)	22 (1-74)

[a]Numbers sum up beyond the total number of hospital admissions (63,579), as cases occasionally fell into multiple categories if different decompensation events occurred within a case (Figure 3).

[b]ICU: intensive care unit.

[c]BG: blood glucose.

[d]N/A: not applicable.

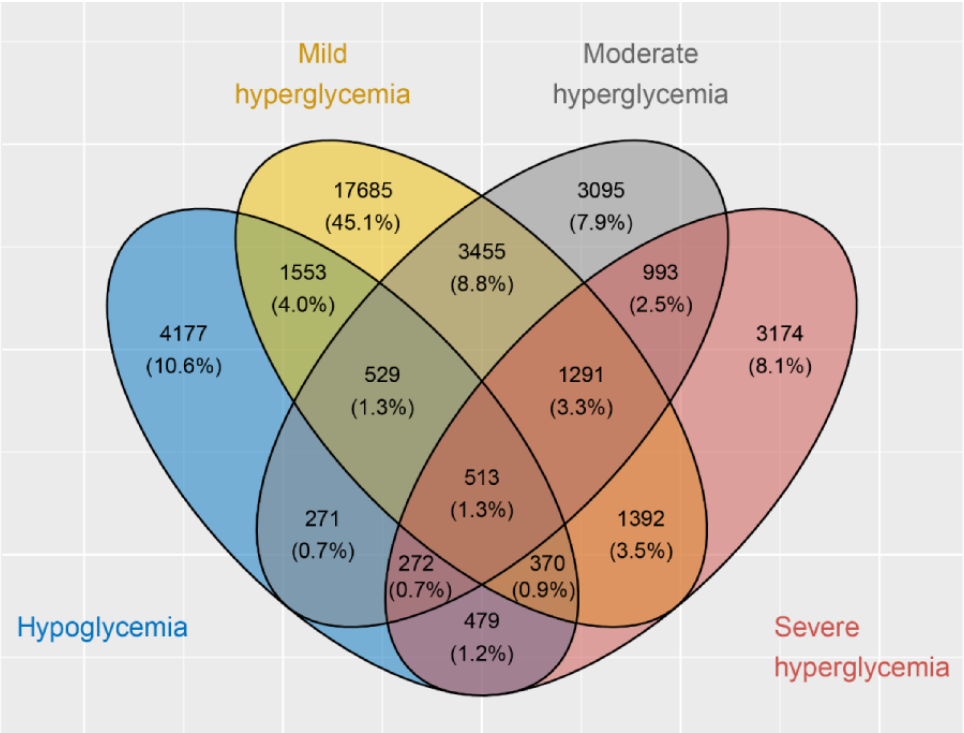
Men represented approximately two-thirds of the patients showing a hyperglycemic decompensation event (Table 2), whereas nondecompensated and hypoglycemic cases mirrored the overall gender distribution of the cohort.

The patients showing hypoglycemia included significantly younger patients compared with other categories, and similarly, patients presenting with hypoglycemia were less obese (ANOVA [type II] using subjects as blocking factors, followed by the Games-Howell post hoc test 95% confidence level; $P<.001$). The likelihood that patients with some form of BG decompensation were admitted to the ICU increased approximately 5-fold.

An overview of the baseline characteristics of the patient cohort is provided in Table 2.

In a subset of cases, the patients showed multiple episodes or categories of decompensation events during admission (Figure 3). This included opposing decompensation types (hypo- and hyperglycemia) during the same admission in approximately 10.16% (3987/39,249) of decompensated cases. The inclusion criteria of the cohort favored a high prevalence of decompensation events (eg, patients with a formal diagnosis of diabetes or diabetes-associated comorbidities); hence, the prevalence of cases without any kind of BG decompensation was appreciably low (Table 2).

Figure 3. Types of decompensation events.



Multiclass Classifier Performance

Several binary models were assembled into a second-level ensemble to create a multiclass classifier. For k classes, a minimum of log_2k binary classifiers are required to represent all classes, (ie, a minimum of 3 classifiers for 5 classes). Adding additional binary classifiers (beyond log_2k) is, in principle, redundant; however, additional bits can act as error-correcting codes (eg, the study by Berger [38]).

As the number of potential ensembles built from binary models grows exponentially, their composition was optimized using a genetic algorithm [36,37], assessing different combinations of the best binary models with respect to precision, sensitivity (recall), AUC ROC, and informedness. Average precision served as a readout (*fitness value*) during the optimization process.

The multiclass classifier ensemble with the best performance comprised binary model types A, B, G, H, and I (compare Figure 2; ie, of models trained to recognize all types of decompensations, only hypoglycemia, hypoglycemia and severe hyperglycemia, all types of hyperglycemia, and moderate and severe hyperglycemia, respectively). The mean AUC ROCs of the contributing binary models were 0.925 (SD 0.001), 0.960 (SD 0.002), 0.867 (SD 0.002), 0.914 (SD 0.001), and 0.863 (SD 0.003), respectively (5-fold cross-validation).

An overview of the true decompensation categories versus the model predictions (confusion matrix) is shown in Figure 4. Our multiclass classifier correctly predicted nondecompensated, hypoglycemic, and hyperglycemic cases in 93.66% (28,042/29,941), 58.99% (1093/1853), and 63.56% (6240/9817) of cases, respectively, in relation to the true category.

The performance metrics for each class are summarized in Table 3. All types of dysglycemia were predicted reasonably but somewhat conservatively.

Figure 4. Classification by multiclass classifier for nondecompensated cases (control), hypoglycemia, and hyperglycemia (confusion matrix; percentages relative to true class; number of cases in parentheses).

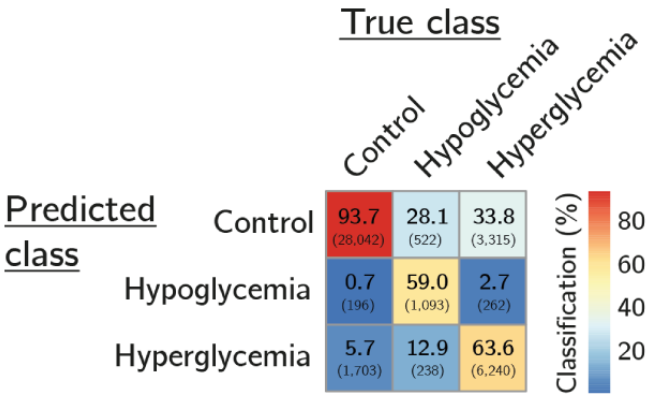


Table 3. Performance of multiclass classifier by class (5-fold cross-validation).

Performance metric	Performance (%), mean (SD)		
	Control	Hypoglycemia	Hyperglycemia
Sensitivity	67.1 (2.3)	59.0 (5.7)	63.6 (1.1)
Specificity	93.7 (0.9)	98.8 (0.5)	93.9 (0.6)
Precision	80.6 (1.9)	71.8 (8.6)	76.3 (1.6)
Balanced accuracy	80.4 (0.7)	78.9 (2.6)	78.7 (0.4)

When the different types of hyperglycemia were considered separately, mild, moderate, and severe hyperglycemia were correctly predicted in 46.69% (2411/5164), 31.22% (797/2553), and 36.87% (774/2099) of the cases in relation to the true category, respectively (Figure 5). For each type of hyperglycemia (mild, moderate, and severe), an additional 20% to 30% of the classifications fell into other categories of hyperglycemia, meaning that the overall state of hyperglycemia was detected in those cases as well.

Figure 5. Classification by multiclass classifier (confusion matrix; percentages relative to true class; case numbers in parentheses).

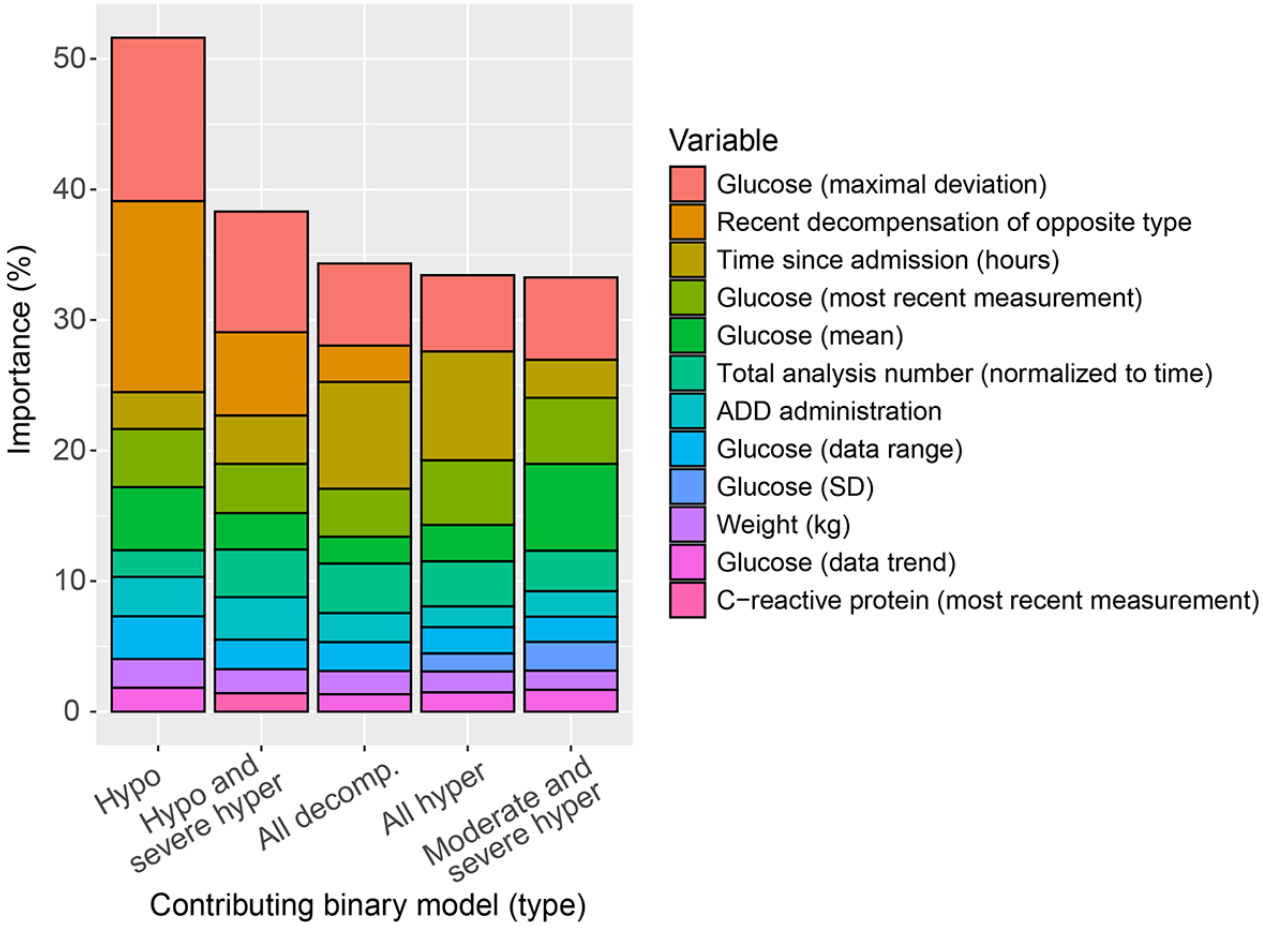
Nondecompensated cases or cases showing hypoglycemia and severe hyperglycemia were predicted reasonably, albeit fairly conservatively, with lower sensitivity. The correct distinction between hyperglycemia types was more challenging. From a clinical perspective, this is understandable as these cases form a continuum. Nevertheless, a warning of any kind of hyperglycemia makes no difference in terms of the measures taken; hence, a classification of hyperglycemia was assumed to be sufficient, as long as this is indeed the underlying condition.

The median look-back window was 4.4 (IQR 1.8-9) days, 1.9 (IQR 0.4-5.7) days, 1.0 (IQR 0.3-2.4) days, 1.5 (IQR 0.6-3.7) days, and 1.8 (IQR 0.7-4.4) days for control class, hypoglycemia, and mild, moderate, and severe hyperglycemia, respectively. The median prediction horizon was 7 (IQR 3-15) hours for hypoglycemia and 4 (IQR 3-7, 3-6, and 3-6 for mild, moderate, and severe hyperglycemia, respectively) hours for all hyperglycemia types. In a clinical setting, such a prediction horizon would allow sufficient time for proactive interventions before BG decompensation occurs.

If the multiclass classifier is converted (reduced) to a binary classifier, it correctly predicts 67.12% (7833/11,670) of decompensated and 93.66% (28,042/29,941) of nondecompensated cases, corresponding to sensitivity (recall) and specificity (selectivity), respectively. In terms of precision (positive predictive value) and balanced accuracy, the model achieved 80.5% and 80.5%, respectively.

Notably, across all predictions, a relatively small set of variables had a comparably large impact on the individual binary models constituting the multiclass classifier (Figure 6). Although the contributions of the variables (feature importance; see the *Methods* section for details) varied slightly across models, they all pointed in the same direction. A set of 12 variables encompassed the 10 most important variables for all 5 binary models. Not unexpectedly, quantitative readouts of glucose measurements had the highest importance; however, *soft* variables, such as information on recent decompensation events or time since admission, were also critical.

Figure 6. Variable importance (population level) of binary models comprising the multiclass classifier (top 10 shown). ADD: antidiabetic drug; all: all types of; decomp: decompensation; hyper: hyperglycemia; hypo: hypoglycemia.



Discussion

Principal Findings

In this study, we set up a prediction model for hypo- and hyperglycemia from routine inpatient clinical data. We used anonymized EHRs of 63,579 hospital admissions of 38,250 patients between January 2014 and December 2019 to derive variables from measurements of common laboratory analytes, as well as patient information, drug administration, and diagnosis history. The data were characterized by a high degree of missingness, particularly for many of the variables originating from less common laboratory tests. To address this issue, we created a second-level ensemble comprising binary decision tree models that could deal with incomplete data. With respect to the overall main categories of interest (nondecompensated cases and hypo- or hyperglycemia), our classifier achieved 93.7% (28,042/29,941), 59% (1093/1853), and 63.6% (6240/9817) correct predictions, respectively. When converted to a binary classifier, it reached correct classification rates of 93.7% (28,042/29,941) and 67.1% (7833/11,670) for decompensated and nondecompensated cases, respectively. The median prediction horizon was 7 and 4 hours for hypo- and hyperglycemia, respectively.

Strengths and Limitations

In our modeling approach, the variables were derived rather than directly taken from the EHRs. This increases the effort in an applied clinical setting as the variables need to be updated when new measurements are added to a patient's EHR. This concerns both the summary statistics of each analyte and variables related to patient history (eg, application of antidiabetic drugs or previous decompensation events), which may change during hospitalization. However, an automatic update from clinical data warehouses should solve this issue. The ultimate aim should be to allow integration into a real-time alert system.

Data were not specifically collected for our retrospective study; hence, our data set may show a potential bias in terms of the timing of blood sampling; for example, hours during the day with higher staffing levels or before meals may be somewhat overrepresented. Hypoglycemia, in particular, may occur at night or in the early morning hours when no measurements are performed. Despite this potential bias in the training data, our model featured a promising prediction horizon. With 7 and 4 hours for hypo- and hyperglycemia, respectively, it provided sufficient time to initiate measures and prevent BG decompensation and its adverse effects. The difference between hypo- and hyperglycemia may be explained by more deliberate monitoring of BG levels in potential hypoglycemic patients by clinical staff.

The main strength of our model's performance lies in the distinction between the extreme ends of the decompensation scale (nondecompensated cases, hypoglycemia, and severe hyperglycemia). However, the discrimination rate of different

types of hyperglycemia, was rather moderate, particularly the rate of correct classification of moderate hyperglycemia (31.2%). In part, this may be because the categorization of different types of hyperglycemia, and the overall distinction into 5 categories in our model, had a clinical relevance rather than a computational one. Different types of hyperglycemia are defined by clinical practice guidelines [26,27] and form a continuum in the daily routine.

This does not affect the model's usefulness drastically when taking into account the intention behind a model like ours, namely serving as decision support for adjustments of patient monitoring or therapeutic strategies (not triggering any treatment such as administration of insulin or glucose). Regarding the potential consequences of misclassification for a patient, mix-ups of hyperglycemia types are clinically irrelevant, and similarly, confusion of hypoglycemia and hyperglycemia has no drastic consequences—they would all *trigger* a BG test by clinical staff. False positives cause an extra workload for clinical staff but are also noncritical for a patient and occurred at a low rate (6.3%). False-negative predictions are more problematic as they may possibly result in the noninitiation of a BG test or countermeasures for BG decompensation. The false-negative rate of our model (32.9%) is suboptimal, meaning that it classifies too conservatively and misses too many decompensated cases. However, in light of these cases being missed completely without decision support, this is a step in the right direction.

Despite having a common denominator—diabetes or diabetes-related comorbidities—patients in our cohort were diagnosed with various primary and secondary diagnoses. Consequently, a plethora of different analyses have been conducted, with some tests done routinely; for example, the assessment of blood count, and other, more specific ones, were missing in most cases, such as analyses related to iron metabolism. Our prediction model, a second-level ensemble using derived variables, deals well with this sparse data (shown in the *Methods* section), in contrast to other popular methods such as logistic regression, support vector machines, and neural networks, which require complete data sets.

Akin to other studies [22], the patient cohort underlying our model features a certain bias toward patients with potential indications of dysglycemia (see the *Methods* section). However, this is not necessarily disadvantageous. A cohort without preselection may lead to trivial models calling for the (nondecompensated) group with an overwhelming frequency, yielding a low overall classification error but limited clinical utility [39]. The actual performance, clinical applicability, and performance for specific patient subgroups (eg, dysglycemia risk patients) must be assessed in broader follow-up studies, as for every model.

Notably, our model uses routine clinical data from standard laboratory tests and patient information, all of which are readily available in hospital settings. This is beneficial as these variables—assuming a proper in-house information flow—are available for free, both in terms of cost (no additional testing required) and effort (once the information is entered). However, such data collected during routine management usually do not feature fixed sampling times or a common set of analyses, which complicates alignment between cases. Given the absence of periodicity in the data, our approach was to mimic a physician's intuitive assessment of the information present across a patient's health record (eg, average and extreme values, trends, and most recent values) rather than individual data points.

This could ensure broad applicability, for example, in contrast to models based on expensive specialized tests [22] or even data from sensor implants [24]. Without a doubt, CGM devices offering frequent subcutaneous glucose measurements are ideal for a narrow set of patients, such as those with type 1 diabetes. However, their wider use in the general context of inpatients is neither practical nor cost-effective. In times of tight budgets for health care systems, the frugality of a model relying only on pre-existing data is of additional value. Our model does not require data sampled at specific time points; it takes whatever data are available. The identification of high-risk patients at no extra cost may lead to a reduction in workload for clinical staff and less frequent blood sampling for average patients. Overalerting should be taken into account as well—it would reduce the benefit of correct predictions; hence, the rather conservative nature of our multiclass classifier may actually not be such a drawback. A model is always an approximation of reality and should serve its purpose in the first place ("All models are wrong but some are useful." [40]).

Comparison With Prior Work

An increasing number of prediction studies are using the growing amount of patient data available in EHRs (for review, see the studies by Woldaregay et al [20], Roca et al [41], and Torkamani et al [42]). Diabetes is a popular and promising target of prediction studies [20] because of both its high prevalence in an aging population and the associated economic burden. The growing number of inpatients with deficiencies in controlling their BG levels calls for a refinement of the existing inpatient dysglycemia management [12-14,43].

Models for predicting a single type of BG decompensation as diabetes-associated complications have been established previously, mainly for hypoglycemia [20-22,44,45]. Few studies only have been published for hyperglycemic events [20].

To the best of our knowledge, this is the first general multiclass prediction model for BG decompensation to date (ie, both hypo- and hyperglycemia). A multiclass model offers the advantage of overcoming the vagueness associated with a binary model in a setting of >2 classes. For example, the prediction of *nonhypoglycemic* by a hypoglycemia model is ambiguous as the affected patient could be *nondecompensated* or *hyperglycemic*, where the latter would require action by the patient or clinical staff. A multiclass model such as ours resolves this conflict by differentiating between nondecompensated and multiple types of decompensated cases.

Notably, our model performed reasonably well when reverted to a binary model. Although it may be counterintuitive to first build a multiclass classifier and then revert it to a binary classifier, this approach takes advantage of the fact that ensembles tend to outperform individual models [46] and can benefit from error correction [38].

Future Directions

In the next step, it would be interesting to assess the multiclass classifier in a follow-up retrospective cohort study for validation purposes using an independent data set. Optimizing the sensitivity and reducing the false-negative rate should be an additional focus to make the model more applicable for clinical use. Further down the road, the incorporation of the model into an alert system or even actionable artificial intelligence [17-19] could be tested. This would allow the evaluation of its real-time effectiveness, ideally leading to a reduction in the incidence of BG decompensations in inpatients.

Conclusions

Our multiclass prediction model based on derived variables can classify both hypo- and hyperglycemia with reasonable sensitivity.

Given the serious adverse health effects of hypo- and hyperglycemia and the associated poor outcome of BG decompensation in inpatients, it is important to prevent dysglycemia whenever possible. Prediction models such as ours may support clinicians in inpatient management by proactively pointing out the necessity for adjustment of patient monitoring or therapeutic strategies. Therefore, this study may serve as a step toward a real-time alarm system and actionable artificial intelligence, which may aid in reducing BG decompensation in inpatients.

Acknowledgments

The authors are indebted to all patients who gave general consent for using their data for research; the Bernese Cantonal Ethics Committee (*Kantonale Ethikkommission*) for approving this study; David Herzig for discussions; Tobias Blatter for input on data processing; and the team at Insel Data Science Center, in particular Jeremy Koch and Noel Frey, for extracting the data set.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Descriptive statistics of laboratory measurements used in this study.
[PDF File (Adobe PDF File), 165 KB - formative_v6i7e36176_app1.pdf]

Multimedia Appendix 2
List of variables used for model setup.
[PDF File (Adobe PDF File), 149 KB - formative_v6i7e36176_app2.pdf]

References

1. IDF diabetes atlas. 10th edition. International Diabetes Federation. Brussels, Belgium; 2021. URL: https://diabetesatlas.org/atlas/tenth-edition/ [accessed 2022-05-28]
2. Diabetes prevalence in the Swiss population in 2017 (Swiss Health Survey 2017). Bundesamt für Statistik. 2018. URL: https://www.bfs.admin.ch/bfs/de/home/statistiken/gesundheit/gesundheitszustand/krankheiten/diabetes.html [accessed 2021-11-24]
3. National Diabetes Inpatient Audit (NaDIA) - 2019. National Health Service Digital. 2020 Nov 13. URL: https://digital.nhs.uk/data-and-information/publications/statistical/national-diabetes-inpatient-audit/2019 [accessed 2021-12-10]
4. National Diabetes Inpatient Audit (NaDIA) - 2017. National Health Service Digital. 2018 Mar 13. URL: https://digital.nhs.uk/data-and-information/publications/statistical/national-diabetes-inpatient-audit/national-diabetes-inpatient-audit-nadia-2017 [accessed 2021-12-10]
5. McAlister FA, Majumdar SR, Blitz S, Rowe BH, Romney J, Marrie TJ. The relation between hyperglycemia and outcomes in 2,471 patients admitted to the hospital with community-acquired pneumonia. Diabetes Care 2005 Apr;28(4):810-815. [doi: 10.2337/diacare.28.4.810] [Medline: 15793178]
6. Umpierrez G, Korytkowski M. Diabetic emergencies - ketoacidosis, hyperglycaemic hyperosmolar state and hypoglycaemia. Nat Rev Endocrinol 2016 Apr;12(4):222-232. [doi: 10.1038/nrendo.2016.15] [Medline: 26893262]
7. Baker EH, Janaway CH, Philips BJ, Brennan AL, Baines DL, Wood DM, et al. Hyperglycaemia is associated with poor outcomes in patients admitted to hospital with acute exacerbations of chronic obstructive pulmonary disease. Thorax 2006 Apr;61(4):284-289 [FREE Full text] [doi: 10.1136/thx.2005.051029] [Medline: 16449265]
8. Pomposelli JJ, Baxter JK, Babineau TJ, Pomfret EA, Driscoll DF, Forse RA, et al. Early postoperative glucose control predicts nosocomial infection rate in diabetic patients. JPEN J Parenter Enteral Nutr 1998;22(2):77-81. [doi: 10.1177/014860719802200277] [Medline: 9527963]
9. McDonnell ME, Umpierrez GE. Insulin therapy for the management of hyperglycemia in hospitalized patients. Endocrinol Metab Clin North Am 2012 Mar;41(1):175-201 [FREE Full text] [doi: 10.1016/j.ecl.2012.01.001] [Medline: 22575413]

10. Umpierrez GE, Isaacs SD, Bazargan N, You X, Thaler LM, Kitabchi AE. Hyperglycemia: an independent marker of in-hospital mortality in patients with undiagnosed diabetes. J Clin Endocrinol Metab 2002 Mar;87(3):978-982. [doi: 10.1210/jcem.87.3.8341] [Medline: 11889147]
11. Ainla T, Baburin A, Teesalu R, Rahu M. The association between hyperglycaemia on admission and 180-day mortality in acute myocardial infarction patients with and without diabetes. Diabet Med 2005 Oct;22(10):1321-1325. [doi: 10.1111/j.1464-5491.2005.01625.x] [Medline: 16176190]
12. Seaquist ER, Anderson J, Childs B, Cryer P, Dagogo-Jack S, Fish L, et al. Hypoglycemia and diabetes: a report of a workgroup of the American Diabetes Association and the Endocrine Society. Diabetes Care 2013 May;36(5):1384-1395 [FREE Full text] [doi: 10.2337/dc12-2480] [Medline: 23589542]
13. Bogun M, Inzucchi SE. Inpatient management of diabetes and hyperglycemia. Clin Ther 2013 May;35(5):724-733. [doi: 10.1016/j.clinthera.2013.04.008] [Medline: 23688537]
14. Pratiwi C, Mokoagow MI, Made Kshanti IA, Soewondo P. The risk factors of inpatient hypoglycemia: a systematic review. Heliyon 2020 May;6(5):e03913 [FREE Full text] [doi: 10.1016/j.heliyon.2020.e03913] [Medline: 32420485]
15. Khazai NB, Hamdy O. Inpatient diabetes management in the twenty-first century. Endocrinol Metab Clin North Am 2016 Dec;45(4):875-894. [doi: 10.1016/j.ecl.2016.06.013] [Medline: 27823609]
16. Colunga-Lozano LE, Gonzalez Torres FJ, Delgado-Figueroa N, Gonzalez-Padilla DA, Hernandez AV, Roman Y, et al. Sliding scale insulin for non-critically ill hospitalised adults with diabetes mellitus. Cochrane Database Syst Rev 2018 Nov 29;11(11):CD011296 [FREE Full text] [doi: 10.1002/14651858.CD011296.pub2] [Medline: 30488948]
17. Boughton CK, Bally L, Martignoni F, Hartnell S, Herzig D, Vogt A, et al. Fully closed-loop insulin delivery in inpatients receiving nutritional support: a two-centre, open-label, randomised controlled trial. Lancet Diabetes Endocrinol 2019 May;7(5):368-377 [FREE Full text] [doi: 10.1016/S2213-8587(19)30061-0] [Medline: 30935872]
18. Leelarathna L, Choudhary P, Wilmot EG, Lumb A, Street T, Kar P, et al. Hybrid closed-loop therapy: where are we in 2021? Diabetes Obes Metab 2021 Mar;23(3):655-660. [doi: 10.1111/dom.14273] [Medline: 33269551]
19. Boughton CK, Hovorka R. The artificial pancreas. Curr Opin Organ Transplant 2020 Aug;25(4):336-342. [doi: 10.1097/MOT.0000000000000786] [Medline: 32618719]
20. Woldaregay AZ, Årsand E, Walderhaug S, Albers D, Mamykina L, Botsis T, et al. Data-driven modeling and prediction of blood glucose dynamics: machine learning applications in type 1 diabetes. Artif Intell Med 2019 Jul;98:109-134 [FREE Full text] [doi: 10.1016/j.artmed.2019.07.007] [Medline: 31383477]
21. Mathioudakis NN, Everett E, Routh S, Pronovost PJ, Yeh HC, Golden SH, et al. Development and validation of a prediction model for insulin-associated hypoglycemia in non-critically ill hospitalized adults. BMJ Open Diabetes Res Care 2018 Mar 2;6(1):e000499 [FREE Full text] [doi: 10.1136/bmjdrc-2017-000499] [Medline: 29527311]
22. Ruan Y, Bellot A, Moysova Z, Tan GD, Lumb A, Davies J, et al. Predicting the risk of inpatient hypoglycemia with machine learning using electronic health records. Diabetes Care 2020 Jul;43(7):1504-1511. [doi: 10.2337/dc19-1743] [Medline: 32350021]
23. Porumb M, Stranges S, Pescapè A, Pecchia L. Precision medicine and artificial intelligence: a pilot study on deep learning for hypoglycemic events detection based on ECG. Sci Rep 2020 Jan 13;10(1):170 [FREE Full text] [doi: 10.1038/s41598-019-56927-5] [Medline: 31932608]
24. Daskalaki E, Nørgaard K, Züger T, Prountzou A, Diem P, Mougiakakou S. An early warning system for hypoglycemic/hyperglycemic events based on fusion of adaptive prediction models. J Diabetes Sci Technol 2013 May 01;7(3):689-698 [FREE Full text] [doi: 10.1177/193229681300700314] [Medline: 23759402]
25. International Statistical Classification of Diseases and Related Health Problems. 10th Revision. World Health Organization. 2019. URL: https://icd.who.int/browse10/2019/en#/ [accessed 2019-12-02]
26. Agiostratidou G, Anhalt H, Ball D, Blonde L, Gourgari E, Harriman KN, et al. Standardizing clinically meaningful outcome measures beyond HbA1c for type 1 diabetes: a consensus report of the American Association of Clinical Endocrinologists, the American Association of Diabetes Educators, the American Diabetes Association, the Endocrine Society, JDRF International, The Leona M. and Harry B. Helmsley Charitable Trust, the Pediatric Endocrine Society, and the T1D Exchange. Diabetes Care 2017 Dec;40(12):1622-1630 [FREE Full text] [doi: 10.2337/dc17-1624] [Medline: 29162582]
27. American Diabetes Association. Hospital admission guidelines for diabetes. Diabetes Care 2004 Jan;27 Suppl 1:S103. [doi: 10.2337/diacare.27.2007.s103] [Medline: 14693939]
28. Clot-Silla E, Argudo-Ramirez A, Fuentes-Arderiu X. Letter to the editor: measured values incompatible with human life. EJIFCC 2011 Jul;22(2):52-54 [FREE Full text] [Medline: 27683391]
29. Vogt W, Oesterle B. [Extreme results in electrolyte determination]. Wien Klin Wochenschr Suppl 1992;192:21-27. [Medline: 1502820]
30. Cuhaci B, Lee J, Ahmed Z. Sodium bicarbonate controversy in lactic acidosis. Chest 2000 Sep;118(3):882-884. [doi: 10.1378/chest.118.3.882] [Medline: 10988226]
31. Rosival V. Evaluating sodium bicarbonate controversy. Chest 2001 May;119(5):1622-1623. [doi: 10.1378/chest.119.5.1622] [Medline: 11348986]
32. Liu FT, Ting KM, Zhou ZH. Isolation forest. In: Proceedings of the 8th IEEE International Conference on Data Mining. 2008 Presented at: ICDM '08; December 15-19, 2008; Pisa, Italy p. 413-422. [doi: 10.1109/icdm.2008.17]

33. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016 Aug 13 Presented at: KDD '16; August 13-17, 2016; San Francisco, CA, USA p. 785-794 URL: https://dl.acm.org/doi/10.1145/2939672.2939785 [doi: 10.1145/2939672.2939785]
34. Belloni M. XGBoost is not black magic. Towards Data Science. 2018 Dec 13. URL: https://towardsdatascience.com/xgboost-is-not-black-magic-56ca013144b4 [accessed 2020-06-10]
35. Brownlee J. Feature importance and feature selection with XGBoost in Python. Machine Learning Mastery. 2016 Aug 31. URL: https://machinelearningmastery.com/feature-importance-and-feature-selection-with-xgboost-in-python/ [accessed 2020-06-08]
36. Scrucca L. GA: a package for genetic algorithms in R. J Stat Soft 2013 Apr;53(4):1-37 [FREE Full text] [doi: 10.18637/jss.v053.i04]
37. Scrucca L. On some extensions to GA package: hybrid optimisation, parallelisation and islands EvolutionOn some extensions to GA package: hybrid optimisation, parallelisation and islands evolution. R J 2017;9(1):187-206 [FREE Full text] [doi: 10.32614/rj-2017-008]
38. Berger A. Error-correcting output coding for text classification. In: Proceedings of IJCAI-99 Workshop on Machine Learning for Information Filtering. 1999 Presented at: IJCAI '99; July 31-August2, 1999; Pittsburgh, PA, USA URL: https://www.cs.cmu.edu/~aberger/pdf/ecoc.pdf
39. Henley SS, Golden RM, Kashner TM. Statistical modeling methods: challenges and strategies. Biostat Epidemiol 2020;4(1):105-139. [doi: 10.1080/24709360.2019.1618653]
40. Box GEP. Robustness in the strategy of scientific model building. In: Launer RL, Wilkinson GN, editors. Robustness in Statistics. Madison, Wisconsin: Academic Press; 1979:201-236. [doi: 10.1016/B978-0-12-438150-6.50018-2]
41. Roca J, Tenyi A, Cano I. Paradigm changes for diagnosis: using big data for prediction. Clin Chem Lab Med 2019 Feb 25;57(3):317-327 [FREE Full text] [doi: 10.1515/cclm-2018-0971] [Medline: 30530879]
42. Torkamani A, Andersen KG, Steinhubl SR, Topol EJ. High-definition medicine. Cell 2017 Aug 24;170(5):828-843 [FREE Full text] [doi: 10.1016/j.cell.2017.08.007] [Medline: 28841416]
43. Clement S, Braithwaite SS, Magee MF, Ahmann A, Smith EP, Schafer RG, American Diabetes Association Diabetes in Hospitals Writing Committee. Management of diabetes and hyperglycemia in hospitals. Diabetes Care 2004 Feb;27(2):553-591. [doi: 10.2337/diacare.27.2.553] [Medline: 14747243]
44. Dagliati A, Marini S, Sacchi L, Cogni G, Teliti M, Tibollo V, et al. Machine learning methods to predict diabetes complications. J Diabetes Sci Technol 2018 Mar;12(2):295-302 [FREE Full text] [doi: 10.1177/1932296817706375] [Medline: 28494618]
45. Mathioudakis NN, Abusamaan MS, Shakarchi AF, Sokolinsky S, Fayzullin S, McGready J, et al. Development and validation of a machine learning model to predict near-term risk of iatrogenic hypoglycemia in hospitalized patients. JAMA Netw Open 2021 Jan 04;4(1):e2030913 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.30913] [Medline: 33416883]
46. Hashem S. Optimal linear combinations of neural networks. Neural Netw 1997 Jun;10(4):599-614. [doi: 10.1016/s0893-6080(96)00098-6] [Medline: 12662858]

Abbreviations

AUC ROC: area under the curve of the receiver operating characteristic curve
BG: blood glucose
CGM: continuous glucose monitoring
EHR: electronic health record
ICU: intensive care unit
XGBoost: extreme gradient boosting

Edited by A Mavragani; submitted 04.01.22; peer-reviewed by HH Kim, N Mathioudakis; comments to author 28.02.22; revised version received 25.04.22; accepted 08.05.22; published 18.07.22.

Please cite as:
Witte H, Nakas C, Bally L, Leichtle AB
Machine Learning Prediction of Hypoglycemia and Hyperglycemia From Electronic Health Records: Algorithm Development and Validation
JMIR Form Res 2022;6(7):e36176
URL: https://formative.jmir.org/2022/7/e36176
doi:10.2196/36176
PMID:35526139

Original Paper

A Web-Based Stress Management Intervention for University Students in Indonesia (Rileks): Feasibility Study Using a Pretest-Posttest Design

Dilfa Juniar[1,2], MPsi; Wouter van Ballegooijen[2,3,4], PhD; Mieke Schulte[2], PhD; Anneke van Schaik[3,4], MD, PhD; Jan Passchier[2], PhD; Elena Heber[5], PhD; Dirk Lehr[6], PhD; Sawitri Supardi Sadarjoen[1], Dr; Heleen Riper[2], PhD

[1]Faculty of Psychology, YARSI University, Jakarta, Indonesia

[2]Department of Clinical, Neuro and Developmental Psychology, Vrije Universiteit, Amsterdam, Netherlands

[3]Department of Psychiatry, Amsterdam University Medical Centers, Amsterdam, Netherlands

[4]Geestelijke Gezondheids Zorg inGeest Specialized Mental Health Care, Amsterdam, Netherlands

[5]HelloBetter – GET.ON Institut für Online Gesundheitstrainings GmbH, Hamburg, Germany

[6]Institute of Psychology, Leuphana University, Lüneburg, Germany

Corresponding Author:
Dilfa Juniar, MPsi
Faculty of Psychology
YARSI University
YARSI Tower 6th Floor
Jl. Letjend Soeprapto
Jakarta, 10510
Indonesia
Phone: 62 21 4223138
Fax: 62 21 4223138
Email: dilfa.juniar@yarsi.ac.id

Abstract

Background: University students are susceptible to excessive stress. A web-based stress management intervention holds promise to improve stress but is still at a novel stage in Indonesia.

Objective: The aim of this paper was to report the feasibility of the intervention we developed—Rileks—among university students in Indonesia in terms of acceptability and usability, and to propose recommendations for future improvements.

Methods: A single-group pretest and posttest design was used. Participants with scores of 15 or higher on the stress subscale of the 42-item Depression Anxiety Stress Scales were given access to the intervention (N=68). The main outcome measures were the 8-item Client Satisfaction Questionnaire (CSQ-8) score, the System Usability Scale (SUS) score, and intervention uptake. Participants' experience in each session was evaluated using closed- and open-ended questions for future improvements. Descriptive statistics were used to examine primary outcome and qualitative session evaluations. Participants' responses to each topic of the open questions were summarized.

Results: The intervention was evaluated as being satisfactory (CSQ-8 mean score 21.89, SD 8.72; range 8-32). However, the intervention's usability was still below expectation (SUS mean score 62.8, SD 14.74; range 0-100). The core modules were completed by 10 out of 68 participants (15%), and the study dropout rate was 63% (43/68) at postassessment. In general, the module content was rated positively, with some notes for improvement covering content and technical aspects.

Conclusions: This study indicates that Rileks is potentially feasible for Indonesian university students. In order to be optimally applied in such a context and before scaling up web-based interventions in Indonesia, in general, further development and refinement are needed.

International Registered Report Identifier (IRRID): RR2-10.2196/11493

(JMIR Form Res 2022;6(7):e37278) doi:10.2196/37278

KEYWORDS

Indonesia; cultural adaptation; feasibility study; internet intervention; telemental health; digital mental health; low- and middle-income countries (LMIC); stress management; university students

Introduction

Globally, an increasing number of university students experience stress [1-3]. To a certain extent, stress can be advantageous in stimulating human thriving [4]. However, ongoing high levels of stress may lead to negative outcomes, such as psychological distress, anxiety, depression, physical illness, substance abuse, and impaired work-related and academic performance [4-6]. The prevalence varies across studies and among countries, but overall, studies suggest that 20% to 25% of university students around the globe from various fields of study [7] suffer from psychological distress [3]. However, most of them do not receive support in reducing their high stress levels [5,8]. This is due to various reasons, including fear of stigma for seeking help for mental health problems [5] and limited availability of skilled mental health professionals within universities [5,9]. This has also been reported by university students in Indonesia [10], despite considerable need for psychological support (eg, two-thirds of Indonesian nursing and medical students experience moderate to severe levels of stress) [11,12].

Web-based interventions may overcome some of the issues related to this treatment gap. They may provide an accessible and potentially less stigmatizing alternative compared with face-to-face treatment, because clients can use web-based interventions privately [13,14]. Studies have shown that a web-based intervention can be a stand-alone intervention or can be an adjunct in a psychological intervention, which is known as a blended strategy [15,16]. A blended strategy offers some benefits, such as increasing an intervention's acceptability [17], lowering dropout rates, and increasing the clinician's efficiency [16]. However, implementing a blended strategy requires an adequate infrastructure and a mental health system that promotes the use of telemental health. In Indonesia, where an internet- or web-based intervention is still a novel approach and mental health service is still limited, a blended strategy might require more time to develop. Thus, as an initial step, we focused on a web-based intervention, as this would be more suitable for the current circumstances in Indonesia. Web-based interventions could be especially suitable for a university student population [18], since younger, well-educated individuals already tend to seek information and help for emotional and mental health problems through the internet [19-21]. A meta-analysis reported that web-based and computer-delivered stress management interventions can effectively diminish university students' stress, with an effect size of 0.73 (95% CI –1.27 to –0.19, *P*=.008) [22]. A web-based stress management intervention culturally adapted for Indonesian university students might be feasible because of the increasing availability of the internet in Indonesia [23]; consequently, there would be greater internet access at their universities and, to a lesser extent, at home.

Several web-based stress management interventions for university students have been developed in high-income countries [22]. GET.ON Stress has been investigated in several randomized controlled trials [24-26], is based on the transactional model by Lazarus and Folkman [27], and consists of problem-focused and emotion-focused coping strategies. It was originally developed for German employees, has been adapted for German-speaking university students, and has been renamed StudiCare Stress [28]. As part of this project, the GET.ON Stress intervention was culturally adapted to the Indonesian context, using the integrative cultural adaptation model by Barrera et al [29] as a guideline. The first two steps of the adaptation process, which have been previously reported [30], have led to the Rileks intervention; Rileks means relax, and means calm in the Indonesian language.

Due to the novelty of this kind of intervention in Indonesia, instead of a pilot study, a feasibility study was conducted as the third step with an emphasis on the process of developing, implementing, and assessing preliminary responses of participants to the new intervention [31,32]. Furthermore, a feasibility study helps to evaluate components (ie, participant recruitment, accuracy of the intervention protocol, and ability to execute the new intervention) necessary for the next large-scale study [33,34]. This paper reports on the third step of the model by Barrera et al [29], which is the feasibility study of the preliminary version of Rileks.

The primary aim of this study was to investigate the feasibility of Rileks among university students in Indonesia in terms of acceptability, usability, and intervention uptake. The secondary aim was to investigate stress, anxiety, and depression reduction and improvement of quality of life, as well as to generate feedback for further refinement of Rileks.

Methods

Participants and Sample Size

Inclusion criteria were for participants to have scored 15 (ie, low stress level) or higher on the 42-item Depression Anxiety Stress Scales (DASS-42) [32], to be enrolled in a university in Indonesia, to be 19 years of age or older, to have access to the internet, and to be able to speak Bahasa Indonesia fluently. Participants were recruited between October 10 and 19, 2018. Information about our study and website was disseminated through social media platforms, such as Facebook, Instagram, and WhatsApp groups, and through presentations by the principal investigator (DJ) at events at two universities in Indonesia: YARSI University and the Indonesian State College of Accountancy.

A formal calculation of sample size used for effectiveness trials is not suitable for a feasibility study [35]. In our study, we intended to include at least 50 participants, with a saturation of 75 participants to ensure sufficiently reliable estimates of our main study parameters. We based this estimate on a systematic study analyzing sample sizes in pilot and feasibility studies in the United Kingdom, which reported a median of 36 participants for a feasibility study sample size [35].

Study Design and Procedure

A single-group pretest and posttest design was used. Interested university students who signed up on our website subsequently received a link to the screening measurement, the DASS-42 stress scale. Eligible university students then had to submit their signed electronic informed consent form before completing the baseline measurements (ie, pretest). All included participants received log-in credentials for the intervention, and only those who logged in received posttreatment measurements (ie, posttest) 10 weeks after the pretest. A duration of 10 weeks was chosen as the postmeasurement time point, as we considered this to be sufficient time for participants to complete the intervention. All measures were self-reported and administered online.

Ethics Approval

The study was reviewed and approved by the Indonesian ethics committee at YARSI University (project No. 193/KEP-UY/BIA/VIII/2017).

Intervention

Rileks consists of six sessions and an optional booster session provided 4 weeks after intervention completion. The first session comprises psychoeducational information about stress, based on the emotion-focused and problem-focused coping strategies by Lazarus and Folkman [27]. The second session comprises the six-step problem-solving method based on problem-solving therapy [36,37]. In the second session, participants work on their problem-solving skills by applying the method to their individual problem. In sessions 3 to 5, participants are introduced to emotional regulation techniques based on affect regulation training [37,38]. The techniques include muscle and breathing relaxation, acceptance and tolerance of emotions, and self-support in difficult situations. These techniques are explained one by one in each session, respectively. In the last session, participants are asked to reassess their goals for the training and to identify their personal warning signs for stress. Furthermore, participants are asked to write a letter to themselves about how they imagined their life would be after applying the methods and techniques they had been taught. In addition, a booster session may be given as an option to evaluate the letter they had written to themselves in the last session, reassess their goals, and make plans to continue applying what they had learned from Rileks.

Each session contains general information, examples related to the exercises, exercises guided by electronic coaches (e-coaches), quizzes, slideshows encompassing explanations related to stress management methods and techniques, audio files, and downloadable material, which were all presented on a secure platform. Access to the platform was given to the participants based on their email addresses and self-designated passwords. Participants were advised to log in once or twice per week. A reminder was sent to the participants if they did not log in within 7 days. Within 48 hours after completion of each session, four trained psychologists acting as e-coaches gave personalized written feedback on the exercises. The e-coaches followed guidelines about the feedback process that are defined according to the standardized manual on feedback writing for the intervention.

Primary Outcome Measures

Acceptability was measured by assessing clients' satisfaction with Rileks using the translated version of the 8-item Client Satisfaction Questionnaire (CSQ-8) [39-41]. The scale consists of eight questions answered using 4-point Likert scales (scored from 1 to 4), with total scores ranging from 8 ("great dissatisfaction") to 32 ("great satisfaction"). We set an average score of above 20 (20 is the median total score) as the criterion for acceptable satisfaction. The CSQ-8 has good reliability, as it has been reported to have a Cronbach α of .92 [42].

The Indonesian version of the System Usability Scale (SUS) [43] was used to assess the usability of Rileks in terms of user friendliness [44,45]. The scale comprises 10 statements scored on a 5-point Likert scale, ranging from 1 ("strongly disagree") to 5 ("strongly agree"). The total scores were then transformed into a scale with scores ranging from 0 to 100, with higher scores representing higher usability. A score of 70 or more was considered adequate as a feasibility criterion [45]. The Cronbach α of the Indonesian version has been reported as .84 [43], which indicates good reliability. Both the CSQ-8 and the SUS were only administered at postintervention.

Intervention uptake was measured by assessing the number of participants who completed the core online sessions (ie, sessions 1 to 5), where participants learn the basic principles of problem solving and emotion regulation. The criterion for acceptable adherence was set at 60% or more participants who completed the core sessions [30]. The 60% threshold was based on previous meta-analyses on adherence to internet-based cognitive behavioral therapy (CBT) for depression, which found that 65.1% of participants completed the entire internet-based CBT sessions [46].

Secondary Outcome Measures

The severities of stress, anxiety, and depression were measured using the Indonesian version of the DASS-42 [47]. The scale consists of 42 items divided into three subscales—depression, anxiety, and stress—where each subscale contains 14 items ranging from 0 to 3 and a higher score indicates a greater degree of severity [48]. The Indonesian version of the DASS-42 shows excellent overall reliability, with a Cronbach α of .95, and high internal consistency in the separate depression, anxiety, and stress subscales (α=.91, .85, and .88, respectively) [47].

Quality of life was measured by the Indonesian version of the brief version of the World Health Organization Quality of Life instrument (WHOQOL-BREF) [49]. It consists of two items that measure overall quality of life and general health, and 24 items that measure how the respondent felt in the last 2 weeks across four domains: physical health, psychological health, social relationships, and environmental health [49,50]. Among the Indonesian population, the WHOQOL-BREF has been reported as having internal consistencies, or Cronbach α values, of .41 to .77 in the four domains, respectively [51], and reliability with intraclass correlation coefficients of 0.70 to 0.79 in the four domains, respectively [49]. The DASS-42 and the WHOQOL-BREF were assessed both at pretest and posttest.

Other Measurements

Demographic variables were collected to describe the characteristics of the study population. Cultural-related aspects of the Rileks modules were evaluated within the aspects of language, case examples, and visual presentation. With regard to language use, we asked whether the participants could understand the words in each session. To evaluate case examples and visual presentation, the participants were asked whether the case examples and pictures used in the sessions represented the Indonesian university student context. Participants' experience with e-coaches was evaluated using six questions concerning participants' satisfaction with contact with their e-coaches. Some of the questions were adapted from another study that assessed participants' satisfaction with online therapeutic contact [52]. A selection of responses was provided for each question. An open question asked for participants' recommendations for future improvement of the e-coaches.

Furthermore, by the end of each session on the intervention platform, participants gave an evaluation about their experience in each session. The participants assessed general aspects of each session (eg, usefulness, easiness, and time needed to complete the module) using a Likert scale ranging from 1 (eg, "very useful" or "very easy") to 5 (eg, "not useful at all" or "very difficult"). They also assessed specific aspects of the sessions (eg, the structure of the module) using a Likert scale ranging from 1 ("positive judgement") to 7 ("negative judgement") and answered some open questions (eg, "What did you like and not like about this session?" and "How might you have benefitted from it?").

Statistical Analyses

Descriptive statistics were used to examine primary outcomes, demographic data, and session evaluations. Participants' satisfaction, using the CSQ-8, and system usability, using the SUS, as feasibility parameters were summarized using means and SDs. Intervention uptake was summarized with frequency of participants who logged in or completed each session. Differences in demographic characteristics between participants who logged in and those who did not were tested using a chi-square test in terms of sex, level of education, field of study, and university location. Independent-sample *t* tests were used to assess differences in mean age, DASS-42 scores, and WHOQOL-BREF scores.

Secondary outcomes were analyzed using 2-tailed paired-samples *t* tests with a level of significance of $P=.05$. A normality test at each time point was conducted beforehand, using the Shapiro-Wilk test. As the distributions significantly differed from normal ($P<.05$), a sensitivity analysis using nonparametric tests (eg, the Wilcoxon signed-rank test) was conducted. Furthermore, within-group effect sizes (Cohen *d*) were calculated. SPSS software (version 26; IBM Corp) was used for the analysis.

In order to understand the participants' experience and provide recommendations relevant to future refinement, participants' responses to the open questions for each topic were summarized by categorizing responses with similar themes. Categorization was done by the principal investigator (DJ) and one of the Indonesian team members as a second rater. The summary was finalized by consensus between the two.

Results

Enrollment and Participant Characteristics

A total of 191 university students registered and completed the screening on the study website. Most of the participants (n=169, 88.5%) learned about the study through social media (eg, Facebook, Instagram, and WhatsApp groups). A small proportion (n=22, 11.5%) were informed through presentations by the principal investigator (DJ) at two universities. Of the 191 registered university students, 40 (20.9%) scored lower than 15 on the DASS-42 stress scale, which can be considered as normal stress, and 151 (79.1%) were eligible for inclusion. See Figure 1 for study flowchart.

Baseline questionnaires were completed by 121 out of 151 (80.1%) participants. The majority were female (n=103, 85.1%), had a bachelor's degree (n=94, 77.7%) in social and behavioral sciences (n=83, 68.6%), and studied on the island of Java (n=79, 65.3%). The average DASS-42 stress score at baseline was 23.11 (SD 5.8), indicating moderate stress. One-third of the participants (n=41, 33.9%) fell into the severe stress category (ie, score 26-33). Those who logged in mostly studied in Java and scored lower on psychological quality of life compared to those who did not log in. See Table 1 for all participants' characteristics at baseline.

Figure 1. Study flowchart.

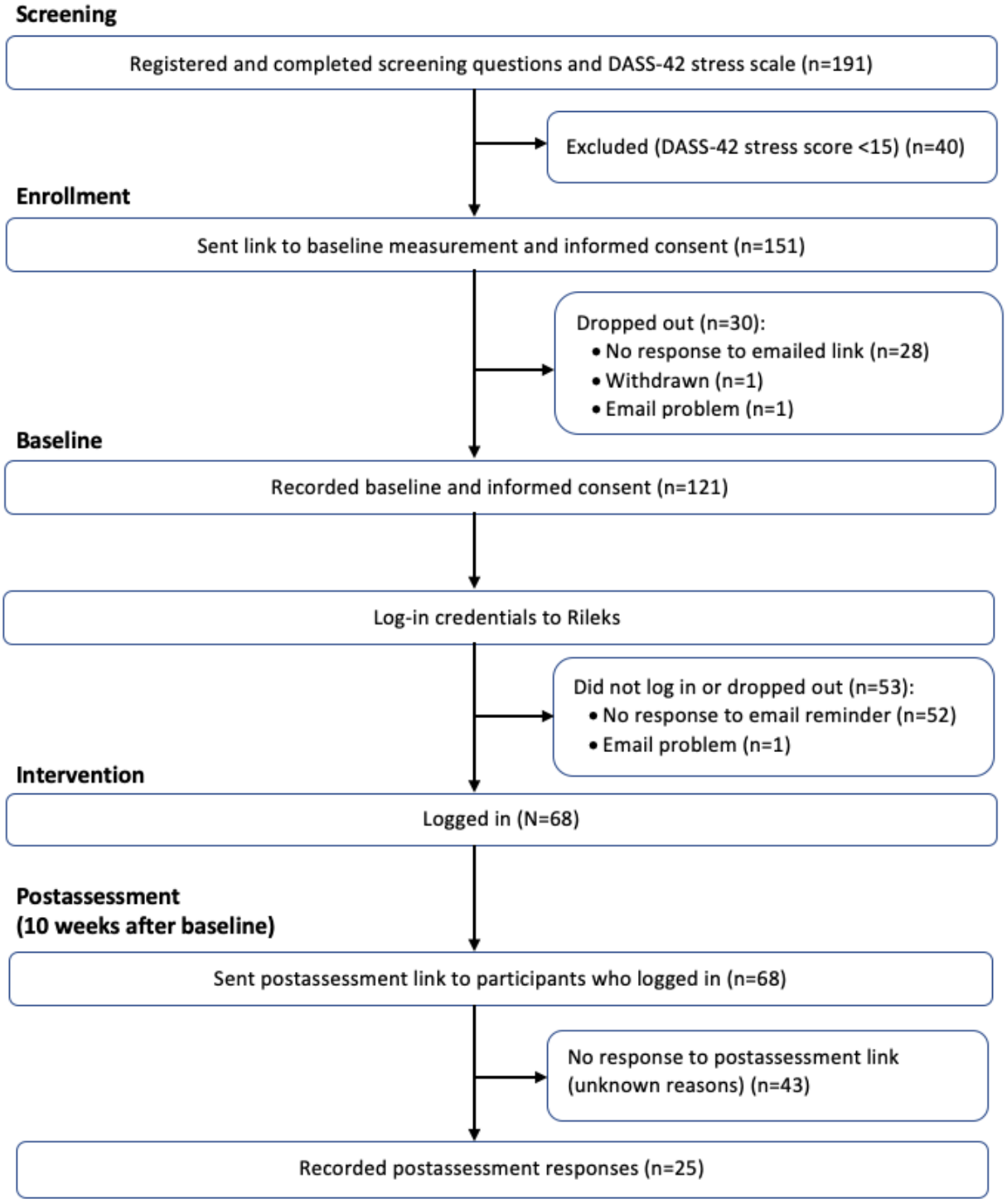


Table 1. Characteristics of participants at baseline and comparisons between those who did not log in and those who did.

Characteristic	Baseline (n=121)	Participants who did not log in (n=53)	Participants who logged in (N=68)	*P* value[a]
Age (years)				
Range	19-42	19-39	19-42	N/A[b]
Mean (SD)	24.03 (4.61)	24.7 (4.41)	23.56 (3.55)	.12
Sex, n (%)				.49
Female	103 (85.1)	44 (83.0)	59 (86.8)	
Male	17 (14.0)	9 (17.0)	8 (11.8)	
Did not fill out	1 (0.8)	0 (0)	1 (1.5)	
Level of education, n (%)				.24
Bachelor's degree and equivalent	94 (77.7)	46 (86.8)	48 (70.6)	
Master's degree	18 (14.9)	5 (9.4)	13 (19.1)	
Doctoral degree	3 (2.5)	1 (1.9)	2 (2.9)	
Did not fill out	6 (5.0)	1 (1.9)	5 (7.4)	
Field of study, n (%)				.29
Social and behavioral science	83 (68.6)	30 (56.6)	53 (77.9)	
Business and administration	12 (9.9)	9 (17.0)	2 (2.9)	
Languages	6 (5.0)	4 (7.5)	0 (0)	
Humanities	0 (0)	3 (5.7)	0 (0)	
Others	20 (16.5)	7 (13.2)	13 (19.1)	
University location, n (%)				.04
Java	79 (65.3)	30 (56.6)	49 (72.1)	
Sumatra	20 (16.5)	10 (18.9)	10 (14.7)	
Kalimantan or Borneo	8 (6.6)	2 (3.7)	6 (8.8)	
Sulawesi	10 (8.3)	8 (15.1)	2 (2.9)	
East Nusa Tenggara	2 (1.7)	2 (3.8)	0 (0)	
Others	2 (1.7)	1 (1.9)	1 (1.5)	
DASS-42[c] subscale score, mean (SD)				
Depression	14.88 (8.70)	13.37 (8.21)	16.16 (8.90)	.92
Anxiety	15.54 (6.72)	15.54 (6.92)	15.58 (6.24)	.99
Stress	23.11 (5.80)	22.89 (5.91)	23.28 (5.53)	.71
Stress level, n (%)				.24
Mild	36 (29.8)	16 (30.2)	20 (29.4)	
Moderate	40 (33.1)	16 (30.2)	24 (35.3)	
Severe	41 (33.9)	20 (37.7)	21 (30.9)	
Extremely severe	4 (3.3)	1 (1.9)	3 (4.4)	
WHOQOL-BREF[d] domain score, mean (SD)				
Physical health	41.83 (7.49)	42.15 (7.23)	41.56 (7.82)	.56
Psychological health	47.90 (14.29)	51 (14.07)	45.34 (13.98)	.02
Social relationship	49.04 (19.29)	48.69 (20.65)	49.33 (18.39)	.43
Environmental health	53.91 (11.67)	53.71 (11.97)	54.07 (11.61)	.98
Overall quality of life	2.98 (0.87)	3.08 (0.88)	2.91 (0.86)	.29
Overall health	2.90 (0.89)	3.04 (0.79)	2.79 (0.97)	.15

[a]*P* values are based on the difference between those who logged in and those who did not.

[b]N/A: not applicable; the *P* value was not calculated for these values.

[c]DASS-42: 42-item Depression Anxiety Stress Scales; scores range from 0 to 42 for each subscale, where a higher score indicates increased severity.

[d]WHOQOL-BREF: brief version of the World Health Organization Quality of Life instrument. Scores range from 1 to 5 for each item in each domain; the total score for each domain was then transformed linearly to a scale ranging from 0 to 100, where a higher score indicates increased quality of life.

Primary Outcomes

A total of 10 weeks after the pretest, posttreatment questionnaires were sent to the 68 participants who had logged in. Questionnaires were completed by 25 participants; thus, the study dropout rate was 63% (43/68). The CSQ-8 mean score was 21.89 (SD 8.72), which met our acceptable satisfaction criterion of a mean score of above 20. Table 2 provides the mean scores of the CSQ-8 items.

Table 2. Scores for all items of the 8-item Client Satisfaction Questionnaire for Rileks (n=25).

Item	Score, mean (SD)[a]
How would you rate the quality of service you have received from Rileks?	3.20 (0.71)
Did you get the kind of service you wanted?	2.92 (0.70)
To what extent has Rileks met your needs?	2.72 (0.61)
If a friend were in need of similar help, would you recommend Rileks to him or her?	3.24 (0.59)
How satisfied are you with the amount of messages you have received from Rileks?	3.16 (0.55)
Has Rileks helped you to deal more effectively with your problems?	3.28 (0.74)
In an overall general sense, how satisfied are you with the help you have received from Rileks?	2.96 (0.73)
If you were to seek help again, would you use Rileks again?	3.04 (0.68)

[a]Items were rated on a 4-point Likert scale, ranging from 1 to 4, where higher scores indicate increased satisfaction.

The SUS mean score was 62.80 (SD 14.74), with the lowest score for the learnability item (Table 3). With regard to intervention uptake, all 121 enrolled participants received log-in credentials for Rileks by email. Of those participants who enrolled, 68 (56.2%) logged in. These 68 participants did not differ from those who never logged in (n=53, 43.8%) on any of the baseline characteristics, with the exception of university location (χ^2_{26}=21.2, *P*=.04) and psychological quality of life (t_{113}=2.14, *P*=.02; Table 1). Of the 68 participants who logged in, the core sessions were completed by 10 (15%) participants. This number is below our acceptable uptake criterion of 60%. Reasons for nonadherence were mostly unknown because those participants could not be reached. The known reasons for nonadherence or withdrawal included time management problems, rare use of email so they missed notifications, unexpected events (eg, internship to a remote village with limited internet coverage), and technical problems, such as unfamiliarity with the log-in system. Table 4 outlines the number of completed sessions in Rileks.

Table 3. Scores for all items of the System Usability Scale for Rileks (n=25).

Item	Score, mean (SD)[a]
I think that I would like to use Rileks frequently.	3.48 (0.82)
I found Rileks unnecessarily complex.[b]	3.40 (1.04)
I thought Rileks was easy to use.	3.56 (0.96)
I think that I would need the support of a technical person to be able to use Rileks.[b]	3.88 (1.01)
I found that the various functions in Rileks were well integrated.	3.72 (0.89)
I thought there was too much inconsistency in Rileks.[b]	3.56 (0.77)
I would imagine that most people would learn to use Rileks very quickly.	3.88 (0.73)
I found Rileks very cumbersome to use.[b]	3.44 (1.01)
I felt very confident using Rileks.	3.32 (0.85)
I needed to learn a lot of things before I could get going with Rileks.[b]	2.88 (1.27)

[a]Items were rated on a 5-point Likert scale, ranging from 1 to 5, where higher scores indicate increased usability.

[b]The scores for this item were reversed.

Table 4. Rileks sessions completed by participants.

Step	Participants (N=68), n (%)
Logged in	68 (100)
Completed module 1	40 (59)
Completed module 2	26 (38)
Completed module 3	16 (24)
Completed module 4	12 (18)
Completed module 5	10 (15)
Completed all 6 modules	9 (13)

Secondary Outcomes

At posttest, out of 25 participants who completed the CSQ-8 and the SUS, the DASS-42 stress scale and the WHOQOL-BREF were completed by 23 (92%) participants, consisting of the core session completers and noncompleters. At posttest, participants reported significantly lower levels of stress (mean –10.04, 95% CI 5.36-14.72), depression (mean –6.85, 95% CI 1.37-12.32), and anxiety (mean –6.45, 95% CI 1.56-11.34), with a high effect size on stress and moderate effect sizes on anxiety and depression. Participants also reported significantly improved quality of life in terms of physical health (mean –21.25, 95% CI –29.05 to –13.44), psychological health (mean –12.50, 95% CI –20.25 to –4.75), overall quality of life (mean –0.55, 95% CI –0.96 to –0.13), and overall health (mean –0.70, 95% CI –1.31 to –0.09). However, we did not find significant differences in the social relationship and environmental health aspects of quality of life (Table 5). Wilcoxon signed-rank tests for nonparametric distributions yielded the same results.

Table 5. DASS-42 and WHOQOL-BREF scores at baseline and posttreatment (n=23).

Measure	Baseline score, mean (SD)	Postassessment score, mean (SD)	*P* value	Cohen *d*
DASS-42[a] subscales				
Stress	24.74 (6.33)	14.70 (11.41)	<.001	0.93
Anxiety	17.95 (8.35)	11.50 (8.34)	.01	0.62
Depression	17.10 (10.74)	10.25 (11.39)	.02	0.58
WHOQOL-BREF[b] domains				
Physical health	42.30 (8.28)	63.55 (13.95)	<.001	1.25
Psychological health	43.85 (17.74)	56.35 (20.14)	.003	0.78
Social relationship	49.40 (17.99)	54.35 (19.64)	.45	0.19
Environmental health	54.55 (13.32)	50.60 (15.14)	.13	0.41
Overall quality of life	2.70 (1.08)	3.25 (1.07)	.01	0.62
Overall health	2.70 (1.03)	3.40 (0.94)	.03	0.54

[a]DASS-42: 42-item Depression Anxiety Stress Scales; scores range from 0 to 42 for each subscale, where a higher score indicates increased severity.

[b]WHOQOL-BREF: brief version of the World Health Organization Quality of Life instrument. Scores range from 1 to 5 for each item in each domain; the total score for each domain was then transformed linearly to a scale ranging from 0 to 100, where a higher score indicates increased quality of life.

Feedback for Future Refinement

Module and Session Evaluation

In general, participants rated the individual modules positively (Table 6). The response summary of the open questions (Multimedia Appendix 1) indicated that, in general, participants were in favor of the clear examples given in the modules and the case examples, to which they felt they could relate. Furthermore, participants suggested that the module content had provided them with new and comprehensive information as well as exercises that could help them manage their stress in terms of recognition, being more reflective and accepting of oneself, learning how to examine problems, and thinking of positive activities that could help them manage their stress. Moreover, Rileks was considered to be a medium where participants were able to express their problems openly without feeling ashamed.

Participants also mentioned things they did not like, mostly technical problems, such as unfamiliarity with the system, poor audio quality, and some files being too large to download, as well as some repeated questions and confusion about how to do the exercises, despite the given examples.

Table 6. Module evaluation.

Evaluation criteria	Evaluations by module					
	Module 1 (n=40)	Module 2 (n=26)	Module 3 (n=16)	Module 4 (n=12)	Module 5 (n=10)	Module 6 (n=9)
General usefulness	Useful	Useful	Useful	Very useful	Useful	Very useful
Easy to complete	Easy	Easy	Easy	Easy	Easy	Very easy
How long did it take? (minutes)	30-60	30-60	30-60	30-60	60-90	30-60
Clarity	Undecided	Clear	Clear	Clear	Clear	Very clear
Subcontent usefulness	Undecided	Useful	Useful	Useful	Useful	Very useful
Pleasantness	Undecided	Pleasant	Undecided	Pleasant	Pleasant	Very pleasant
Comprehensibility	Undecided	Comprehensible	Comprehensible	Comprehensible	Comprehensible	Very comprehensible

Suggestions for improvements included the following: adding interactive functions for communication with the e-coach or other professionals to get direct help, such as live chat, especially in parts where participants need to remember and deal with negative events they had; making the intervention simpler and shorter; adding links to music and video guidance, especially for relaxation; providing downloadable materials; and developing a better display for those accessing Rileks from a mobile phone.

Cultural and Technical Aspects

At posttest, all participants (n=23) could understand the language used in the Rileks modules, including idioms and metaphors. Most participants (n=20, 87%) thought that Rileks used suitable media, such as audio guidance for relaxation or a slide show to explain theory. Case examples in Rileks were considered to represent the Indonesian university student context (n=22, 96%). Furthermore, one participant suggested adding a case example that speaks for university students who come from a family with a low economic background.

e-Coach Evaluation

Of the 23 participants at posttest, 13 (57%) said they missed face-to-face communication and 7 (30%) had hoped for more support. The quality of support was generally considered positive. The e-coach support helped participants gain insight into managing their problems and go through the modules with confidence and motivation. Moreover, the presence of e-coaches made them feel understood and less lonely.

Open questions about suggestions for future refinement were responded to by 22 (96%) participants. Suggestions for refinement from the open questions can be summarized as follows. Firstly, the feedback process could be more interactive and personal (n=3, 13%). Participants hoped that they would be able to ask further questions by replying to the feedback they had received. Secondly, some participants suggested the use of a medium other than email to deliver feedback (eg, using chat; n=2, 9%). Thirdly, they would like the possibility of maintaining a future connection with the e-coaches for further counseling after the intervention had ended (n=2, 9%). Lastly, there was a suggestion to combine offline and online treatment to meet the needs of a participant who felt that they could not fully express themselves to the e-coach through written text (n=1, 4%).

Discussion

Principal Findings

The main aim of this study was to evaluate the feasibility of Rileks as part of an adaptation process of a web-based stress management intervention among university students in Indonesia. Rileks was reported as being acceptable, even though its usability and intervention uptake were still below our expected criteria levels. Study findings showed that participants' stress level and quality of life improved at posttest. The intervention was appreciated by participants in terms of content usefulness, easiness to complete, comprehensibility, suitability for the Indonesian university student context, and the quality of e-coach support, with more e-coach interaction being desired.

Primary Outcomes

Rileks was rated as generally satisfactory, which indicates that the intervention was acceptable. The most satisfying aspect of Rileks as perceived by the participants was that the intervention helped them deal with their problems effectively and that they would recommend it to their friends. With regard to usability, Rileks received a lower than expected threshold score, with the learnability aspect as a main challenge. As reported, participants needed to learn and become familiar with a number of new technical aspects related to the system before they could engage with the intervention. We considered this to be due to the fact that a web-based intervention was relatively new in Indonesia, hence, participants did not have much prior experience in using such an intervention, which may have affected their perception of its usability [53]. Another possible explanation is that the user interface of the system is not user friendly enough. Even though learnability was a challenge, participants felt positive about being able to learn it in a relatively short period of time.

Usability issues were also found in internet-delivered mental health treatments in developed countries [54]. While usability is an essential part of system development, assessing usability of a web-based intervention system is still challenging. As usability is closely connected with interaction design, we are challenged with human-computer interaction issues that are still poorly understood. Furthermore, studies also revealed the need for a guideline for testing the usability of internet-delivered treatment systems [54] and a standardized usability questionnaire for such interventions [55].

The study findings revealed that the uptake of the intervention's core modules was below our criterion level. However, the number of participants who completed the core modules in our study still fell within the range of the reported number of module completers in a systematic review on computerized CBT (12%-100%) [56]. Moreover, a systematic review reported that only 30% of patients adhered to treatment until the third session of face-to-face interventions [57], making an uptake of 24% for module 3 sufficient, considering that computerized CBT is more likely to have a lower adherence rate [56]. Thus, we consider that Rileks still has potential, even though it did not yet meet the uptake criterion level.

Compared to the GET.ON Stress and StudiCare Stress interventions, Rileks' adherence rate of 15% for its core modules still falls behind. Studies of the GET.ON Stress intervention revealed an adherence rate range of 41.9% to 71.8% for its modules [24-26,58,59]. While the StudiCare Stress adherence rate was reported to be high, on average, the participants completed 74.7% of the intervention [60]. The findings seem in line with participants' overall satisfaction with the intervention. Based on the CSQ-8, the StudiCare Stress intervention had a very high satisfaction rate [60], and the GET.ON Stress intervention had high to very high satisfaction rates in its studies [24-26]. Rileks itself had an acceptable overall satisfaction level. Participants perceived that the intervention had helped them to deal more effectively with their problems, but they still considered Rileks to be suboptimal in meeting their needs or to be the kind of service they wanted.

A few reported reasons for nonadherence included time management issues (eg, being too busy), having competing activities, and technical problems, such as unfamiliarity with the log-in system. This is supported by a systematic review that argued that unfamiliarity with computers or the internet and feeling too busy to complete treatment may contribute to dropout from internet-based treatment [61,62]. Other reasons for intervention dropout in our study were that participants missed notifications because they seldom used or checked their email. This implies the need for future studies to use means other than email to send reminders to participants in order to boost intervention uptake, such as personal messages via preferred platforms (eg, instant chat messaging). Furthermore, problems with internet connections in rural areas were also experienced, which indicate that other forms of communication or channels, such as a mobile phone apps, may also be explored as potential options, as the use of mobile apps does not always require constant internet connectivity.

Furthermore, nearly 50% of participants dropped out by the end of module 1, which was the largest number of dropouts compared to other modules. Among all modules, module 1 also received the most "undecided" responses in the evaluation categories of clarity, subcontent usefulness, pleasantness, and comprehensibility. This result may be due to several possibilities. One possibility is that those who dropped out after module 1 were interested in and curious about the intervention at first, but found out that a module-based intervention was not suitable for them after finishing module 1. Another possibility is that participants' experience was suboptimal during module 1 due to a delivery method of psychoeducational content that was less than ideal for university students. This is in line with a systematic review that reported that young people tend to perceive educational material as unengaging, which caused dropout [63]. Further investigation on how to deliver psychoeducational material in engaging ways for university students in Indonesia is needed for future refinement.

Interestingly, it was found that stress severity level was associated with adherence, as those who logged in and engaged tended to have higher stress levels and significantly poorer psychological quality of life compared to those who did not log in. This is supported by other studies that confirmed that participants with less severe problems and difficulties may be less motivated and are subsequently more likely to drop out of internet-based treatment [61,64].

Secondary Outcomes

Participants' stress levels at the posttreatment assessment were significantly lower than at the pretreatment assessment. In addition, the levels of anxiety and depression were also significantly lower among participants at postintervention. This finding is in line with a previous meta-analysis that reported that face-to-face and internet-based interventions do reduce stress levels as well as symptoms of anxiety and depression among university students [22,65,66]. Furthermore, participants' overall quality of life as well as their health and psychological quality of life were better at posttreatment compared with pretreatment. Our study findings may give initial implications for the clinical impact of Rileks, but should be interpreted with caution due to our study design and small sample size.

Future Web-Based Intervention Studies and Rileks Refinements

Our findings demonstrate that web-based intervention studies targeting university students in Indonesia may be feasible. We reached 191 potential participants from several islands in Indonesia in 10 days, with 80.1% (121/151) of invited participants giving their consent and filling in baseline questionnaires. This indicates that university students in Indonesia were interested in our study. The use of the internet, social media, and particularly WhatsApp groups play an important role in the process of reaching targeted potential participants. These strategies allowed us to reach potential participants from many different cities, islands, and fields of study with relatively little effort. Furthermore, taking part in activities involving the target group was also useful in disseminating information on the study. Such activities allowed face-to-face interaction between the principal investigator (DJ) and the target group, where the principal investigator could give information on stress and our study, thus increasing university students' awareness of stress and the credibility of our study. However, our study had a relatively high dropout rate (63%), which is commonly found in other internet-based interventions [56]. It is unfortunate that we did not have sufficient data to explain the reasons for dropout due to the unresponsiveness of our participants. This may be because we approached the participants who dropped out via email, which was not the most convenient medium for them. Thus, for future studies, we recommend using other ways to approach study participants

(eg, sending a short survey about reasons for dropout through any of their preferred devices or platforms).

Participants' feedback for improving future versions of Rileks related to both content (eg, scope of case examples and wording) and technical aspects (eg, suggesting a medium other than email to send notifications and a better display on the mobile phone version). Involvement of relevant stakeholders in the process of refining the content and technical aspects would be valuable for the purpose of achieving an optimal form of Rileks. One highlight of our findings was the request for the availability of interactive communication with the e-coach. This suggested that even though university students in Indonesia are open to the use of internet-based interventions, face-to-face or interactive communication is still preferred. This outcome is in line with other research that suggests that face-to-face interaction is still considered an essential and significant element of mental health services [64,67,68]. According to the latter study, Indonesian people in general would still prefer face-to-face contact, especially when they can access it nearby. This condition may make the use of a web-based intervention to overcome the mental health gap in Indonesia seem challenging. One alternative is to use lay counselors (eg, psychology students trained as e-coaches). Furthermore, reaching out and involving local community health centers to support interventions may be a possible means for future dissemination and implementation of web-based interventions in Indonesia [69]. Another alternative is using a blended strategy by combining modules and instant chat or video conference for more interactive communication between the client and the e-coach.

Our findings indicate that a web-based psychological intervention such as Rileks is acceptable among Indonesian university students and has the potential benefit of clinical effectiveness. Thus, it may provide a good opportunity for university students to have psychological support when there is a mental health service gap within universities and stigma for seeking mental health care. A refinement incorporating all input from participants to overcome challenges of usability and dropout rates is needed. In order to incorporate all input related to the content, technical aspects, and e-coach support into the intervention, involvement and collaboration of relevant stakeholders will be very important (eg, university students, clinicians, lay counselors, software developers, and user interface and user experience experts). Input from participants that we highlight from this study is their desire to have more interactive e-coach support. Further study should make sure that the amount and type of e-coach support that is needed throughout the intervention is provided (eg, chat, video call, email, time of support, and duration). Another shortcoming that we need to tackle is optimizing participants' experience while they are going through the modules, especially the psychoeducational module. Future studies may experiment to find the most engaging way to deliver psychoeducational content to university students in Indonesia.

Strengths and Limitations

To the best of our knowledge, Rileks is the first web-based intervention to provide stress management in the Indonesian university student context. Furthermore, it is among the first web-based psychological interventions being culturally adapted from Western culture to Asian culture. In addition, other than quantitative assessment, our study also considered qualitative data from open questions to give more insight into participants' experience in working with the Rileks modules and system, including e-coaches. However, our study also has a number of limitations. The dropout rate from the enrollment stage to postassessment was relatively high. The data obtained at the postassessment only came from 37% of the participants who were logged in to the intervention and, thus, did not equally represent all participants. Furthermore, 85.1% of the participants were female and, thus, not likely to represent the real university student population, which consists of 56.1% female students and 43.9% male students [70].

Conclusions

Rileks shows potential feasibility for Indonesian university students. However, our findings also underscore the need for further development of this kind of intervention in order to optimize the experience for Indonesian university students. Further refinements are needed regarding content and technical aspects. Despite the potential of web-based interventions and telemental health, in general, to minimize the mental health gap among the Indonesian population, our study implies that more work needs to be done before we can scale up this kind of intervention in Indonesia. In sum, the intervention has potential, but it needs refinement before it can be optimally applied in this setting.

Acknowledgments

This project was supported by the Indonesian Endowment Fund for Education in 2015, funded as part of the principal investigator's (DJ) PhD trajectory.

Conflicts of Interest

EH is a shareholder of HelloBetter – GET.ON Institut für Online Gesundheitstrainings GmbH, which aims to implement scientific findings related to digital health interventions into routine care.

Multimedia Appendix 1
Response summary of open questions on the evaluation of the modules.
[PDF File (Adobe PDF File), 91 KB - formative_v6i7e37278_app1.pdf]

References

1. Robotham D. Stress among higher education students: Towards a research agenda. High Educ 2008 Apr 4;56(6):735-746. [doi: 10.1007/s10734-008-9137-1]
2. Stallman HM. Psychological distress in university students: A comparison with general population data. Aust Psychol 2010 Nov 06;45(4):249-257. [doi: 10.1080/00050067.2010.482109]
3. Kumaraswamy N. Academic stress, anxiety and depression among college students-A brief review. Int Rev Soc Sci Humanit 2013;5(1):135-143 [FREE Full text]
4. Alzahem AM, van der Molen HT, Alaujan AH, Schmidt HG, Zamakhshary MH. Stress amongst dental students: A systematic review. Eur J Dent Educ 2011 Feb;15(1):8-18. [doi: 10.1111/j.1600-0579.2010.00640.x] [Medline: 21226800]
5. Storrie K, Ahern K, Tuckett A. A systematic review: Students with mental health problems--A growing problem. Int J Nurs Pract 2010 Feb;16(1):1-6. [doi: 10.1111/j.1440-172X.2009.01813.x] [Medline: 20158541]
6. Beiter R, Nash R, McCrady M, Rhoades D, Linscomb M, Clarahan M, et al. The prevalence and correlates of depression, anxiety, and stress in a sample of college students. J Affect Disord 2015 Mar 01;173:90-96. [doi: 10.1016/j.jad.2014.10.054] [Medline: 25462401]
7. Hilger-Kolb J, Diehl K, Herr R, Loerbroks A. Effort-reward imbalance among students at German universities: Associations with self-rated health and mental health. Int Arch Occup Environ Health 2018 Nov;91(8):1011-1020. [doi: 10.1007/s00420-018-1342-3] [Medline: 30022255]
8. 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]
9. Drum DJ, Denmark AB. Campus suicide prevention: Bridging paradigms and forging partnerships. Harv Rev Psychiatry 2012;20(4):209-221. [doi: 10.3109/10673229.2012.712841] [Medline: 22894730]
10. Lukito Setiawan J. Willingness to seek counselling, and factors that facilitate and inhibit the seeking of counselling in Indonesian undergraduate students. Br J Guid Counc 2006 Aug;34(3):403-419. [doi: 10.1080/03069880600769654]
11. Purwati S. Tingkat Stres Akademik pada Mahasiswa Reguler Angkatan 2010. Bachelor's Thesis. Kota Depok, Indonesia: Fakultas Ilmu Keperawatan, Universitas Indonesia; 2012. URL: http://lib.ui.ac.id/file?file=digital/20299163-S1958-Tingkat%20stres.pdf [accessed 2019-01-13]
12. Augesti G, Lisiswanti R, Saputra O, Nisa K. Differences in stress level between first year and last year medical students in Medical Faculty of Lampung. J Majority 2015;4:50-56 [FREE Full text]
13. Hedman E, Ljótsson B, Lindefors N. Cognitive behavior therapy via the internet: A systematic review of applications, clinical efficacy and cost-effectiveness. Expert Rev Pharmacoecon Outcomes Res 2012 Dec;12(6):745-764. [doi: 10.1586/erp.12.67] [Medline: 23252357]
14. Sobowale K, Nguyen M, Weiss B, Hai Van TT, Trung LT. Acceptability of internet interventions for youth mental health in Vietnam. Glob Ment Health (Camb) 2016;3:e22 [FREE Full text] [doi: 10.1017/gmh.2016.18] [Medline: 28596890]
15. Berger R, Benatov J, Cuadros R, VanNattan J, Gelkopf M. Enhancing resiliency and promoting prosocial behavior among Tanzanian primary-school students: A school-based intervention. Transcult Psychiatry 2018 Dec;55(6):821-845. [doi: 10.1177/1363461518793749] [Medline: 30091688]
16. Erbe D, Eichert H, Riper H, Ebert DD. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: Systematic review. J Med Internet Res 2017 Sep 15;19(9):e306 [FREE Full text] [doi: 10.2196/jmir.6588] [Medline: 28916506]
17. Lindegaard T, Wasteson E, Demetry Y, Andersson G, Richards D, Shahnavaz S. Investigating the potential of a novel internet-based cognitive behavioural intervention for Dari and Farsi speaking refugee youth: A feasibility study. Internet Interv 2022 Apr;28:100533 [FREE Full text] [doi: 10.1016/j.invent.2022.100533] [Medline: 35433279]
18. Chan JK, Farrer LM, Gulliver A, Bennett K, Griffiths KM. University students' views on the perceived benefits and drawbacks of seeking help for mental health problems on the internet: A qualitative study. JMIR Hum Factors 2016 Jan 19;3(1):e3 [FREE Full text] [doi: 10.2196/humanfactors.4765] [Medline: 27026140]
19. Escoffery C, Miner KR, Adame DD, Butler S, McCormick L, Mendell E. Internet use for health information among college students. J Am Coll Health 2005;53(4):183-188. [doi: 10.3200/JACH.53.4.183-188] [Medline: 15663067]
20. Burns K. Community college student success variables: A review of the literature. Community Coll Enterp 2010;16(2):33-61 [FREE Full text]
21. Batterham PJ, Calear AL. Preferences for internet-based mental health interventions in an adult online sample: Findings from an online community survey. JMIR Ment Health 2017 Jun 30;4(2):e26 [FREE Full text] [doi: 10.2196/mental.7722] [Medline: 28666976]
22. Davies EB, Morriss R, Glazebrook C. Computer-delivered and web-based interventions to improve depression, anxiety, and psychological well-being of university students: A systematic review and meta-analysis. J Med Internet Res 2014 May 16;16(5):e130 [FREE Full text] [doi: 10.2196/jmir.3142] [Medline: 24836465]
23. Penetrasi & Perilaku Pengguna Internet Indonesia. South Jakarta, Indonesia: Asosiasi Penyelenggara Jasa Internet Indonesia (APJII); 2017. URL: https://web.kominfo.go.id/sites/default/files/Laporan%20Survei%20APJII_2017_v1.3.pdf [accessed 2019-01-13]

24. Ebert DD, Lehr D, Heber E, Riper H, Cuijpers P, Berking M. Internet- and mobile-based stress management for employees with adherence-focused guidance: Efficacy and mechanism of change. Scand J Work Environ Health 2016 Sep 01;42(5):382-394 [FREE Full text] [doi: 10.5271/sjweh.3573] [Medline: 27249161]
25. Heber E, Lehr D, Ebert DD, Berking M, Riper H. Web-based and mobile stress management intervention for employees: A randomized controlled trial. J Med Internet Res 2016 Jan 27;18(1):e21 [FREE Full text] [doi: 10.2196/jmir.5112] [Medline: 26818683]
26. Ebert DD, Heber E, Berking M, Riper H, Cuijpers P, Funk B, et al. Self-guided internet-based and mobile-based stress management for employees: Results of a randomised controlled trial. Occup Environ Med 2016 May;73(5):315-323. [doi: 10.1136/oemed-2015-103269] [Medline: 26884049]
27. Lazarus RS, Folkman S. Stress, Appraisal, and Coping. New York, NY: Springer Publishing Company, Inc; 1984.
28. Fleischmann R, Harrer M, Zarski A, Baumeister H, Lehr D, Ebert D. Patients' experiences in a guided internet- and app-based stress intervention for college students: A qualitative study. Internet Interv 2018 Jun;12:130-140 [FREE Full text] [doi: 10.1016/j.invent.2017.12.001] [Medline: 30135777]
29. Barrera M, Castro FG, Strycker LA, Toobert DJ. Cultural adaptations of behavioral health interventions: A progress report. J Consult Clin Psychol 2013 Apr;81(2):196-205 [FREE Full text] [doi: 10.1037/a0027085] [Medline: 22289132]
30. Juniar D, van Ballegooijen W, Karyotaki E, van Schaik A, Passchier J, Heber E, et al. Web-based stress management program for university students in Indonesia: Systematic cultural adaptation and protocol for a feasibility study. JMIR Res Protoc 2019 Jan 25;8(1):e11493 [FREE Full text] [doi: 10.2196/11493] [Medline: 30681970]
31. Gitlin LN. Introducing a new intervention: An overview of research phases and common challenges. Am J Occup Ther 2013;67(2):177-184 [FREE Full text] [doi: 10.5014/ajot.2013.006742] [Medline: 23433272]
32. Orsmond GI, Cohn ES. The distinctive features of a feasibility study: Objectives and guiding questions. OTJR (Thorofare N J) 2015 Jul;35(3):169-177. [doi: 10.1177/1539449215578649] [Medline: 26594739]
33. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol 2010 Jul 16;10:67 [FREE Full text] [doi: 10.1186/1471-2288-10-67] [Medline: 20637084]
34. Tickle-Degnen L. Nuts and bolts of conducting feasibility studies. Am J Occup Ther 2013;67(2):171-176 [FREE Full text] [doi: 10.5014/ajot.2013.006270] [Medline: 23433271]
35. Billingham SAM, Whitehead AL, Julious SA. An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database. BMC Med Res Methodol 2013 Aug 20;13:104 [FREE Full text] [doi: 10.1186/1471-2288-13-104] [Medline: 23961782]
36. Nezu AM, Nezu CM, D'Zurilla TJ. Problem-Solving Therapy: A Treatment Manual. New York, NY: Springer Publishing Company, LLC; 2012.
37. Heber E, Ebert DD, Lehr D, Nobis S, Berking M, Riper H. Efficacy and cost-effectiveness of a web-based and mobile stress-management intervention for employees: Design of a randomized controlled trial. BMC Public Health 2013 Jul 15;13:655 [FREE Full text] [doi: 10.1186/1471-2458-13-655] [Medline: 23855376]
38. Berking M, Ebert D, Cuijpers P, Hofmann SG. Emotion regulation skills training enhances the efficacy of inpatient cognitive behavioral therapy for major depressive disorder: A randomized controlled trial. Psychother Psychosom 2013;82(4):234-245. [doi: 10.1159/000348448] [Medline: 23712210]
39. Larsen DL, Attkisson C, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: Development of a general scale. Eval Program Plann 1979 Jan;2(3):197-207. [doi: 10.1016/0149-7189(79)90094-6]
40. Attkisson CC, Zwick R. The client satisfaction questionnaire. Eval Program Plann 1982 Jan;5(3):233-237. [doi: 10.1016/0149-7189(82)90074-x]
41. Utoyo DB, Lubis DU, Jaya ES, Arjadi R, Hanum L, Astri K, et al. Preliminary study on the effectiveness of short group cognitive behavioral therapy (GCBT) on Indonesian older adults. PLoS One 2013;8(2):e57198 [FREE Full text] [doi: 10.1371/journal.pone.0057198] [Medline: 23437339]
42. De Wilde EF, Hendriks VM. The Client Satisfaction Questionnaire: Psychometric properties in a Dutch addict population. Eur Addict Res 2005;11(4):157-162. [doi: 10.1159/000086396] [Medline: 16110221]
43. Sharfina Z, Santoso HB. An Indonesian adaptation of the System Usability Scale (SUS). In: Proceedings of the International Conference on Advanced Computer Science and Information Systems. 2016 Presented at: The International Conference on Advanced Computer Science and Information Systems; October 15-16, 2016; Malang, Indonesia p. 145-148. [doi: 10.1109/icacsis.2016.7872776]
44. Brook J. SUS: A 'quick and dirty' usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation In Industry. London, UK: Taylor & Francis Ltd; 1996:189-194.
45. Chen X, Wu H, Feng J, Li Y, Lv J, Shi W, et al. Transcriptome profiling unveils GAP43 regulates ABC transporters and EIF2 signaling in colorectal cancer cells. BMC Cancer 2021 Jan 05;21(1):24 [FREE Full text] [doi: 10.1186/s12885-020-07728-x] [Medline: 33402155]
46. van Ballegooijen W, Cuijpers P, van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to internet-based and face-to-face cognitive behavioural therapy for depression: A meta-analysis. PLoS One 2014;9(7):e100674 [FREE Full text] [doi: 10.1371/journal.pone.0100674] [Medline: 25029507]

47. Damanik ED. Pengujian Reliabilitas, Validitas, Analisis Item dan Pembuatan Norma Depression Anxiety Stress Scale (DASS): Berdasarkan Penelitian Pada Kelompok Sampel Yogyakarta dan Bantul Yang Mengalami Gempa Bumi dan Kelompok Sampel Jakarta dan Sekitarnya Yang Tidak Mengalami Gempa Bumi. Master's Thesis. Depok City, Indonesia: Fakultas Psikologi, Universitas Indonesia; 2006. URL: https://lib.ui.ac.id/detail?id=94859&lokasi=lokal [accessed 2019-01-13]
48. Lovibond PF, Lovibond SH. The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther 1995 Mar;33(3):335-343. [doi: 10.1016/0005-7967(94)00075-u]
49. Purba FD, Hunfeld JAM, Iskandarsyah A, Fitriana TS, Sadarjoen SS, Passchier J, et al. Quality of life of the Indonesian general population: Test-retest reliability and population norms of the EQ-5D-5L and WHOQOL-BREF. PLoS One 2018;13(5):e0197098 [FREE Full text] [doi: 10.1371/journal.pone.0197098] [Medline: 29750806]
50. Skevington S, Lotfy M, O'Connell K. The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A Report from the WHOQOL Group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/b:qure.0000018486.91360.00]
51. Salim OC, Sudharma NI, Kusumaratna RK, Hidayat A. Validity and reliability of World Health Organization Quality of Life-BREF to assess the quality of life in the elderly. Univ Med 2007;26(1):27-38 [FREE Full text] [doi: 10.18051/UnivMed.2007.v26.27-38]
52. Knaevelsrud C, Maercker A. Internet-based treatment for PTSD reduces distress and facilitates the development of a strong therapeutic alliance: A randomized controlled clinical trial. BMC Psychiatry 2007 Apr 19;7:13 [FREE Full text] [doi: 10.1186/1471-244X-7-13] [Medline: 17442125]
53. McLellan S, Muddimer A, Peres SC. The effect of experience on system usability scale ratings. J Usability Stud 2012 Feb;7(2):56-67 [FREE Full text]
54. Yogarajah A, Kenter R, Lamo Y, Kaldo V, Nordgreen T. Internet-delivered mental health treatment systems in Scandinavia - A usability evaluation. Internet Interv 2020 Apr;20:100314 [FREE Full text] [doi: 10.1016/j.invent.2020.100314] [Medline: 32426241]
55. Inal Y, Wake JD, Guribye F, Nordgreen T. Usability evaluations of mobile mental health technologies: Systematic review. J Med Internet Res 2020 Jan 06;22(1):e15337 [FREE Full text] [doi: 10.2196/15337] [Medline: 31904579]
56. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: A systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
57. 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]
58. Nixon P, Boß L, Heber E, Ebert DD, Lehr D. A three-armed randomised controlled trial investigating the comparative impact of guidance on the efficacy of a web-based stress management intervention and health impairing and promoting mechanisms of prevention. BMC Public Health 2021 Aug 05;21(1):1511 [FREE Full text] [doi: 10.1186/s12889-021-11504-2] [Medline: 34353294]
59. Ebert DD, Franke M, Zarski A, Berking M, Riper H, Cuijpers P, et al. Effectiveness and moderators of an internet-based mobile-supported stress management intervention as a universal prevention approach: Randomized controlled trial. J Med Internet Res 2021 Dec 22;23(12):e22107 [FREE Full text] [doi: 10.2196/22107] [Medline: 34941541]
60. Harrer M, Apolinário-Hagen J, Fritsche L, Salewski C, Zarski A, Lehr D, et al. Effect of an internet- and app-based stress intervention compared to online psychoeducation in university students with depressive symptoms: Results of a randomized controlled trial. Internet Interv 2021 Apr;24:100374 [FREE Full text] [doi: 10.1016/j.invent.2021.100374] [Medline: 33718001]
61. 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. [doi: 10.1348/014466509X472138] [Medline: 19799804]
62. Beatty L, Binnion C. A systematic review of predictors of, and reasons for, adherence to online psychological interventions. Int J Behav Med 2016 Dec;23(6):776-794. [doi: 10.1007/s12529-016-9556-9] [Medline: 26957109]
63. Garrido S, Millington C, Cheers D, Boydell K, Schubert E, Meade T, et al. What works and what doesn't work? A systematic review of digital mental health interventions for depression and anxiety in young people. Front Psychiatry 2019;10:759 [FREE Full text] [doi: 10.3389/fpsyt.2019.00759] [Medline: 31798468]
64. Arjadi R, Nauta MH, Bockting CLH. Acceptability of internet-based interventions for depression in Indonesia. Internet Interv 2018 Sep;13:8-15 [FREE Full text] [doi: 10.1016/j.invent.2018.04.004] [Medline: 30206513]
65. Yusufov M, Nicoloro-SantaBarbara J, Grey NE, Moyer A, Lobel M. Meta-analytic evaluation of stress reduction interventions for undergraduate and graduate students. Int J Stress Manag 2019 May;26(2):132-145. [doi: 10.1037/str0000099]
66. Harrer M, Adam SH, Baumeister H, Cuijpers P, Karyotaki E, Auerbach RP, et al. Internet interventions for mental health in university students: A systematic review and meta-analysis. Int J Methods Psychiatr Res 2019 Jun;28(2):e1759 [FREE Full text] [doi: 10.1002/mpr.1759] [Medline: 30585363]
67. Sillence E, Briggs P, Harris PR, Fishwick L. How do patients evaluate and make use of online health information? Soc Sci Med 2007 May;64(9):1853-1862. [doi: 10.1016/j.socscimed.2007.01.012] [Medline: 17328998]

68. Harper Shehadeh MJ, Abi Ramia J, Cuijpers P, El Chammay R, Heim E, Kheir W, et al. Step-by-step, an e-mental health intervention for depression: A mixed methods pilot study from Lebanon. Front Psychiatry 2019;10:986 [FREE Full text] [doi: 10.3389/fpsyt.2019.00986] [Medline: 32116815]
69. Kohrt B, Asher L, Bhardwaj A, Fazel M, Jordans M, Mutamba B, et al. The role of communities in mental health care in low- and middle-income countries: A meta-review of components and competencies. Int J Environ Res Public Health 2018 Jun 16;15(6):1279 [FREE Full text] [doi: 10.3390/ijerph15061279] [Medline: 29914185]
70. Higher Education Statistics 2020. Jakarta, Indonesia: PDDikti Kemendikbud; 2020. URL: https://pddikti.kemdikbud.go.id/asset/data/publikasi/Statistik%20Pendidikan%20Tinggi%202020.pdf [accessed 2022-07-09]

Abbreviations

CBT: cognitive behavioral therapy
CSQ-8: 8-item Client Satisfaction Questionnaire
DASS-42: 42-item Depression Anxiety Stress Scales
e-coach: electronic coach
SUS: System Usability Scale
WHOQOL-BREF: brief version of the World Health Organization Quality of Life instrument

Edited by A Mavragani; submitted 14.02.22; peer-reviewed by C Pretorius, L Balcombe; comments to author 11.03.22; revised version received 24.06.22; accepted 27.06.22; published 19.07.22.

Please cite as:
Juniar D, van Ballegooijen W, Schulte M, van Schaik A, Passchier J, Heber E, Lehr D, Sadarjoen SS, Riper H
A Web-Based Stress Management Intervention for University Students in Indonesia (Rileks): Feasibility Study Using a Pretest-Posttest Design
JMIR Form Res 2022;6(7):e37278
URL: https://formative.jmir.org/2022/7/e37278
doi:10.2196/37278
PMID:35852838

Original Paper

An Urban Population Health Observatory for Disease Causal Pathway Analysis and Decision Support: Underlying Explainable Artificial Intelligence Model

Whitney S Brakefield[1,2], MSc, PhD; Nariman Ammar[1,3], MSc, PhD; Arash Shaban-Nejad[1], MPH, PhD

[1]Center for Biomedical Informatics, Department of Pediatrics, College of Medicine, University of Tennessee Health Science Center, Memphis, TN, United States

[2]Bredesen Center for Data Science, University of Tennessee, Knoxville, TN, United States

[3]Ochsner Xavier Institute for Health Equity and Research, Ochsner Clinic Foundation, New Orleans, LA, United States

Corresponding Author:
Arash Shaban-Nejad, MPH, PhD
Center for Biomedical Informatics
Department of Pediatrics, College of Medicine
University of Tennessee Health Science Center
50 N Dunlap Street, R492
Memphis, TN, 38103
United States
Phone: 1 901 287 5863
Email: ashabann@uthsc.edu

Abstract

Background: Many researchers have aimed to develop chronic health surveillance systems to assist in public health decision-making. Several digital health solutions created lack the ability to explain their decisions and actions to human users.

Objective: This study sought to (1) expand our existing Urban Population Health Observatory (UPHO) system by incorporating a semantics layer; (2) cohesively employ machine learning and semantic/logical inference to provide measurable evidence and detect pathways leading to undesirable health outcomes; (3) provide clinical use case scenarios and design case studies to identify socioenvironmental determinants of health associated with the prevalence of obesity, and (4) design a dashboard that demonstrates the use of UPHO in the context of obesity surveillance using the provided scenarios.

Methods: The system design includes a knowledge graph generation component that provides contextual knowledge from relevant domains of interest. This system leverages semantics using concepts, properties, and axioms from existing ontologies. In addition, we used the publicly available US Centers for Disease Control and Prevention 500 Cities data set to perform multivariate analysis. A cohesive approach that employs machine learning and semantic/logical inference reveals pathways leading to diseases.

Results: In this study, we present 2 clinical case scenarios and a proof-of-concept prototype design of a dashboard that provides warnings, recommendations, and explanations and demonstrates the use of UPHO in the context of obesity surveillance, treatment, and prevention. While exploring the case scenarios using a support vector regression machine learning model, we found that poverty, lack of physical activity, education, and unemployment were the most important predictive variables that contribute to obesity in Memphis, TN.

Conclusions: The application of UPHO could help reduce health disparities and improve urban population health. The expanded UPHO feature incorporates an additional level of interpretable knowledge to enhance physicians, researchers, and health officials' informed decision-making at both patient and community levels.

International Registered Report Identifier (IRRID): RR2-10.2196/28269

(JMIR Form Res 2022;6(7):e36055) doi:10.2196/36055

KEYWORDS

health surveillance system; explainable AI; decision support; machine learning; obesity; chronic disease; precision health prevention; semantic inference

Introduction

Background

Enhanced health surveillance systems for chronic disease support could mitigate factors that contribute to the incline of morbidity and mortality of diseases such as obesity. Obesity is linked to increased overall mortality and has reached pandemic proportions, being responsible for approximately 2.8 million deaths annually [1,2]. Obesity represents an excessive and abnormal accumulation of body fat, which leads to adverse health effects that impose a health and financial toll on individuals and society [2]. More than half of the US population has at least one chronic condition, and 27% are living with multimorbidity [3]. These conditions cause more than 1.7 million deaths per year in the United States, where obesity is associated with the top leading causes of death (eg, diabetes, heart disease, stroke, and cancer) [4].

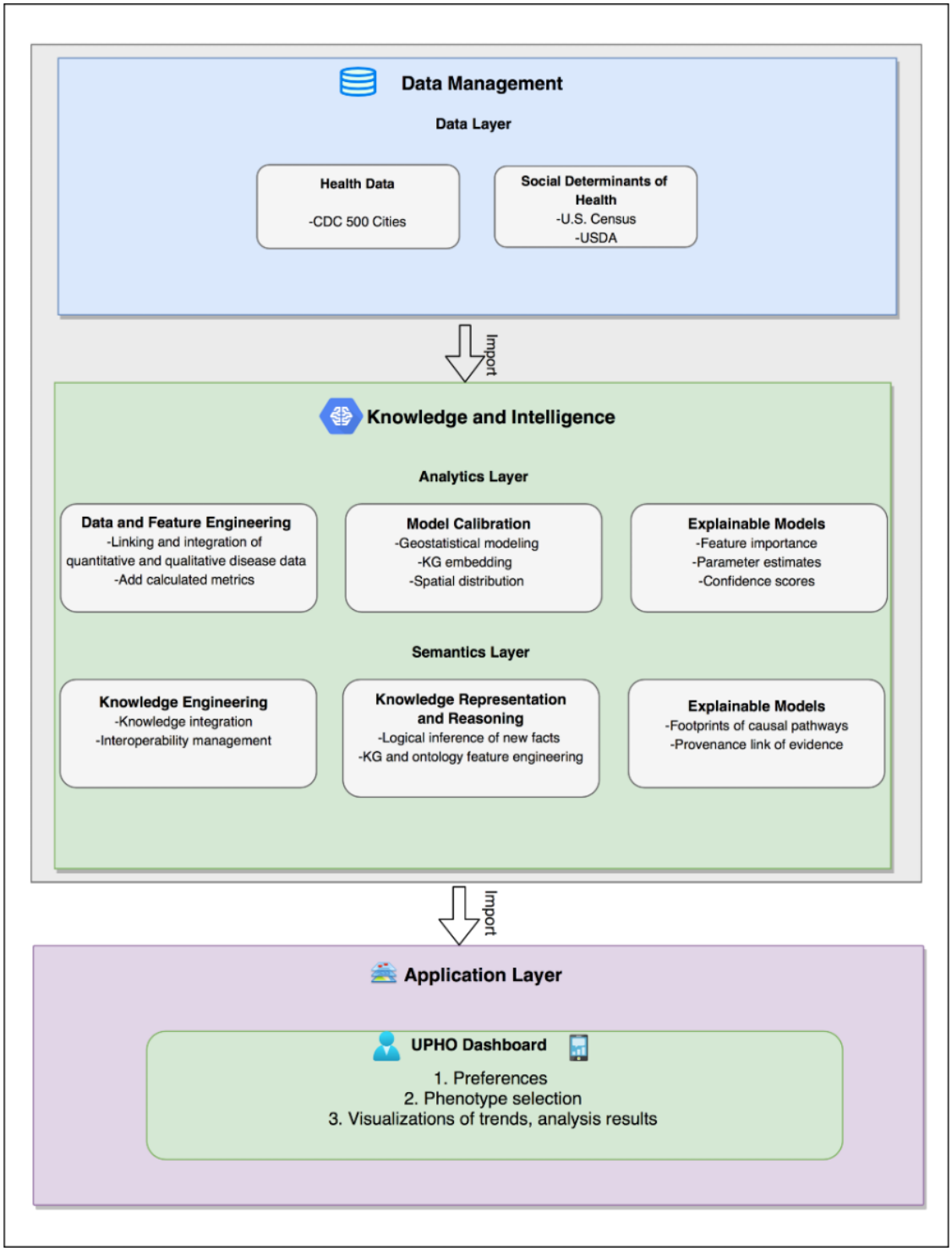
Neighborhood factors such as socioenvironmental determinants of health (SDoH) significantly contribute to these statistics [5-8]. Implementation of an intelligent health surveillance platform that incorporates SDoH can improve preparedness, prevention, and management of this obesity pandemic by assisting in the implementation of effective treatment and interventions.

Health surveillance involves the "ongoing systematic collection, analysis, and interpretation of data essential to the planning, implementation, and evaluation of public health practice, closely integrated with the timely dissemination of these data to those who need to know" [9]. Researchers have aimed to develop chronic health surveillance systems to assist in chronic health decision-making [10-16]. The World Health Organization (WHO) developed a conceptual framework for an Urban Public Health Observatory (UPHO) comprised of 3 domains: mission, governance, and knowledge and intelligence, the latter of which incorporates a data management component [17]. This framework from WHO therefore provides a strategic model for health surveillance.

Many current digital health solutions and electronic health record (EHR) systems lack the ability to incorporate machine learning algorithms into their decision-making process, and even if they do, the algorithms used do not have appropriate capabilities to explain the suggested decisions and actions to human users [18]. Machine learning approaches, so-called black-box statistics, should be trustworthy, transparent, interpretable, and explainable when making decisions in the clinical or health science setting [18-20]. A system's explanation constitutes its interpretability [18,20-22]. Explainable AI (XAI) increases the intelligence delivered to the user by providing explanations, thereby enhancing the interpretability of outcomes and findings. Researcher efforts have been shifting toward applying algorithms that can aid in explaining the results of machine learning models. For instance, the SHAP (Shapley Additive Explanations) analysis [23] is an approach that assigns each model feature an importance score for making a particular prediction. Compared to traditional feature importance analyses, the novelty of SHAP lies in its ability to assess importance at the individual patient level. In this paper, we propose a novel approach to explainability that uses knowledge graphs as a semantic infrastructure explainable by design and enriches those graphs with results from machine learning algorithms as metrics and scores. The semantic causal relationships on the graph provide contextual knowledge around a population, and the metrics support those relationships, which provides 2 levels of evidence: knowledge level and statistical level.

We implement a UPHO platform as a knowledge-based surveillance system that provides better insight to improve decision-making by incorporating SDoH and providing XAI and interpretability functions [24]. Our UPHO consists of 3 layers: data, analytics, and application. In this work, we refine the initial design by incorporating data management, knowledge, and intelligence domains (Figure 1) that are in alignment with the conceptual model by WHO and a focus on the semantics layer.

Figure 1. Expanded Urban Population Health Observatory framework. CDC: US Centers for Control and Prevention; USDA: US Department of Agriculture; KG: knowledge graph; UPHO: Urban Population Health Observatory.



Objectives

The objectives of this article are to (1) expand UPHO by incorporating a semantics layer, (2) cohesively employ machine learning and semantic/logical inference to provide measurable evidence and detect pathways that lead to undesirable health outcomes, (3) provide clinical case scenarios and design case studies on identifying SDoH associated with obesity prevalence, and (4) provide a dashboard design that demonstrates the use of UPHO in the context of obesity, using the provided case scenario.

Methods

UPHO Expansion

Figure 1 shows the expansion of the UPHO to incorporate the semantics layer. In the following section, we provide a detailed description of the UPHO platform expansion design.

Data Management Domain

The data management domain comprises the data layer. The UPHO collects population-level health and SDoH data and

individual-level clinical and demographic data from EHRs through regional registries.

Data Layer

To obtain population-level health data, we used the US Centers for Control and Prevention (CDC) 500 Cities Behavior Risk Factors Surveillance System, which includes data regarding chronic diseases and their behavioral risk factors [25]. These variables are model-based estimates of crude prevalence among adults aged ≥18 years in 2018. We extracted variables pertaining to obesity, lack of physical activity, lack of insurance, and diabetes mellitus at the census tract level.

We extracted population-level SDoH variables that pertain to food insecurity, transportation, and socioeconomic stability at zip code, census tract, census block, and census block group levels from the US Census Bureau 2018 American Community Survey [26] and the US Department of Agriculture Research Atlas [27].

Knowledge and Intelligence Domain

Analytics Layer

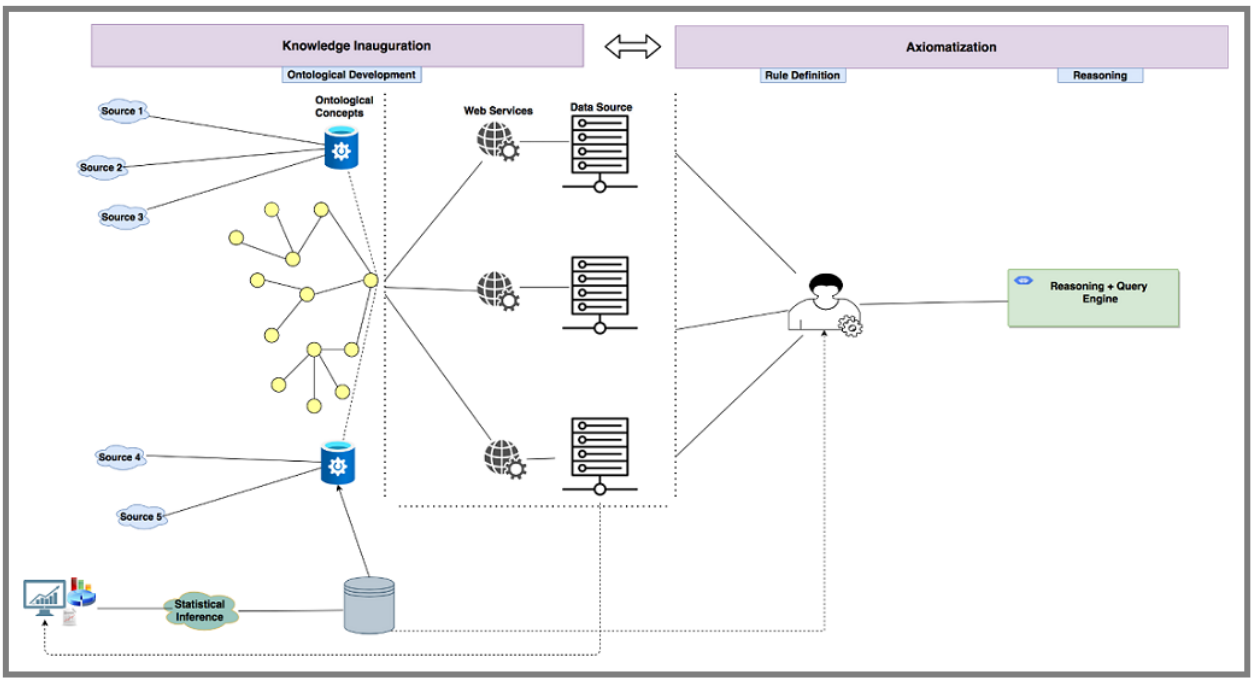
The analytics layer pulls raw data from different sources in the data layer and analyzes it to classify it, predict new relations, conduct spatial pattern detection, and calculate new metrics. The analytics layer also performs feature engineering by deriving new metrics and using them to enrich the original data sets.

Semantics Layer

The stages of the UPHO semantics layer are shown in Figure 2. In the knowledge representation and axiomatization stages, we use semantic web technologies to develop several domain and application ontologies from relevant domains of interest to provide necessary contextual knowledge. Ontologies are systematic representations of knowledge that can be used to integrate and analyze large amounts of heterogeneous data, thereby allowing classification of knowledge [28]. In those ontologies, we define concept hierarchies and rule axioms by using existing domain knowledge, such as the WHO/CDC guidelines, and federal and local sources. We develop new ontologies by reusing several existing domain ontologies. For this study, we adopted concepts, properties, and axioms from 5 different ontologies that we used in our prior work [29-32], specifically (1) disease ontology (DO) [31] (eg, obesity, diabetes); (2) the Childhood Obesity Prevention (Knowledge) Enterprise (COPE) ontology [29] that defines SDoH concepts such as socioeconomic issues (eg, food deserts, income) and behavioral issues (eg, lack of physical activity, purchasing preference); (3) geographical information system ontology (GISO) [32] (eg, zip code, census tract); (4) health indicators ontology (HIO) [32]; and (5) the adverse childhood experiences (ACEs) ontology (ACESO), which defines concepts related to ACEs, health outcomes (eg, mental and physical health), interventions, and SDoH, including axioms that define issues like lack of transportation (eg, limited access to a vehicle and limited access to public transit) and food and how they affect routine follow-up activities (eg, missing medical appointments) [30].

We start our semantic analysis using concepts defined in our ontologies and web services to align concepts to actual data resources, allowing us to construct a population knowledge graph structure that abides by an ontology and contains both data and concepts [33]. We enrich that knowledge graph using a logical reasoner that uses facts derived from existing knowledge, new knowledge extracted from the analytics layer, and the generic rule axioms defined in the domain ontologies that trigger specific actions under certain conditions.

Figure 2. Urban Population Health Observatory semantic layer framework.



Explainable AI

An effective explainable system accounts for the target user group (eg, physician, researcher). Knowledge of the end user is very important for the delivery of decisions, recommendations, and actions. Each analytics and semantics layer contains an explainability component that can be leveraged in the uppermost health applications layer. To maintain features such as data integration, XAI, and interpretability, we must achieve interoperability by using semantics and ontologies. Explanations adaptable to the user can decrease errors in interpretation by enhancing the interpretability of outcomes and findings.

Application Domain

The UPHO platform can be used as a basis to develop several applications, some of which we have already developed, including dashboards [24], mobile health (mHealth) apps [34], digital assistants, and recommender systems [35]. In this article, we leverage the UPHO to implement a dashboard for real-time surveillance. By accessing dynamic knowledge discovered through the UPHO, the dashboard can provide real-time early warnings that are based both on content and context. The platform is accessible to policymakers, physicians, researchers, public health officials, and the public.

Clinical Scenarios

The following sections present 2 clinical case scenarios that focus on a physician and a researcher as users to demonstrate the methodology used in the knowledge and intelligence domain layers and the corresponding dashboard design in the application.

Scenario 1: A physician seeks an effective intervention for an adult African American patient diagnosed with obesity. The physician focuses on how SDoH in the patient's neighborhood can influence the doctor's management plans.

Scenario 2: A researcher investigating the impact of SDoH on obesity seeks an effective intervention for the adult obese populations in Memphis, TN.

Analytics Layer: Machine Learning Model Development

We trained a machine learning–based support vector regression (SVR) machine model [36,37]. We linked CDC 500 Cities population-level obesity and behavior data [25] to the population-level social DoH [26,27] data set at the census tract level in Memphis, TN. We analyzed 8 features and used a Spearman rank test to assess the positive or negative relationship between each feature. We used a variance inflation factor (VIF) to detect multicollinearity between features. To examine patient neighborhood-level exposure, we used SHAP analysis. Table 1 shows the summary statistics for features considered for this study. We trained our support vector model on the 85% of randomly selected training data and tested the model on the 15% of remaining data to ensure the generalizability of the model, and we applied the linear kernel function. We scaled our data to have a mean of zero and standard deviation of one. We applied the grid search optimization method to seek optimal hyperparameters to improve model performance using the Caret package in R software (R Foundation for Statistical Computing) [38]. In addition, to avoid overfitting 5-fold cross-validation was applied to the training data set. We used root mean square error (RMSE) and R^2 to evaluate the performance of the model.

Table 1. Summary statistics for obesity and related risk factors in Memphis, TN, census tract (n=178 census tracts).

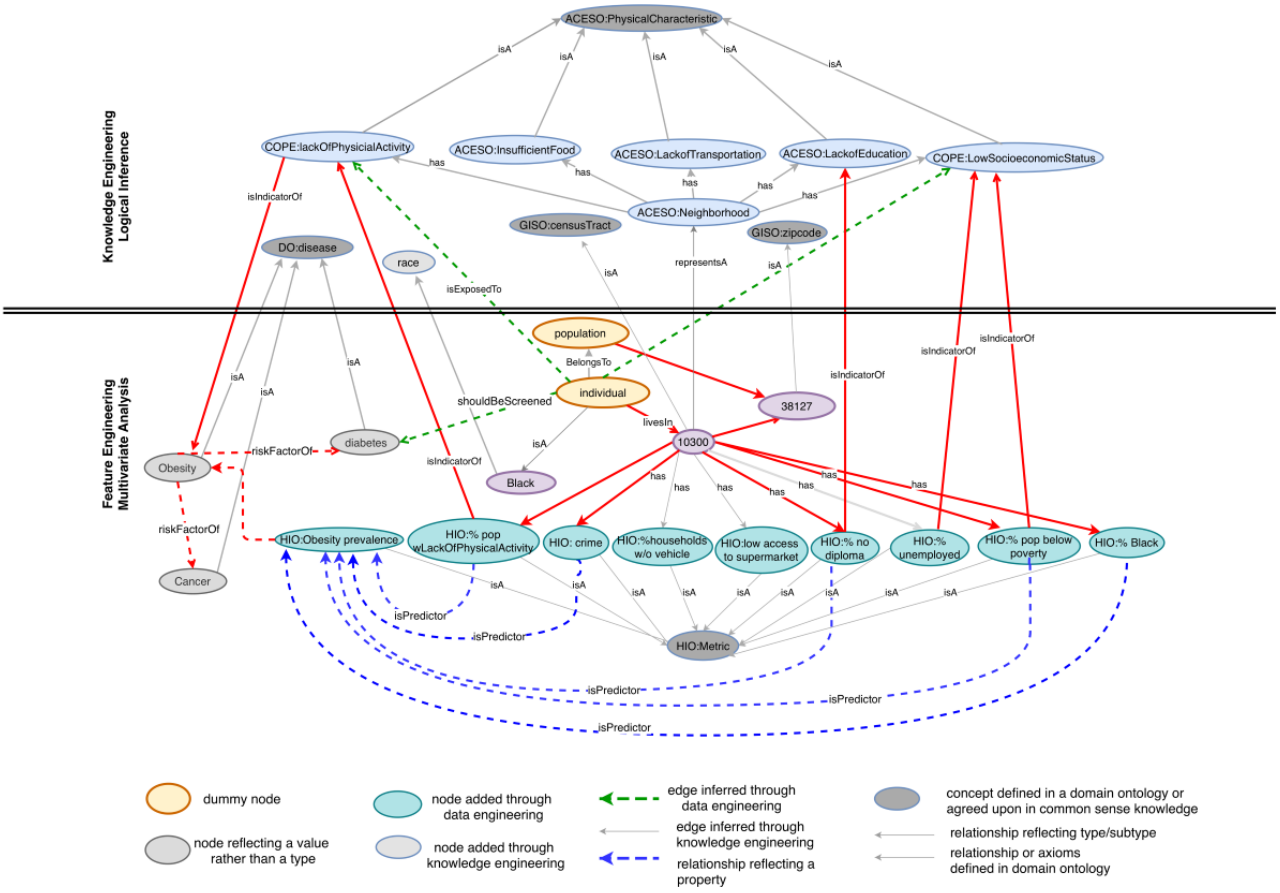
Features	Operationalization	Original, mean (SD)	Training, mean (SD)	Test, mean (SD)
Obesity	Model-based estimate for crude prevalence of obesity among adults aged ≥18 years, 2018	37.50 (7.84)	37.42 (7.54)	37.97 (6.95)
Low access to supermarket	Count of low income population more than half mile from a supermarket in the census tract	1382.20 (108.37)	1345.68 (967.83)	1616.17 (1120.23)
Black	Percentage of population that is Black or African American	63.17 (32.70)	62.22 (33.04)	63.72 (31.88)
Poverty	Percentage of population living below the federal poverty line	28.65 (16.28)	28.27 (16.18)	31.06 (17.06)
Unemployment	Percentage of unemployed population	15.73 (9.31)	15.97 (9.67)	14.16 (6.52)
High school diploma	Percentage of population 25 years or older without high school diploma	10.38 (6.59)	10.23 (6.70)	11.35 (5.89)
Lack of physical activity	Model-based estimate for crude prevalence of lack of physical activity among adults aged ≥18 years, 2018	36.16 (9.80)	35.97 (9.79)	37.34 (9.99)
Crime	Crime rate per thousand people	350.20 (126.26)	160.99 (337.65)	111.93 (80.40)
Lack of access to insurance	Model-based estimate for crude prevalence of lack of insurance among adults aged ≥18 years, 2018	20.21 (6.78)	20.10 (6.81)	20.88 (6.67)

Semantics Layer: Knowledge Graph Generation

We followed the following ordered steps to generate the semantics layer knowledge graph from concepts defined in our domain ontologies.

1. We use concepts, relations, and axioms from domain ontologies to construct a preliminary population knowledge graph. For our scenario, we start by adding a dummy node that represents either a patient or a population (Figure 3). We begin connecting that node to concepts like disease, risk factors, physical characteristics, etc. For example, obesity falls into the disease type represented by the *isA* relation in Figure 3, where *isA* reflects a subtype. For instance, SDoH *isA RiskFactr*, and lackOfTransporation is an SDoH subtype. These different hierarchies are encoded in the ACESO ontology. We also add a prefix before each type to reflect the namespace where that concept is defined (eg, the term DO:Disease reflects that the concept disease is defined in the DO). Relations can also reflect the properties of a node. For example, a patient *livesIn* 38127, which *isA* zip code, and that zip code has 8 census tracts of the type *CensusTract*, defined in the GISO ontology.
2. We populate the generated graph structure with evidence from the data layer. For instance, our data set contains a variable that shows the prevalence of obesity as a percentage metric in specific neighborhoods. We use that information to add edges to our graph that link obesity (as a disease) to prevalence (as a metric).
3. We further enrich and refine the initial graph by performing knowledge engineering using the logical reasoner (Figure 2) and feature engineering using the results from the analytics layer. The logical reasoner uses a set of rule axioms to perform logical inference on concepts already existing in the graph. For instance, our COPE ontology encodes epidemiological causal axioms that link SDoH to negative health outcomes. Textbox 1 shows how we encode the generic axioms R1-R3. When we combine those generic rule axioms with facts about a specific census tract, we can infer all the risk factors associated with living in that census tract, eg, knowing the facts F1 can tell us that the population living in that area may be exposed to risk factors that lead to obesity.
4. After performing the logical inference on the initial graph structure, we incorporate new nodes and edges in the graph corresponding to new concepts (eg, the lackOfPhysicalActivity concept from the COPE ontology) or new relations (eg, *isExposedTo*). The knowledge graph refinement is an iterative process, so we can repeat step 2 until we reach a stable state of the graph after which we can populate the graph with more evidence from our engineered data that we pull from the analytics layer. For that purpose, we use the population-level data about SDoH risk factors collected from US Census, CDC, and USDA. For instance, to capture the lack of physical activity we use the CDC 500 Cities data set. The machine learning analysis performed by the analytics layer provides edges that pertain to prediction (eg, *isPredictorOf*, Figure 3). The final knowledge graph is shown in Figure 3, which provides a generic view of all possible assumptions we can make about this patient or population.
5. To gather the most important information from this graph, a user can trace a specific pathway based on both logical inference and machine learning results. The red arrows in Figure 3 reflect the pathway in our scenario.

Figure 3. Knowledge graph that links concepts defined in domain ontologies (eg, GISO: CensusTract) to data resources stored in databases (eg, percentage Black population) or those derived from the analytics layer. The upper part of the figure shows the nodes and edges produced through semantic inference during the knowledge engineering phase. The lower part of the figure shows the nodes and edges added through ML analysis during the feature engineering phase. GISO: geographical information system ontology; HIO: health indicators ontology; ACESO: adverse childhood experiences ontology; COPE: Childhood Obesity Prevention (Knowledge) Enterprise; DO: disease ontology.



Textbox 1. Encoding axioms as general rules, initial facts, or new facts derived from feature engineering or logical reasoning.

Generic rule axioms

- COPE:lackOfPhysicialActivity *leadsTo* DO:Obesity (R1)
- %ObesityPrevalence:Metric *isHealthIndicatorFor* DO:Obesity (R2)
- Obesity:Disease *isRiskFactorOf* Diabetes:Disease (R3)

Facts

- individual:Patient *livesIn* “10300”:CensusTract (F1)
- “10300”:CensusTract *has* “49”:%PopWLackOfPhysicialActivity (F2)
- “10300”:CensusTract *has* “21”:%PopNoHighSchoolDiploma (F3)
- “10300”:CensusTract *has* “60”:%UnderPovertyLine (F4)
- “10300”:CensusTract *has* “46”:%ObesityPrevalence (F5)

Feature engineered through multivariate analysis

- %PopWLackOfPhysicialActivity:Metric *isPredictorOf* ObesityPrevalence:Metric [using F2-F5] (F6)

Logical reasoning

- individual:Patient *isExposedTo* LackPhysicalActivity:PhysicalCharacteristic [using F1 and F2] (F7)
- individual:Patient *shouldBeScreenedFor* Diabetes:Disease [using R1, R2, R3, F6, F7]

Ethics Approval

No ethics review board assessment was required for this study because we used publicly available data.

Results

Machine Learning Analysis

The significant Spearman rank coefficient and VIF of the 7 features included in this study are shown in Table 2. Any feature exhibiting a VIF of greater than 10 was removed. For the SVR model results, we obtained an RMSE of 0.312 for the training set and 0.203 for the test set, while the R^2 for the training set was 0.91 and that for the testing data set was 0.95 . Since the model provides similar results for training and test data sets, the proposed model does not overfit. The SVR feature importance results range on a scale of 0 to 100, and the greater the score, the most important the feature (Table 3). We found that the percentage of the population lacking physical activity, percentage of population below poverty level, percentage of population without high school diploma, percentage of population unemployed, and percentage of Black population were the most important variables when predicting obesity prevalence in Memphis, TN.

Figure 4 shows SHAP's value plot of feature contribution at the patient neighborhood level (census tract: 10300), which indicates the most important features such as the percentage of the population that lack physical activity and the percentage of population below the poverty level, from the point of view of the prediction of obesity prevalence in the patient neighborhood. The lack of physical activity and poverty had the largest positive (increased) contributions to obesity prevalence. On the other hand, the population of low income and more than a half-mile from the supermarket showed a negative (decreased) contribution but was the least important variable when predicting the patient neighborhood obesity prevalence. The knowledge extracted from our analysis will be used to detect the obesity prevalence pathways, which are defined by the top 5 most important features.

Table 2. Spearman rank coefficient and variance inflation factor for each feature.

Features	Spearman rank coefficient	VIF[a]
Low access to supermarket	0.37	1.70
Black	0.77	2.80
Poverty	0.83	3.66
Unemployment	0.73	3.02
No high school diploma	0.81	3.55
Lack of physical activity	0.92	8.82
Crime	0.37	1.68

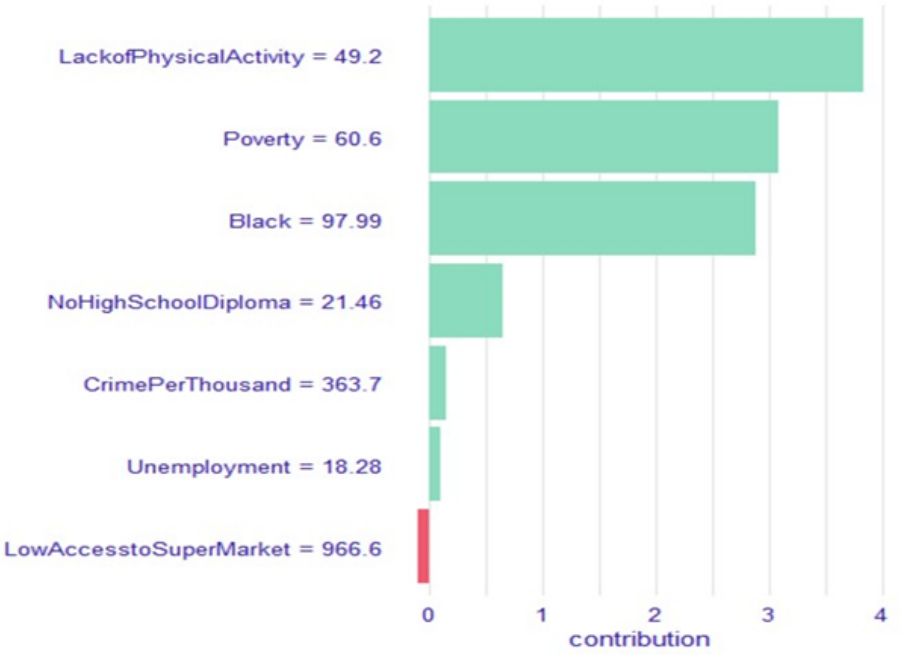
[a]VIF: variance inflation factor.

Table 3. Support vector regression data set–level feature importance score.

Features	SVR[a] feature importance
Low access to supermarket	4.39
Black	68.20
Poverty	78.60
Unemployment	70.16
No high school diploma	73.41
Lack of physical activity	100
Crime	0

[a]SVR: support vector regression.

Figure 4. The Shapley Additive Explanations (SHAP) value plot of the feature contribution (unscaled) for the patient's neighborhood (census tract:10300). The x-axis represents the SHAP's value, and the y-axis represents the features. The lack of physical activity and poverty had the largest positive (increase) contributions to obesity prevalence in the patient's neighborhood.



UPHO: Dashboard Design

Use Case Scenarios

In this section, we describe the semantics feature provided by UPHO through a proof-of-concept prototype that will display the different features of the expanded system by implementing the clinical scenarios described in the previous section.

First, the user will sign into the UPHO platform dashboard, which will determine their specific role and establish the proper access permissions. The user will make the selections from the following menu items:

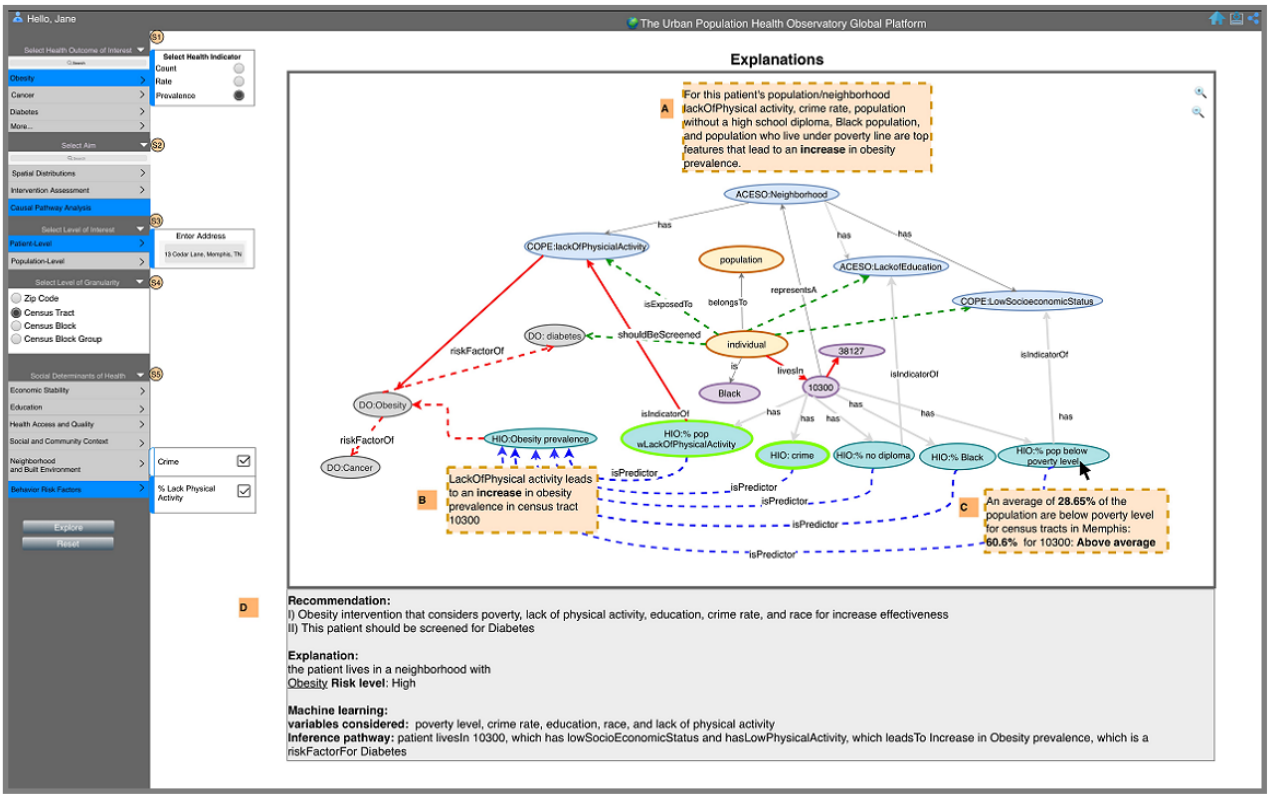
- S1. Select an outcome of interest (eg, obesity prevalence, cancer,)
- S2. Select analytics aim
- S3. Select level of analysis and enter address/location (patient's address [patient-level], city, county, or state [population-level])
- S4. Select geographical level of granularity (eg, zip code, census tract)
- S5. Select SDoH domain-specific risk factors

After making these selections, the system will present on-demand explanations of risk level calculations, based on the selected level of geographic granularity.

Scenario 1

The physician selects "obesity prevalence" as the outcome of interest (S1), and "causal pathway analysis" (S2) as the analytics aim, "patient-level" as the level of interest (entering patient's address, S3), and "census tract" as the geographical level of granularity (S4). The system provides risk-level calculations and descriptive statistics based on the census tract of the patient's address. The physician also has the option to select a particular SDoH of interest in S5, in which case the system will highlight these nodes in the graph. Finally, the user selects "Explore" to generate the results and a corresponding knowledge graph. These results are tailored to the user's interest in patient-level analysis and provide an explanative overview of the analysis results (Figure 5A). The system also allows the user to hover over pathways and nodes to explore explanative knowledge (Figure 5B, 5C) and offers a summary of recommendations and knowledge (Figure 5D).

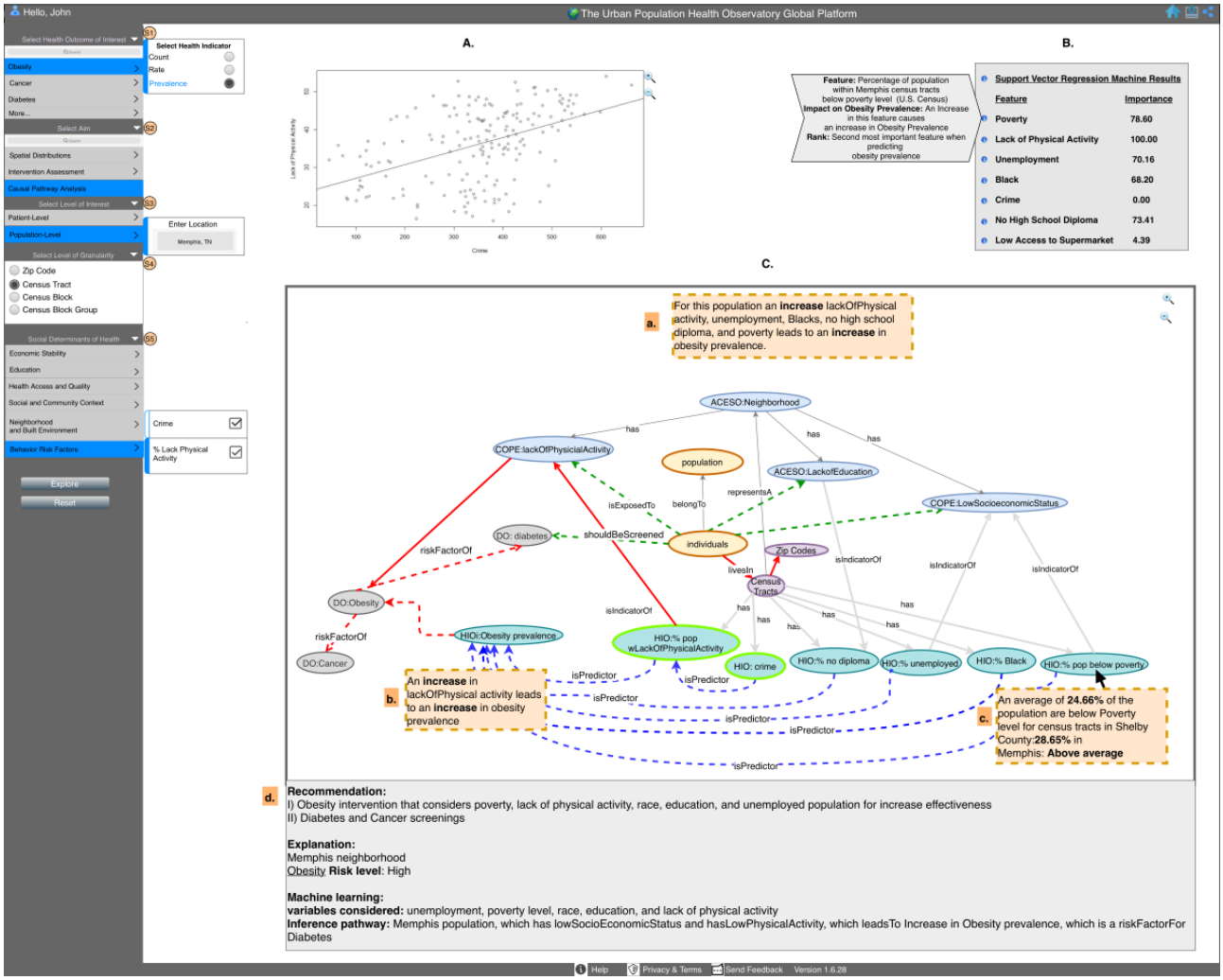
Figure 5. The dashboard of the Urban Population Health Observatory displays a physician user interested in obesity prevalence in her patient's neighborhood with an overview of analysis results (A), explanations displayed when user hovers over a particular pathway (B), knowledge displayed when user hovers over a particular node (C), and summary of recommendations and knowledge (D). ACESO: adverse childhood experiences ontology; GISO: geographical information system ontology; DO: disease ontology; HIO: health indicators ontology.



Scenario 2

Here the researcher has access to more features. The researcher explores the causal pathway analysis aim in a population-level analysis and enters Memphis, TN, as a location of interest at the census tract–level (S1-S3), as shown in Figure 6, and the system provides risk-level calculations for the city of Memphis, TN. The researcher also has access to regression plots (Figure 6, A), which reflect the selection in S5. In section B, the system reports the results from the SVR machine model and provides explanations for each feature included in the model (Figure 6, B). In section C, the explanation pane presents a knowledge graph showing results tailored to the user's interest in population-level analysis (Figure 6, C). The researcher can also hover over pathways and nodes for knowledge (Figure 6, C, a, b, and c), like the physician in scenario 1. The system also offers the researcher a summary of recommendations and knowledge (Figure 6, C, d).

Figure 6. The dashboard of the Urban Population Health Observatory displays a researcher as the user interested in obesity prevalence in Memphis, TN, with univariate regression plot (A), multivariate analysis (B), and (C) which contains an overview of analysis results (a), explanations displayed when user hovers over a particular pathway (b), knowledge displayed when user hovers over a particular node (c), and summary of recommendations and knowledge (d). ACESO: adverse childhood experiences ontology; GISO: geographical information system ontology; DO: disease ontology; HIO: health indicators ontology.



Explanations

The graph part of the dashboard can serve as a tool for researchers and physicians to semantically explain the recommendations that we made about a specific patient or population. The current version of the graph provides 2 different visual cues, as follows.

- Tracing pathways on the graph provides visual cues. The red arrows in Figure 3 show the edges that are part of a causal pathway that leads from risk factors to negative health outcomes for the specific patient, zip code, or census tract. While this path is specific to the selected patient or population, it can be used as a generic metapath. For example, Individual *livesIn* CensusTract→*representsA* Neighborhood→*hasPhysicalCharacteristic*→*RiskFactorFor* Disease. Depending on the level of sophistication desired, the user can trace paths on the graph and click on certain nodes or edges to obtain more insights, including statistically derived evidence or semantically inferred knowledge. They can also track the sources of that knowledge including the ontologies used.
- Clicking on a node or edge on the graph displays analysis results or knowledge. The user can hover over a certain edge (eg, lackOfPhysicalActivity isPredictorOf ObesityPrevalence; Figure 5B) to obtain an explanation of the data that show lackOfPhysical activity in the patient's census tract leads to an increase in the prevalence of obesity. Similarly, the user can hover over a metric node (eg, percentage of the population below the poverty line, Figure 5C) to explain that this patient lives in a neighborhood in which nearly 61% of the population lives below the poverty line, compared to the average in their city, county, or state.

UPHO's metrics can be implemented into the backend of EHR systems (eg, Epic), and the results of those metrics can be rendered on the EHR interface in the form of risk scores on dashboards with severity indicators based on thresholds. Physicians can examine these metrics at the population level or individual patient level. UPHO can alternatively be used in a standalone approach by allowing a physician to extract more details about a single patient by providing the patient's address or a population of patients by providing their city, state, or

county. The input is coded to a geographical level of granularity that can be aligned with the population-level data to gain insights into the patient's environment.

Discussion

Principal Findings

Previous studies provided evidence that socially disadvantaged communities are disproportionately affected by chronic diseases such as obesity [5-8], which is a risk factor for developing diabetes, heart disease, and cancer. The significance of UPHO lies in its ability to provide a multifaceted surveillance system design that serves as an apparatus for actualizing effective interventions, addressing concerns in health disparities, providing awareness to the public, and equipping health officials with a surveillance system that will improve population health decision-making and planning [24]. Using the semantics layer, the UPHO platform provides contextual knowledge by reusing several ontologies focusing on public health (eg, diseases, transportation, geography).

The incorporation of semantics provides the user with an additional layer of explainability and interpretability, which could decrease errors in intervention or treatment due to misinterpretation or misunderstanding. The semantics layer can also use ontologies to overcome the challenge of scattered data sources, thereby assisting in the achievement of interoperability, which will be used to maintain features such as data integration, XAI, and interpretability. We apply logical reasoners to extract and supply knowledge despite limited data.

Similar chronic disease surveillance systems [10-16] have offered approaches to assist in the efforts of improving chronic disease surveillance. Several published systems do not incorporate SDoH data [10-14]. One did not provide an implementation of the systematic framework [16] and several of them did not include XAI as a feature.

We followed the conceptual framework for UPHOs [17] and sought to improve the quality of disease surveillance by incorporating advances in AI and Bog Data, including interactive dashboard design, explainability, data integration, and interoperability, and incorporation of multimodal SDoH data. Developing a multidimensional scalable surveillance system to monitor and detect trends and deliver rapid early warnings and recommendations could assist health officials, physicians, and researchers in mitigating a health crisis such as the ongoing COVID-19 pandemic [24].

Limitations

One of the major limitations of UPHO is that it collects population data so that neighborhood or population assumptions are made for an individual in a clinical setting. For instance, individuals or patients who live in a particular population or neighborhood might not have the same characteristics as other individuals residing in the same neighborhood or population. However, our platform provides an end-to-end approach to examining the environment one resides and incorporates information that is important for the implementation of effective interventions for a given disease.

The future work will be focusing on the further development of the UPHO platform, so it can enable timely, insight-driven decisions and inform immediate or long-term health policy responses [15] to current and future public health crises.

Conclusions

This study leveraged semantic technology and presented a proof-of-concept prototype design for our knowledge-based surveillance system, UPHO, which aims to reduce health disparities and improve population health. The expanded feature incorporates another level of interpretable knowledge needed to inform physicians, researchers, and health officials' decision-making process at the community level. Incorporating XAI helps with the explainability and interpretability of the relevant data, information, and knowledge. Users who are not equipped with domain knowledge could extract common sense knowledge from a system that incorporates XAI [35]. We as humans need a clear visualization and understanding of relationships between parameters in a system to make informed decisions. The lack of understandability and explainability in the health care and public health domains often leads to poor transparency, lack of accountability, and ultimately lower quality of care and biased health policies [39]. Thus, the incorporation of semantics and XAI can improve fairness, accountability, transparency, and trust in health care and public health.

Acknowledgments

The funding for this study is provided by the University of Tennessee Health Science Center.

Conflicts of Interest

None declared.

References

1. Obesity 2021. Geneva: World Health Organization URL: https://www.who.int/news-room/facts-in-pictures/detail/6-facts-on-obesity [accessed 2021-09-09]
2. Abdelaal M, le Roux CW, Docherty NG. Morbidity and mortality associated with obesity. Ann Transl Med 2017 Apr;5(7):161 [FREE Full text] [doi: 10.21037/atm.2017.03.107] [Medline: 28480197]
3. Health and economic costs of chronic diseases. Atlanta: Centers for Disease Control and Prevention; 2021 Jun 23. URL: https://www.cdc.gov/chronicdisease/about/costs/index.htm [accessed 2021-09-09]

4. Boersma P, Black LI, Ward BW. Prevalence of multiple chronic conditions among US adults, 2018. Prev Chronic Dis 2020 Sep 17;17:E106 [FREE Full text] [doi: 10.5888/pcd17.200130] [Medline: 32945769]
5. Lee A, Cardel M, Donahoo W. Social and environmental factors influencing obesity. In: Feingold K, Anawalt B, Boyce A, editors. Endotext. South Dartmouth: MDText.com, Inc; 2000.
6. Suglia SF, Shelton RC, Hsiao A, Wang YC, Rundle A, Link BG. Why the neighborhood social environment is critical in obesity prevention. J Urban Health 2016 Feb 15;93(1):206-212 [FREE Full text] [doi: 10.1007/s11524-015-0017-6] [Medline: 26780582]
7. Reidpath DD, Burns C, Garrard J, Mahoney M, Townsend M. An ecological study of the relationship between social and environmental determinants of obesity. Health Place 2002 Jun;8(2):141-145. [doi: 10.1016/s1353-8292(01)00028-4]
8. Yusuf ZI, Dongarwar D, Yusuf RA, Bell M, Harris T, Salihu HM. Social determinants of overweight and obesity among children in the United States. Int J MCH AIDS 2020 Dec 28;9(1):22-33 [FREE Full text] [doi: 10.21106/ijma.337] [Medline: 32123625]
9. Thacker S, Berkelman R. Public health surveillance in the United States. Epidemiol Rev 1988;10:164-190. [doi: 10.1093/oxfordjournals.epirev.a036021] [Medline: 3066626]
10. Rotmensch M, Halpern Y, Tlimat A, Horng S, Sontag D. Learning a health knowledge graph from electronic medical records. Sci Rep 2017 Jul 20;7(1):5994 [FREE Full text] [doi: 10.1038/s41598-017-05778-z] [Medline: 28729710]
11. Sheth P, Anantharam P, Thirunarayan K. kHealth: proactive personalized actionable information for better healthcare. 2014 Presented at: 40th Int'l Conf Very Large Databases Workshop Personal Data Analytics in the Internet of Things (PDA@IOT); 2014; online.
12. Chari S, Qi M, Agu N, Seneviratne O, McCusker J. Making study populations visible through knowledge graphs. Lecture Notes Comput Sci 2019:53-68. [doi: 10.1007/978-3-030-30796-7_4]
13. Seneviratne O, Rashid S, Chari S. Knowledge integration for disease characterization: a breast cancer example. Int Semantic Web Conf 2018:223-238. [doi: 10.1007/978-3-030-00668-6_14]
14. Gyrard A, Manas G, Saeedeh S, Krishnaprasad T, Amit S. Personalized health knowledge graph. ISWC 2018 Contextualized Knowledge Graph Workshop 2018:1 [FREE Full text]
15. Shaban-Nejad A, Lavigne M, Okhmatovskaia A, Buckeridge DL. PopHR: a knowledge-based platform to support integration, analysis, and visualization of population health data. Ann N Y Acad Sci 2017 Jan 17;1387(1):44-53. [doi: 10.1111/nyas.13271] [Medline: 27750378]
16. Castillo-Salgado C. Developing an academia-based public health observatory: the new global public health observatory with emphasis on urban health at Johns Hopkins Bloomberg School of Public Health. Cad. Saúde Pública 2015 Nov;31(suppl 1):286-293. [doi: 10.1590/0102-311x00132914]
17. Caiaffa WT, Friche AAL, Dias MAS, Meireles AL, Ignacio CF, Prasad A, et al. Developing a conceptual framework of urban health observatories toward integrating research and evidence into urban policy for health and health equity. J Urban Health 2014 Feb 22;91(1):1-16 [FREE Full text] [doi: 10.1007/s11524-013-9812-0] [Medline: 23974945]
18. Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. Mach Learn 2017:2. [doi: 10.1201/9780367816377-16]
19. Battaglia W, Hamrick J. Relational inductive biases, deep learning, and graph networks. 2018. URL: https://arxiv.org/abs/1806.01261 [accessed 2022-06-27]
20. Angelov PP, Soares EA, Jiang R, Arnold NI, Atkinson PM. Explainable artificial intelligence: an analytical review. WIREs Data Mining Knowl Discov 2021 Jul 12;11(5):1. [doi: 10.1002/widm.1424]
21. Rosenfeld A, Richardson A. Explainability in human–agent systems. Auton Agent Multi-Agent Syst 2019 May 13;33(6):673-705. [doi: 10.1007/s10458-019-09408-y]
22. Biemann C. What do we need to build explainable AI systems for the medical domain?. 2017. URL: https://arxiv.org/abs/1712.09923 [accessed 2022-06-27]
23. Lundberg M, Lee S. A unified approach to interpreting model predictions. Proc Adv Neural Inf Proc Syst 2017:4768-4777 [FREE Full text]
24. Brakefield WS, Ammar N, Olusanya OA, Shaban-Nejad A. An Urban Population Health Observatory system to support COVID-19 pandemic preparedness, response, and management: design and development study. JMIR Public Health Surveill 2021 Jun 16;7(6):e28269 [FREE Full text] [doi: 10.2196/28269] [Medline: 34081605]
25. 500 cities project: 2016 to 2019. Atlanta: Centers for Disease Control and Prevention; 2020 Dec 08. URL: https://www.cdc.gov/places/about/500-cities-2016-2019/index.html [accessed 2021-09-20]
26. About the ACS. U. S. Census Bureau. 2021 Jan 04. URL: https://www.census.gov/programs-surveys/acs/about.html [accessed 2021-09-20]
27. Economic Research Service: data sources. U.S. Department of Agriculture. URL: https://www.ers.usda.gov/topics/farm-economy/farm-sector-income-finances/data-sources/ [accessed 2021-09-20]
28. Haendel MA, Chute CG, Robinson PN. Classification, ontology, and precision medicine. N Engl J Med 2018 Oct 11;379(15):1452-1462. [doi: 10.1056/nejmra1615014]
29. Shaban-Nejad A, Buckeridge D, Dubé L. COPE: Childhood obesity prevention [knowledge] enterprise. Artific Intell Med 2011:225-229. [doi: 10.1007/978-3-642-22218-4_28]

30. Brenas JH, Shin EK, Shaban-Nejad A. Adverse childhood experiences ontology for mental health surveillance, research, and evaluation: advanced knowledge representation and semantic web techniques. JMIR Ment Health 2019 May 21;6(5):e13498 [FREE Full text] [doi: 10.2196/13498] [Medline: 31115344]
31. Schriml LM, Arze C, Nadendla S, Chang YW, Mazaitis M, Felix V, et al. Disease Ontology: a backbone for disease semantic integration. Nucleic Acids Res 2012 Jan 12;40(D1):D940-D946 [FREE Full text] [doi: 10.1093/nar/gkr972] [Medline: 22080554]
32. Shaban-Nejad A, Okhmatovskaia A, Izadi MT, Naderi N, Mondor L, Jauvin C, et al. PHIO: a knowledge base for interpretation and calculation of public health indicators. Stud Health Technol Inform 2013;192:1207. [Medline: 23920981]
33. Ehrlinger L. Towards a definition of knowledge graphs. 2016. URL: http://ceur-ws.org/Vol-1695/paper4.pdf [accessed 2022-06-27]
34. Ammar N, Bailey JE, Davis RL, Shaban-Nejad A. Using a personal health library-enabled mHealth recommender system for self-management of diabetes among underserved populations: use case for knowledge graphs and linked data. JMIR Form Res 2021 Mar 16;5(3):e24738 [FREE Full text] [doi: 10.2196/24738] [Medline: 33724197]
35. Ammar N, Shaban-Nejad A. Explainable artificial intelligence recommendation system by leveraging the semantics of adverse childhood experiences: proof-of-concept prototype development. JMIR Med Inform 2020 Nov 04;8(11):e18752 [FREE Full text] [doi: 10.2196/18752] [Medline: 33146623]
36. Awad M, Khanna R. Support vector regression. In: Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers. online: apress; 2015:67-80.
37. Drucker H, Burges C, Kaufman L, Smola A, Vapnik V. Support vector regression machines. 1997. URL: http://papers.neurips.cc/paper/1238-support-vector-regression-machines.pdf [accessed 2022-06-27]
38. Kuhn M. Building predictive models in R using the caret package. J Stat Soft 2008;28(5):1-26. [doi: 10.18637/jss.v028.i05]
39. Shaban-Nejad A, Michalowski M, Brownstein J, Buckeridge D. Guest editorial explainable AI: towards fairness, accountability, transparency and trust in healthcare. IEEE J Biomed Health Inform 2021 Jul;25(7):2374-2375. [doi: 10.1109/jbhi.2021.3088832]

Abbreviations

ACE: adverse childhood experience
ACESO: adverse childhood experiences ontology
CDC: US Centers for Control and Prevention
COPE: Childhood Obesity Prevention (Knowledge) Enterprise
DO: disease ontology
EHR: electronic health record
GISO: geographical information system ontology
HIO: health indicators ontology
mHealth: mobile health
RMSE: root mean square error
SDoH: socioenvironmental determinants of health
SHAP: Shapley Additive Explanations
SVR: support vector regression
UPHO: Urban Population Health Observatory
VIF: variance inflation factor
WHO: World Health Organization
XAI: explainable artificial intelligence

Edited by A Mavragani; submitted 29.12.21; peer-reviewed by S Mukherjee, J Ye, D Surian; comments to author 13.04.22; revised version received 03.05.22; accepted 07.06.22; published 20.07.22.

Please cite as:
Brakefield WS, Ammar N, Shaban-Nejad A
An Urban Population Health Observatory for Disease Causal Pathway Analysis and Decision Support: Underlying Explainable Artificial Intelligence Model
JMIR Form Res 2022;6(7):e36055
URL: https://formative.jmir.org/2022/7/e36055
doi:10.2196/36055
PMID:35857363

Original Paper

Exploring an Artificial Intelligence–Based, Gamified Phone App Prototype to Track and Improve Food Choices of Adolescent Girls in Vietnam: Acceptability, Usability, and Likeability Study

Bianca C Braga[1], MA; Phuong H Nguyen[2], MD, PhD; Noora-Lisa Aberman[3], PhD; Frank Doyle[4], BSc; Gloria Folson[5], PhD; Nga Hoang[6], PhD; Phuong Huynh[6], PhD; Bastien Koch[2], MIDP; Peter McCloskey[4], BSc; Lan Tran[7], MPH; David Hughes[4], PhD; Aulo Gelli[2], PhD

[1]Friedman School of Nutrition Science and Policy, Tufts University, Boston, MA, United States

[2]Poverty, Health, and Nutrition Division, International Food Policy Research Institute, Washington, DC, United States

[3]The Global Alliance for Improved Nutrition, Washington, DC, United States

[4]College of Agricultural Sciences, Pennsylvania State University, University Park, PA, United States

[5]Department of Nutrition, College of Health Sciences, Noguchi Memorial Institute for Medical Research, University of Ghana, Legon, Ghana

[6]National Institute of Nutrition, Ha Noi, Vietnam

[7]Hubert Department of Global Health, Rolling School of Public Health, Emory University, Atlanta, GA, United States

Corresponding Author:
Bianca C Braga, MA
Friedman School of Nutrition Science and Policy
Tufts University
150 Harrison Ave
Boston, MA, 02111
United States
Phone: 1 6176363777
Email: curi.bianca@tufts.edu

Abstract

Background: Adolescents' consumption of healthy foods is suboptimal in low- and middle-income countries. Adolescents' fondness for games and social media and the increasing access to smartphones make apps suitable for collecting dietary data and influencing their food choices. Little is known about how adolescents use phones to track and shape their food choices.

Objective: This study aimed to examine the acceptability, usability, and likability of a mobile phone app prototype developed to collect dietary data using artificial intelligence–based image recognition of foods, provide feedback, and motivate users to make healthier food choices. The findings were used to improve the design of the app.

Methods: A total of 4 focus group discussions (n=32 girls, aged 15-17 years) were conducted in Vietnam. Qualitative data were collected and analyzed by grouping ideas into common themes based on content analysis and ground theory.

Results: Adolescents accepted most of the individual- and team-based dietary goals presented in the app prototype to help them make healthier food choices. They deemed the overall app wireframes, interface, and graphic design as acceptable, likable, and usable but suggested the following modifications: tailored feedback based on users' medical history, anthropometric characteristics, and fitness goals; new language on dietary goals; provision of information about each of the food group dietary goals; wider camera frame to fit the whole family food tray, as meals are shared in Vietnam; possibility of digitally separating food consumption on shared meals; and more appealing graphic design, including unique badge designs for each food group. Participants also liked the app's feedback on food choices in the form of badges, notifications, and statistics. A new version of the app was designed incorporating adolescent's feedback to improve its acceptability, usability, and likability.

Conclusions: A phone app prototype designed to track food choice and help adolescent girls from low- and middle-income countries make healthier food choices was found to be acceptable, likable, and usable. Further research is needed to examine the feasibility of using this technology at scale.

(JMIR Form Res 2022;6(7):e35197) doi:10.2196/35197

KEYWORDS

adolescent; dietary quality; food choice; gamification; low- and middle-income country; smartphone app; mobile phone

Introduction

Background

The consumption of healthy foods and nutrients is suboptimal worldwide, particularly in low- and middle-income countries. Although the daily intake of unhealthy foods (such as red and processed meats and sugar-sweetened beverages) exceeds the optimal consumption level, the daily intake of healthy foods (such as fruits, vegetables, grains, nuts, seeds, and fiber) is far below the recommended levels globally [1]. The importance of diets rich in fruits, vegetables, and whole grains should not be underestimated [2]. Plant-based diets prevent cancers [3] and cardiovascular diseases [4], and the consumption of fruits and vegetables has been associated with increased happiness, life satisfaction, and psychological well-being [5-7]. In contrast, sugar-sweetened beverage consumption is associated with increased adiposity, long-term weight gain, and a higher risk of diabetes [8-11]. The population of Southeast Asia has a low consumption of grains, legumes, fruits, vegetables, nuts, and seeds and a high consumption of processed meat, sugar-sweetened beverages, trans fats, and sodium [1]. Adolescents in this region have especially low consumption of fruits and vegetables [12].

Interventions intended to improve adolescents' knowledge of nutrition and health behaviors can take advantage of new technology to self-monitor diets with feedback on performance [13,14]. Young people are early adopters of mobile phone apps, games, and social media and are highly influenced by peer norms [13,15,16]. Mobile phones and computers are now widely used by adolescents from countries with all income levels [17-19]. The implementation of nutrition interventions based on apps holds great potential to improve the dietary quality of this population, but its impact is yet to be demonstrated [17].

Prior Work

Vietnam is a country in nutrition transition, with great changes in food supply, food prices, household expenditure, diet, and nutrition outcomes [20]. We previously reviewed the apps available in the Vietnamese market that could serve as tools for nutrition interventions [18]. Most of the apps found focused on tracking and changing food choices with the ultimate goal of losing weight [18]. Food consumption is usually tracked by searching for and selecting each food and beverage separately from a dropdown menu and their portion sizes [18]. This was considered so time consuming that even adolescents from Vietnam interested in changing their diets stopped using these apps [18]. There is a lack of free, user-friendly apps that focused on changing dietary quality [18]. We also conducted a literature review for detailed descriptions of eHealth and mobile health technologies used in nutrition interventions [18]. Intervention studies to promote healthy diets in adolescence often ignore techniques to keep participants engaged, leading to low participation and high drop-offs [18]. Adherence to these interventions is poorly reported in the literature [18].

Research Objectives

Our goal was to develop a new app with artificial intelligence (AI) technology to recognize images of foods that can be used to track and influence the quality of food choices of adolescent girls in Vietnam using gamified nudges. Previous apps based on image recognition of foods have successfully estimated the carbohydrate content of meals with higher accuracy than individuals, predicted glycemic dynamics, and helped improve diabetes self-management [21-23]. First, we prepared a food database and an image library by (1) developing a food inventory with priority foods; (2) preparing, cooking, and taking graduated pictures of foods; and (3) annotating pictures and linking them to the food database. Second, we used annotated pictures to train a semantic segmentation model to recognize foods and estimate portion sizes (details of the AI model development are under review elsewhere). Third, we developed the *Food Recognition Assistance and Nudging Insights* (FRANI) app, which included (1) conducting formative research (2 rounds of focus group discussions [FGDs]) with users to develop its interface; (2) validating the image recognition AI technology against the 2 gold standard forms of dietary data collection: 24-hour recalls and weigh food records; and (3) integrating the validated AI technology with the user interface. The app was developed by engineers from our team.

This study aimed to examine the acceptability, usability, and likability of the FRANI prototype. Our three specific objectives were to obtain qualitative inputs on (1) the acceptability of its daily individual- and team-based dietary goals to motivate healthier food choices, (2) the usability and likability of the image capturing feature and the likability of the feedback to users (badges, notifications, and statistics), and (3) to use the inputs of participants to improve the user interface design and functionality. We hypothesized that FRANI would be acceptable, usable, and likable.

Methods

Study Design, Participants, and Setting

We conducted 2 rounds of FGDs with 32 adolescent girls aged 15 to 17 years in Thai Nguyen, Vietnam. The first round of FGDs was aimed at developing the FRANI prototype, and the second aim was to understand its acceptability, usability, and likability and make changes so as to improve its dietary goals, camera, and feedback (badges, notifications, and statistics). We purposely selected an urban public high school because of its good relationship with the research team. All 12th graders were excluded, because they would graduate within a few months from the recruitment period, which could undermine follow-up studies. A total of 2 classes from each of the remaining 2 grades (10th and 11th) were randomly selected, and 8 (25%) girls were randomly selected from each of the 4 classes.

They were approached by the school teachers and divided into 4 FGDs with 8 participants, each according to their school year and specialization (social science and language vs natural and

biological sciences). Two of the girls initially selected did not participate because their parents thought that the study would take time and interfere with their studies. A total of 2 additional adolescents were randomly selected and replaced those who refused to participate. Randomization led to a sample of participants that varied in scholastic performance and socioeconomic characteristics as described in Table 1. The study design and reporting of findings were based on content analysis, ground theory, and the consolidated criteria for reporting qualitative research, which includes a 32-item checklist for interviews and FGDs [24].

Table 1. Characteristics of focus group participants (N=32).

Characteristics	Participants, n (%)
Parents	
Parent's level of education	
Mother	
Less than high school	9 (25)
High school	18 (50)
College	8 (22)
Postgraduate (master's or PhD)	1 (3)
Father	
Less than high school	16 (45)
High school	15 (41)
College	5 (14)
Postgraduate (master's or PhD)	0 (0)
Parents' occupation	
Mother	
Farmer	6 (16)
Blue-collar worker	6 (16)
White-collar worker	8 (22)
Unskilled worker	13 (37)
Stay-at-home parent	1 (3)
Other	2 (6)
Father	
Farmer	5 (14)
Blue-collar worker	7 (21)
White-collar worker	5 (14)
Unskilled worker	15 (41)
Stay-at-home parent	0 (0)
Other	4 (10)
Household and participants	
Household with assets	
Television	36 (100)
Computer	25 (69)
Refrigerator or freezer	36 (100)
Air conditioners	21 (59)
Washing machine	32 (88)
Gas cooker or stove	36 (100)
Water heater	33 (91)
Electric bicycle	21 (59)
Motorcycle	36 (100)
Car	10 (28)
Participants with excellent school performance[a]	
Math	10 (28)
Physics	12 (32)

Characteristics	Participants, n (%)
Chemistry	13 (36)
Biology	21 (59)
Literature	5 (13)
History	20 (55)
Geography	13 (37)
Foreign language	3 (7)
Overall	8 (23)

[a]Excellent school performance was defined as an average grade of ≥9 out of 10.

Research Team, Data Collection, and Analysis

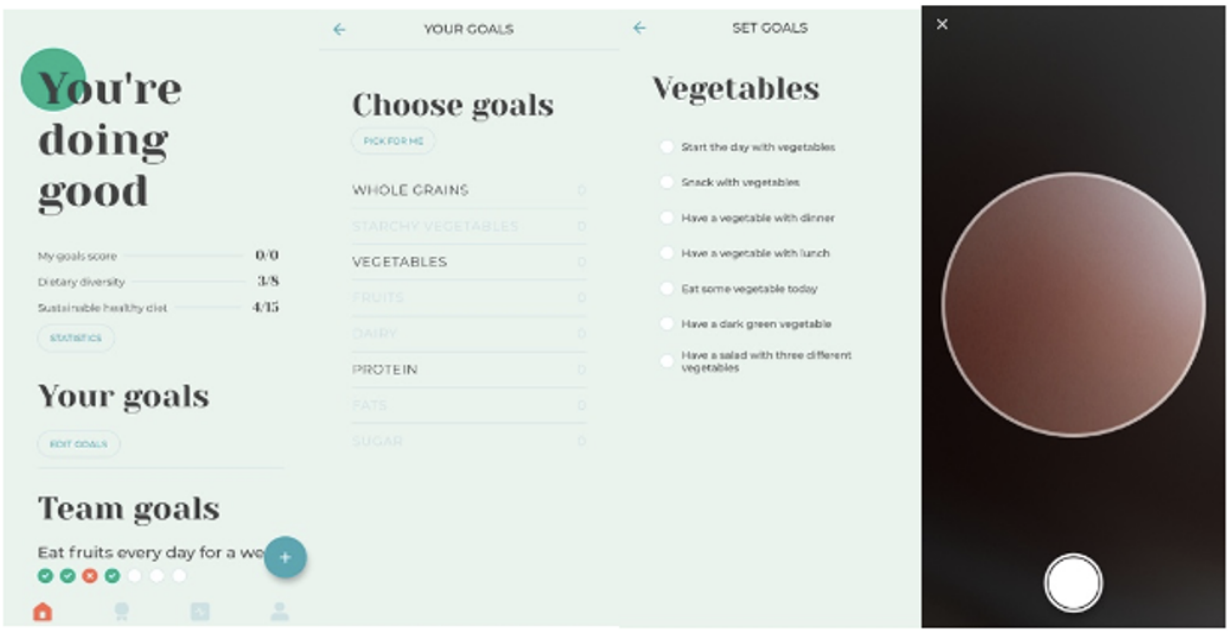
A total of 2 researchers designed a semistructured questionnaire based on an analysis of the content of themes that emerged from previous FGDs [18]. The questionnaire was pilot-tested and changed according to the feedback of the pilot participants. FGDs were then facilitated in Vietnamese by a female senior researcher with a PhD degree and experience in qualitative methods from the National Institute of Nutrition of Vietnam. Moreover, 2 female researchers from the Preventive Medicine Department of the Thai Nguyen University of Pharmacy and Medicine assisted in taking notes, recording the audio, and keeping the time. The researchers did not know the participants before the FGDs. Each FGD took approximately 2 hours, and all interviews were completed in November 2020. Another researcher coded the data.

The researchers presented the wireframes of the FRANI prototype (Figures 1A-1I, Multimedia Appendix 1) along with short instructions to guide the participants to explore the core functionality of their project-provided smartphones. Participants were guided to choose individual-based dietary goals and discussed the acceptability of each goal option. These goals for 3 food groups (vegetables, whole grains, and proteins) were selected from the lists described in Textbox 1. After selecting goals, participants took and uploaded pictures of foods using the camera tool in the FRANI prototype and then received feedback from the FRANI prototype. The facilitator asked questions on acceptability, usability, and likability, as described in Textbox 2. The FGDs were conducted in the school participants' study.

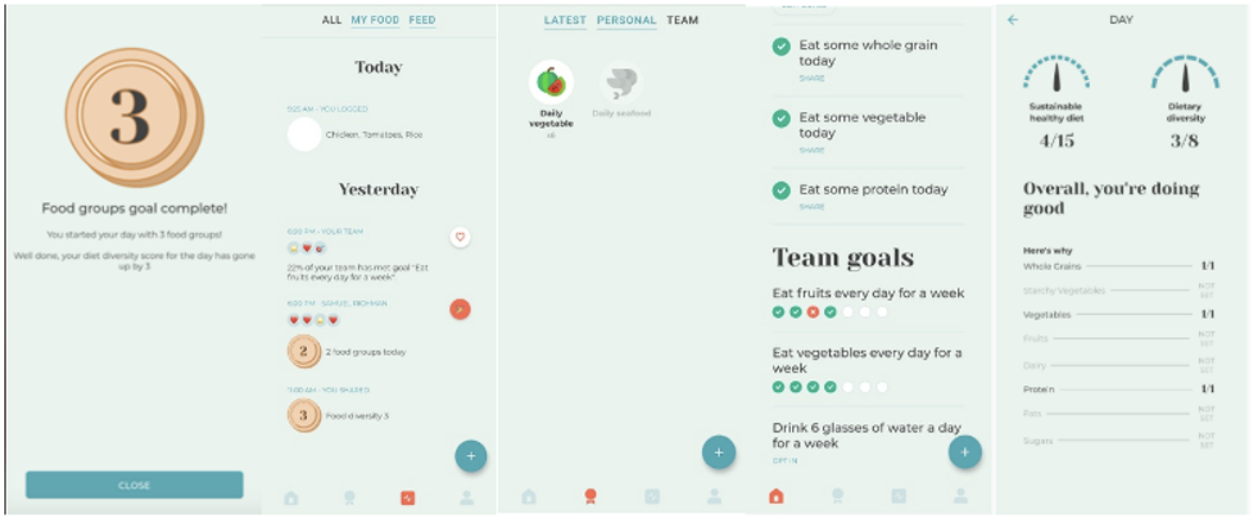
All FGDs were tape recorded and transcribed nonverbatim. Field notes and interviewers' observations were also incorporated into transcripts. Transcribed data from the early interviews were reviewed and discussed by the team to identify gaps in data exploration, which could be further investigated during subsequent interviews. No new information was generated in the fourth FGD, suggesting theoretical saturation. All transcripts were translated into English and randomly checked by a senior researcher who is proficient in English and Vietnamese.

Patterns within and across FGDs were reviewed and organized into the most important common themes using NVivo (version 11; QSR International) software [25]. These themes were acceptability of FRANI prototype's individual- and team-based dietary goals; usability and likability of the FRANI prototype's camera; likability of the feedback, including badges, notifications, and statistics; and likability of the graphic design. Transcripts did not return to the participants for comments and corrections, and interviews with the same group were not repeated. Each FGD lasted approximately 2 hours.

Figure 1. Examples of the Food Recognition Assistance and Nudging Insights wireframes presented for focus groups participants. (A) Home screen, which represents the individual- and team-based scores. The denominators were mistakenly presented in the focus group discussions (FGDs) as 8 and 15 instead of 10 and 14 signifying the food groups for the Dietary Diversity Score (DDS) and subgroups Sustainable Healthy Diet Score (SHDS), respectively. (B) All food groups for which participants could choose goals during the FGDs. (C) All possible goals for vegetables. (D) Camera frame. (E) Confirmation screen with a golden badge for completing 3 goals. (F) Activity screen in which participants can react to what other users posted. (G) Team-based badge. (H) Progress bars for team-based goals (bottom of home screen [A]). (I) Summary of daily statistics. The wireframes presented in the FGDs were written in Vietnamese.



1a 1b 1c 1d



1e 1f 1g 1h 1i

Textbox 1. List of the Food Recognition Assistance and Nudging Insights (FRANI) individual-based dietary goals that were shown to participants and discussed during the focus group discussions. The FRANI includes other food groups, but only those mentioned in this textbox were discussed with the participants.

Vegetables

- Start the day with vegetables
- Snack with vegetables
- Have a vegetable with dinner
- Have a vegetable with lunch
- Eat some vegetables today
- Have a dark green vegetable
- Have a salad with 3 different vegetables

Proteins

- Start the day with proteins
- Snack with proteins
- Have proteins with dinner
- Have proteins with lunch
- Eat some protein today
- Vary your protein routine with beans and peas
- Have a protein food with another food group
- Bake, roast, or grill a protein food

Whole grains

- Start the day with whole grain
- Snack with whole grain
- Have whole grain with dinner
- Have whole grain with lunch
- Eat some whole grain today
- Try a new whole grain
- Have whole grain with another food group

Textbox 2. Questions asked during the focus group discussions.

Acceptability

- What do you think about the dietary goals?
- What do you think about the process of setting dietary goals?
- How does having goals related to food groups help you achieve your personal goals (eg, getting in shape or being healthier)?
- How do you like having team-based goals?
- What do you think about the team-based goals?
- Would you change or improve the way the team-based goals are presented on the home screen?
- On the activity feed, how do you like seeing the combined progress of your team ("A total of 22% of your team has met the goal...")?
- Do you like how much information can be shared on the activity feed?
- How does FRANI compare with other apps you already used to improve your habits in general, and what you eat specifically?
- How was FRANI's information relevant for you to understand the basic principles of a healthy diet?

Likability and usability

- How do you like the tool for taking pictures of your meals?
- How would you change or improve the design of the confirmation screens (the immediate feedback after taking the photo)?
- How do you like the way the feedback on what you ate was provided?
- In general, how do you like the idea of receiving feedback each time right after you logged food?
- How do you like the idea of receiving medals (badges) each time you reach a goal?
- How do you like the design of the medals (badges)?
- How do you like the idea of having daily statistics about what you eat?
- How do you like the way the 3 different scores are presented on the statistics screen?
- Is the progress bar clear?
- Should the layout of that screen be changed or improved?
- Do you like using the app? Why?
- What features do you like best? Why?
- What features do you like least? Why?
- What kind of problems do you see with using the app as it's currently designed?
- What other features do you miss in this app?
- What kind of benefits do you see resulting from using this app?
- What kind of problems do you see resulting from using this app?

Ethics Approval

Methods and materials for the FGDs were approved by the institutional review boards of the International Food Policy Research Institute (protocol code 00007490) and Thai Nguyen National Hospital (protocol code 274/ĐĐĐ-BVTWTN). All participants assented to participate, and their parents consented to their participation. The participants were paid $50,000 Vietnamese dongs (US $2.14). All procedures were performed in accordance with the Declaration of Helsinki.

Results

Acceptability of the Individual-Based Goals

Acceptability of the Vegetable Goals

Participants generally considered the FRANI dietary goals listed in Textbox 2 achievable and suitable for a variety of people because of numerous alternatives. Respondents were confident that they could eat vegetables every day but not in all meals. All alternatives for vegetable goals are shown in Figure 1C (Multimedia Appendix 1). The goal *Eat some vegetables today* was considered easy to achieve, whereas *Eat vegetable for a snack* and *Start the day with vegetables* were considered hard to achieve. *Have a salad with 3 different types of vegetables* was seen as part of a *Western diet*, unsuited for the Vietnamese culture, because they understood a salad to be exclusively

composed of leafy greens. When informed that root vegetables could be part of the salad group, they suggested the goal of having 2 vegetables, because 3 were an uncommonly high variety of vegetables for 1 meal. A group of participants suggested changing the goal *Have a salad with 3 different types of vegetables* for *eat n m*, meaning *a mix of vegetables* in a local dialect, although the term may not be understood in other regions of Vietnam. The goal of *Having a dark green vegetable* was seen as meaningless, because they did not comprehend how the color of food could impact nutrition and had difficulty distinguishing dark green from light green vegetables. There were also concerns about intoxication from pesticide residues in the vegetables.

Acceptability of the Protein Goals

Respondents were unclear about what constitutes a plant-based protein. The facilitators explained that the objective of the *Vary your protein routine with beans and peas* goal was to facilitate the replacement of animal proteins with plant proteins, for example, with soybeans, peas, or tofu. Some participants held that FRANI should show examples of plant-based proteins to facilitate comprehension; others suggested changing the guidance to *Eat plant-based proteins instead of animal-based protein*. This goal and *Snack with protein* were seen as difficult to achieve, because adolescents mostly eat pastries, cakes, or candies as snacks. The healthfulness of the cooking methods of the goal *Bake, roast, or grill protein-rich foods* divided opinions. Some asked to include an explanation in FRANI about why proteins should be roasted, grilled, or baked. Others proposed boiling, air frying, and steaming, noting that well-known health influencers used these cooking methods.

Acceptability of the Whole-Grain Goals

Whole grains, including rice, wheat, corn, and others [26], were widely mistaken for packaged breakfast cereals:

> *I thought that [whole grain] was a packaged cereal.* [FGD 1]

Others thought whole grains were a synonym for unprocessed food or snacks:

> *I think that whole grains are foods that have not undergone many preparations and are pre-processed.* [FGD 3]

> *[A whole grain] is everything that is considered a snack, for example, bread, or milk.* [FGD 1]

Most whole-grain goals were either achievable or redundant after the leads explained their meaning. The only exception was the goal *Start the day with a whole grain*, which was perceived as difficult to achieve for those who did not eat breakfast. Participants thought that the goal *Try a new whole grain every day* required too much cooking time and creativity. To avoid misunderstanding, it was suggested that *whole grains* should be renamed.

The Relationship Among the Dietary, Health, and Fitness Goals

Overall, participants were interested in healthy eating but felt rightly unequipped to make healthy food choices. All participants said that the FRANI should ask for weight and height, medical history, and fitness and health goals. On the basis of this information, personalized, suitable dietary goals that consider both the quality and quantity of food groups should be suggested:

> *There must be a complete set of information. It [FRANI] cannot help us with anything if it's too vague. When people have filled out all the information [weight, height, medical history, and health condition], the app should make personalized suggestions so people can pay attention to what they eat.* [FGD 3]

Most participants would only use an app that could help them achieve fitness goals, so there was a high demand for a feature to set goals such as weight gain or loss and building muscle mass. They said that FRANI should suggest what dietary goal users should set to achieve their health and fitness goals and indicate what and how many food groups, nutrients, and calories were missing from each meal:

> *The app's purpose is to promote healthy eating, [...] but I still don't know [after choosing dietary goals on FRANI] if the way I eat is right or wrong, if it's nutritionally inappropriate... The app has to give me advice so I can follow it.* [FGD 3]

> *Choosing goals should not be based on "users' opinions.*" [FGD 4]

Other comments suggested difficulty in deciding among multiple goals, and some said that FRANI should choose their dietary goals. They also stated that FRANI should provide information about what foods are contained within each food group to avoid misunderstanding. FRANI was interpreted as an app better suited for helping with weight gain than weight loss, because receiving rewards for achieving multiple goals was associated with eating more.

Acceptability of the Team-Based Dietary Goals and Competition

The team-based goals included *eat fruit*, *eat vegetables*, and *drink 6 glasses of water*. At the bottom of the home screen, there was a progress bar for active goals, with 1 box for each day of the week (Figure 1H, Multimedia Appendix 1). Individual members of a team would have their boxes checked if they met the goal that day. For example, if a user chose to opt in to a team-based goal of having fruits every day, they would be grouped with other users who opted in to the same goal. If the individual user achieved that team goal for that day, they would have the box checked, turning it green, or leave it red if they did not. The participants liked the idea of competing in teams. Some participants suggested adding a search engine to find users by their names, inviting contacts to join FRANI, and receiving friendship suggestions so that they could have friends as teammates.

Some participants were more excited about team-based than individual-based goals, saying that competition could trigger motivation, create connections, and foster social support. They wanted more options for team-based goals, because the team-based goals available were seen as easy to achieve:

> *The team-based goals are easy to achieve because these are foods that we eat every day and that are essential.* [FGD 1]

They would like to see team members individually ranked, with the lowest achievers at the bottom and highest achievers at the top:

> *Here [home screen] it must have a specific team name, and when I click on the team name, I see what goals people have achieved today and the rankings of the team members below.* [FGD 3]

Many claimed that underachievement would make them try harder and that they were not worried about peer pressure.

Usability and Likability of the Image Capture Camera

Many participants said that families would support the use of FRANI as long as it helped improve the quality of their diet. However, several challenges were raised. The first perceived usability barrier of the FRANI was the size of the FRANI camera window, which is considered too small to capture images of whole family food trays (Figure 1D, Multimedia Appendix 1), because meals are typically served and eaten communally in Vietnam. Another perceived barrier to usability was the inconvenience for family members in waiting for the process of taking pictures completely before starting to eat their meals. Furthermore, separating what they eat from the rest of the family meal would be seen as disrespectful by most families, especially in the presence of older adults, house guests, and in restaurants:

> *If there are many people of my family [during mealtime], including the elders, it would be strange for me to separate my food from everyone else.* [FGD 3]

An additional usability limitation was that participants did not know how much they ate during the shared family meals:

> *When eating with others, we share food so it [food consumption attributed by FRANI to user] can't be accurate. It's complicated when many people are eating at once.* [FGD 3]

Similarly, after taking and uploading photos, they may consume additional servings. Thus, participants requested a wider camera frame and feature to digitally separate what users ate. Finally, participants were concerned about how much the picture could illustrate the food:

> *There are many things in a sandwich, but the app won't know all the ingredients unless I take it apart.* [FGD 2]

They also talked about the difficulty of taking pictures of foods that come inside packages, such as potato chips. They suggested that FRANI recognizes food packages. They also wanted the ability to upload food pictures to the app after meals to avoid the inconvenience of waiting for other family members or in case there was no internet access. They would prefer to log the same picture for multiple similar meals to save time and phone memory.

Likability of the Feedback, Including Notifications, Badges, and Statistics

Participants said they would like FRANI to give immediate feedback on the quantity and quality of foods eaten, including indications on which foods are unhealthy so as to help them make healthier choices for their next meal:

> *It [FRANI] should provide an immediate response so that I can learn from the previous experience for the next meal.* [FGD 4]

Some said they would rather receive a brief notification reminding them what they still have to eat to achieve individual-based dietary goals, than to get a daily report at night with a summary of what was eaten. They said that FRANI should also notify users about dietary goals that were not met at the end of each day and remind them to take pictures at the time of their meals. Others would like to choose the time of day to receive notifications.

Badges, or visual representations of individual- and team-based dietary goal achievements (Figure 1E, Multimedia Appendix 1), are a form of feedback that can be shared with FRANI friends on an activity screen (Figure 1F, Multimedia Appendix 1). The badge system ranges from bronze to silver to gold and streak badges praise for high dietary diversity multiple days in a row. They liked the badge system but preferred to receive only one badge per day:

> *I think there should only be one badge a day. This means that if you accomplished all three goals very well, you'll get a gold badge, if you did not do too well [...] you would get a silver badge. If you didn't accomplish any goals, you will not get a badge.* [FGD 3]

Participants liked the fact that friends could react (send *hearts*) when they shared their badges in the feed and agreed that comments should not be allowed to avoid cyberbullying (Figure 1F, Multimedia Appendix 1).

Participants were also presented with statistics on food intake using simple charts designed to minimize cognitive load [27]. Most participants understood the 3 different dietary scores calculated by the FRANI prototype, as explained in Multimedia Appendix 2 [28,29]. Participants liked daily statistics on dietary scores for individual goals (Figure 1I, Multimedia Appendix 1). They suggested diagrams to show if users met goals, including how much was eaten from each food item and how much they would still need to eat to achieve the goals. They would like to have a history of daily statistics permanently saved in the FRANI.

Likability of the Graphic Design

Participants wanted the graphic design of the FRANI prototype to be more colorful and the icons to be brighter. The graphic design of the badges was considered unattractive (Figure 1E and 1G, Multimedia Appendix 1). Any food logged by participants during the FGDs showed a watermelon icon as a badge (Figure 1G, Multimedia Appendix 1); therefore, they thought the FRANI prototype did not recognize the foods correctly. This was not a recognition mistake but an unclear

graphic design. They suggested that each food item should have its own differentiated badge. They also stated that graphic design should be more intuitive. The plus (+) icon used to take pictures on the home screen (Figure 1A, Multimedia Appendix 1) should be replaced by a camera icon.

Discussion

Principal Findings

This formative study investigated the acceptability, usability, and likability of an AI-based, gamified mobile phone app prototype to track and improve the nutritional quality of food choices of adolescent girls in Vietnam. The FRANI prototype was designed to self-monitor and influence food choices by considering the capabilities of and opportunities for potential users [30]. We included game elements, such as setting dietary goals and giving feedback on performance, including a badge system, notifications, and statistics, to motivate app use [31-35]. We also added elements of social media, such as sharing information with other users, a newsfeed, and the possibility to react positively to what others posted [34]. Participants of the 4 FGDs provided feedback on the app prototype, informing platform designers of their expectations [36].

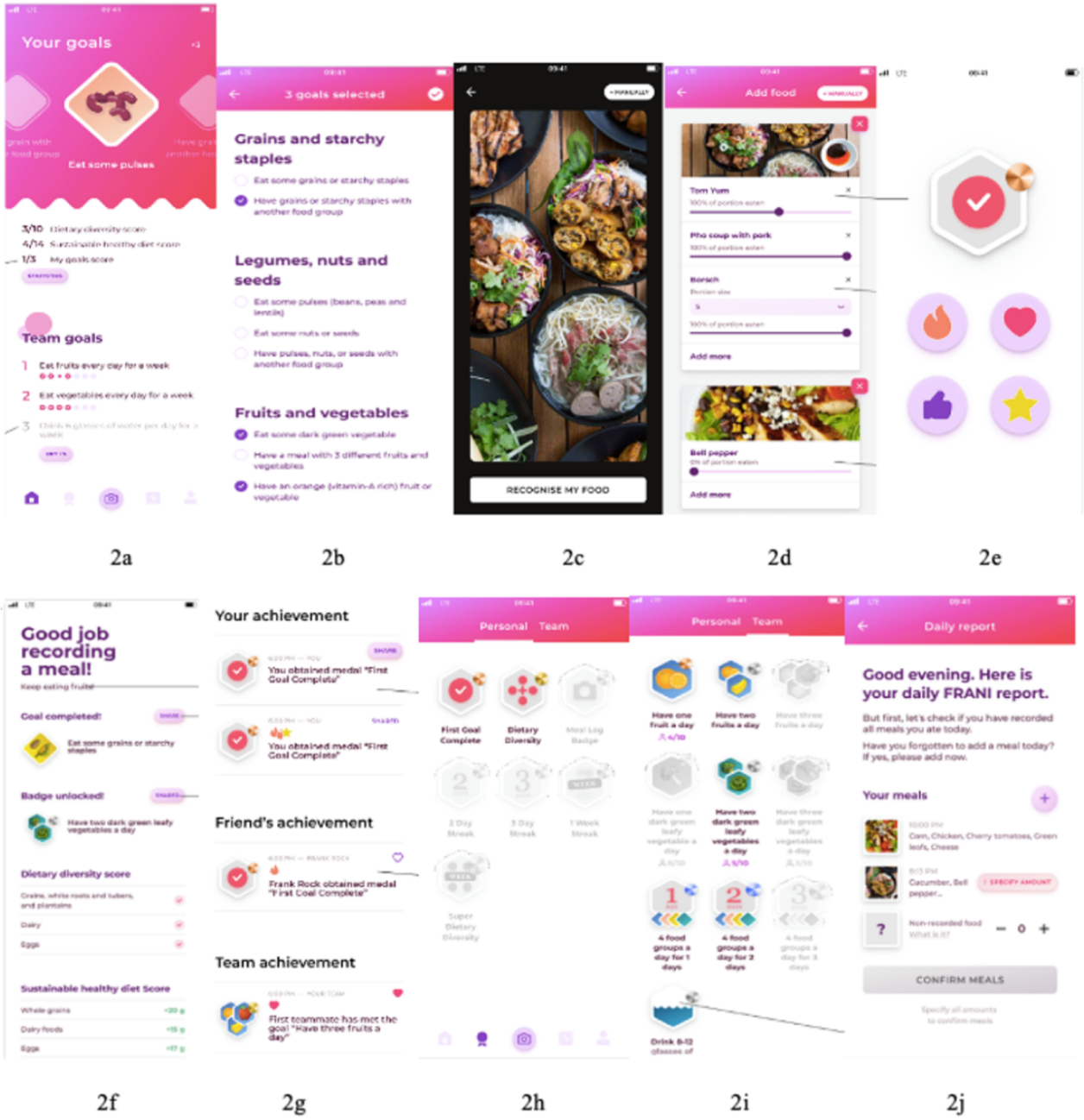
Acceptability of the Individual-Based Goals

Most participants had previously tried using various nutrition and health apps to track physical activity, diet, and health. However, they deleted these apps within a few weeks after downloading, owing to excessive battery consumption (ie, for counting steps) and time spent to complete the required tasks (ie, for logging food consumption). Adolescents preferred apps focused on behavior changes that would help them achieve immediate, potentially unhealthy fitness goals (ie, losing weight quickly) rather than focusing on health benefits (ie, improving dietary diversity). This is in line with the idea of *young invincibles* described in the literature: adolescents tend to make unhealthy food choices because of their low likelihood of developing diseases at their life stage [37]. The FRANI was carefully designed to avoid reinforcing this tendency. We opted to inspire continued use of the app by reinforcing the comprehension of the goals [34] and by increasing the saliency of healthy food choices.

Participants accepted most of the FRANI dietary goals related to vegetable, protein, and whole-grain consumption. The dietary goals that were not understood had their language changed in the updated version of the FRANI (Figure 2, Multimedia Appendix 3). Some participants preferred FRANI to assign goals based on medical history and anthropometric characteristics, instead of choosing dietary goals themselves. This is because they rightly did not feel knowledgeable about nutrition to make choices that would optimize health and fitness. Despite their lack of confidence in setting appropriate dietary goals, FRANI was designed with the expectation that achieving the goal set would improve dietary quality, as long as the dietary goals are changed daily, and goals are achieved [18]. Furthermore, the advantages of fostering habit formation by setting specific goals [38] could be undermined by automatizing this process. Therefore, the updated version of the FRANI does not suggest dietary goals for users. This was an example of incongruency between what users said they wanted and research objectives and theory, a problem that has been previously described in the literature [39].

Following the participants' suggestions, the updated version of FRANI's image recognition software estimated the weight of foods and beverages and provided feedback on quantities. However, goal setting was purely qualitative because maintaining simplicity may be important to the user's experience. The updated version of the FRANI also includes information pages with infographics and detailed explanations for food groups, goals, scores, and health implications of achieving goals. The incorporation of user goals and expectations should lead to higher motivation to use this technology and more effective future digital interventions [40,41].

Figure 2. Examples of the Food Recognition Assistance and Nudging Insights (FRANI) wireframes after changes based on feedback from the focus groups participants. (A) Home screen, which represents the individual- and team-based scores with brighter colors and the right denominators for the Dietary Diversity Score (DDS) and Sustainable Healthy Diet Score (SHDS). (B) Three of the food groups for which users will be able to choose goals (they can also choose goals for dairy and meat when scrolling down). (C) Wider camera frame. (D) Slide bars to digitally indicate how much users ate from foods recognized. (E) The confirmation screen with a bronze badge. (F) Completed goals, badges, and quantities. (G) Activity screen. (H) Individual-based badges. (I) Team-based badges with 1 design for each type of food. (J) The FRANI daily report sent at night for users so they can include complete or correct information uploaded throughout the day. There are English and Vietnamese options for FRANI.



Acceptability of Team-Based Dietary Goals and Competition

Team-based challenges were widely accepted, as they were seen as a means of self-improvement driven by competition. Participants suggested a form of within-team competition: FRANI should show individual achievement information to all team members and rank them accordingly. However, the updated version does not do so, because we want to avoid excessive peer pressure, given that changes in the brain and increased testosterone during this life stage have been associated with a greater sensitivity to social evaluation and influence [42].

Usability and Likability of the Camera

The updated version of FRANI has a wider camera frame that fits the whole family food tray and has a function to digitally separate the food portions eaten by users from the rest of the family meal. Concerning the difficulty of capturing the multiple ingredients of foods (eg, a sandwich) and foods inside packages, FRANI now also recognizes images of packages, and it is possible to manually log the ingredients that were not captured

by AI technology from the picture. The manual log-in of ingredients is not only more time consuming than taking pictures of foods but also increases the overall accuracy of the app.

Likability of the Feedback, Including Notifications, Badges, and Statistics

FRANI provides personally relevant feedback using credible nutrition information in the form of notifications, badges, and statistics. Participants demonstrated extensive interest in receiving personalized feedback from FRANI, which is important to keep users motivated and engaged [40,43], and it works better than *one size fits all* feedback [41,44]. Receiving feedback has been shown to have the potential to break undesired habits [45,46]. However, some adolescents wanted to receive negative feedback when recording unhealthy food.

In accordance with the desire of participants, the updated version of FRANI sends *just-in-time* notifications reminding users to take pictures at breakfast and lunchtime, which can lead to behavior change [47-49] and habit formation [50]. Notifications can now be turned on and off, but it is not yet possible to choose the time of the day and length of the notifications received. We chose not to send feedback disapproving unhealthy food choices or suggesting substitution for healthy foods, because it can be counterproductive and encourage unhealthy food consumption [51]. Adolescents are also particularly prone to eating disorders [52,53]; therefore, we tried to minimize the stress generated by negative feedback.

Study participants saw badges as an effective way to support behavior change, but the effects of badges, although promising, have not yet been teased out from other incentives in the nutrition behavior literature [31]. Using badges as rewards can be effective in encouraging positive responses, such as supporting desirable studying practices [33], increasing participation and engagement in schooling [54], supporting professional development [55], and stimulating voluntary contributions in web-based communities [56]. The participants were interested in self-monitoring through daily statistics and suggested that the FRANI should save historical dietary data. The updated version of the FRANI displays the dietary records.

Likability of the Graphic Design

In line with the ground theory, participants showed a preference for a simple interface with an appealing design [18,57-59]. The extent of the engagement benefits of eHealth and mobile health depends on the appropriate design that should be persuasive and personally relevant [60]. The updated interface was simple and visually appealing. The badges sent to praise achievements related to any food group during the FGDs had a watermelon design, which was confusing for the participants. The updated version of FRANI has badges with distinctive designs for accomplishments based on each food group (Figure 3).

Figure 3. Design of new badges based on remarks from the focus group discussions. Colors are brighter than in previous badge versions, food groups have different badges, and users are leveled up from bronze to silver to gold depending on what they achieved.

Strengths and Limitations

The main strength of this study was the in-depth qualitative assessment of the FRANI prototype. Although this assessment was based on the presentation of the prototype during the FGDs and not on real user experience, the process we describe herein can help other researchers build acceptable, usable, and likable tools for interventions. This may increase participation and adherence to future studies. As our study had a small sample size from a limited geographic location, the findings cannot be generalized to other countries, cultures, genders, and age groups.

Conclusions

Participants accepted FRANI and deemed it usable and likable conditionally on the following modifications: (1) tailored feedback based on users' medical history, anthropometric characteristics, and fitness goals; (2) new language on dietary goals; (3) provision of information about each of the food group's dietary goals; (4) wider camera frame to fit the whole

family food tray, as meals are shared in Vietnam; (5) possibility of digitally separating food consumption on shared meals; and (6) more appealing graphic design, including unique badge designs for each food group.

The findings on FGDs served as a guide to improve the FRANI prototype, which may also help other researchers design tools for different interventions. The acceptability, usability, and likability of the new version of FRANI, along with an assessment of the effects of using FRANI to improve food choices and dietary quality, will be quantitatively examined in a randomized controlled pilot study. To the best of our knowledge, FRANI is the first AI-based gamified self-monitoring app focused on improving the diets of adolescents in low- and middle-income countries. If successful, this tool will help to improve female adolescents' diets and will make high-frequency data collection on food consumption possible, minimize errors, decrease long-term research costs, and help fill the gap in adolescents' food consumption data collection [61].

Acknowledgments

This research was supported by a Fondation Botnar grant (REG 19-018) and by the Consultative Group for International Agricultural Research (CGIAR) program on Agriculture, Nutrition, and Health (A4NH). BCB is a PhD candidate supported by the Friedman Nutrition and Citizenship Fellowship from the Friedman School of Nutrition Science and Policy at Tufts University. The authors would like to thank Alexei Stulikov and Rohit Gangupantulu for their support in designing this prototype. The sponsors were not involved in the study design or in the collection, analysis, and interpretation of data or in the writing of the report and its submission to publication.

Data Availability

The data are available on request from the authors. Thus, a formal data sharing agreement is required. The lead author has full access to the data reported in this manuscript.

Authors' Contributions

AG, BCB, and N-LA designed this study. AG, BCB, BK, DH, and PM provided inputs for the wireframe design and prototype of the Food Recognition Assistance and Nudging Insights. FD programmed the prototype. AG, BCB, and BK wrote the focus group discussion (FGD) scripts. BCB and BK designed the discussion guide. GF designed the code book. PHN, LT, NH, and PH recruited the participants and ran the FGDs. BCB coded and analyzed the FGD transcripts. BCB wrote the manuscript. PNH, N-LA, BK, and AG revised the manuscript. All the authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary Figure 1.
[DOCX File , 374 KB - formative_v6i7e35197_app1.docx]

Multimedia Appendix 2
An explanation for the dietary quality indicators present in the Food Recognition Assistance and Nudging Insights.
[DOCX File , 16 KB - formative_v6i7e35197_app2.docx]

Multimedia Appendix 3
Supplementary Figure 2.
[DOCX File , 1111 KB - formative_v6i7e35197_app3.docx]

References

1. GBD 2017 Diet Collaborators. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019 May 11;393(10184):1958-1972 [FREE Full text] [doi: 10.1016/S0140-6736(19)30041-8] [Medline: 30954305]
2. Hall JN, Moore S, Harper SB, Lynch JW. Global variability in fruit and vegetable consumption. Am J Prev Med 2009 May;36(5):402-9.e5. [doi: 10.1016/j.amepre.2009.01.029] [Medline: 19362694]
3. Cancer risk factors. World Cancer Research Fund International. URL: https://www.wcrf.org/diet-and-cancer/exposures/ [accessed 2022-03-07]
4. Micha R, Shulkin ML, Peñalvo JL, Khatibzadeh S, Singh GM, Rao M, et al. Etiologic effects and optimal intakes of foods and nutrients for risk of cardiovascular diseases and diabetes: systematic reviews and meta-analyses from the Nutrition and

Chronic Diseases Expert Group (NutriCoDE). PLoS One 2017 Apr 27;12(4):e0175149 [FREE Full text] [doi: 10.1371/journal.pone.0175149] [Medline: 28448503]
5. Holder MD. The contribution of food consumption to well-being. Ann Nutr Metab 2019;74 Suppl 2:44-52 [FREE Full text] [doi: 10.1159/000499147] [Medline: 31234181]
6. Mujcic R, Oswald AJ. Evolution of well-being and happiness after increases in consumption of fruit and vegetables. Am J Public Health 2016 Aug;106(8):1504-1510. [doi: 10.2105/AJPH.2016.303260] [Medline: 27400354]
7. Tuck NJ, Farrow C, Thomas JM. Assessing the effects of vegetable consumption on the psychological health of healthy adults: a systematic review of prospective research. Am J Clin Nutr 2019 Jul 01;110(1):196-211. [doi: 10.1093/ajcn/nqz080] [Medline: 31152539]
8. Ebbeling CB, Feldman HA, Chomitz VR, Antonelli TA, Gortmaker SL, Osganian SK, et al. A randomized trial of sugar-sweetened beverages and adolescent body weight. N Engl J Med 2012 Oct 11;367(15):1407-1416 [FREE Full text] [doi: 10.1056/NEJMoa1203388] [Medline: 22998339]
9. Singh GM, Micha R, Khatibzadeh S, Shi P, Lim S, Andrews KG, Global Burden of Diseases Nutrition and Chronic Diseases Expert Group (NutriCoDE). Global, regional, and national consumption of sugar-sweetened beverages, fruit juices, and milk: a systematic assessment of beverage intake in 187 countries. PLoS One 2015 Aug 5;10(8):e0124845 [FREE Full text] [doi: 10.1371/journal.pone.0124845] [Medline: 26244332]
10. Qi Q, Chu AY, Kang JH, Jensen MK, Curhan GC, Pasquale LR, et al. Sugar-sweetened beverages and genetic risk of obesity. N Engl J Med 2012 Oct 11;367(15):1387-1396 [FREE Full text] [doi: 10.1056/NEJMoa1203039] [Medline: 22998338]
11. de Ruyter JC, Olthof MR, Seidell JC, Katan MB. A trial of sugar-free or sugar-sweetened beverages and body weight in children. N Engl J Med 2012 Oct 11;367(15):1397-1406. [doi: 10.1056/NEJMoa1203034] [Medline: 22998340]
12. Beal T, Morris SS, Tumilowicz A. Global patterns of adolescent fruit, vegetable, carbonated soft drink, and fast-food consumption: a meta-analysis of global school-based student health surveys. Food Nutr Bull 2019 Dec;40(4):444-459 [FREE Full text] [doi: 10.1177/0379572119848287] [Medline: 31617415]
13. Hargreaves D, Mates E, Menon P, Alderman H, Devakumar D, Fawzi W, et al. Strategies and interventions for healthy adolescent growth, nutrition, and development. Lancet 2022 Jan 08;399(10320):198-210. [doi: 10.1016/S0140-6736(21)01593-2] [Medline: 34856192]
14. Chen JL, Wilkosz ME. Efficacy of technology-based interventions for obesity prevention in adolescents: a systematic review. Adolesc Health Med Ther 2014 Aug 7;5:159-170 [FREE Full text] [doi: 10.2147/AHMT.S39969] [Medline: 25177158]
15. 81% of Teens Have Access to Gaming Consoles. Pew Research Center. 2015. URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2015/04/PI_2015-04-09_teensandtech_08.png [accessed 2021-07-16]
16. Anderson M, Jiang J. Teens, social media and technology 2018. Pew Research Center. 2018 May 31. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed 2021-07-16]
17. Wickham CA, Carbone ET. What's technology cooking up? A systematic review of the use of technology in adolescent food literacy programs. Appetite 2018 Jun 01;125:333-344. [doi: 10.1016/j.appet.2018.02.001] [Medline: 29471069]
18. Braga BC, Aberman NL, Arrieta A, Bannerman B, Burns A, Folson G, et al. Design of a mobile phone-based Artificial Intelligence (AI) application to assess dietary intake and provide nudges to improve healthy eating choices: formative research in Ghana and Vietnam. IFPRI Discussion Papers 2021:02024. [doi: 10.2499/p15738coll2.134412]
19. Taylor K, Silver L. Smartphone ownership is growing rapidly around the world, but not always equally. Pew Research Center. 2019 Feb 5. 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 2021-07-16]
20. Harris J, Nguyen PH, Tran LM, Huynh PN. Nutrition transition in Vietnam: changing food supply, food prices, household expenditure, diet and nutrition outcomes. Food Sec 2020 Sep 01;12(5):1141-1155. [doi: 10.1007/s12571-020-01096-x]
21. Rhyner D, Loher H, Dehais J, Anthimopoulos M, Shevchik S, Botwey RH, et al. Carbohydrate estimation by a mobile phone-based system versus self-estimations of individuals with type 1 diabetes mellitus: a comparative study. J Med Internet Res 2016 May 11;18(5):e101 [FREE Full text] [doi: 10.2196/jmir.5567] [Medline: 27170498]
22. Waki K, Aizawa K, Kato S, Fujita H, Lee H, Kobayashi H, et al. DialBetics with a multimedia food recording tool, FoodLog: smartphone-based self-management for type 2 diabetes. J Diabetes Sci Technol 2015 May;9(3):534-540 [FREE Full text] [doi: 10.1177/1932296815579690] [Medline: 25883164]
23. Konstantakopoulos F, Geroga EI, Klampanas K, Rouvalis D, Iaonnou N, Fotiadis DI. Automatic estimation of the nutritional composition of foods as part of the GlucoseML type 1 diabetes self-management system. In: Proceedings of the IEEE 19th International Conference on Bioinformatics and Bioengineering. 2019 Presented at: BIBE '19; October 28-30, 2019; Athens, Greece p. 470-473. [doi: 10.1109/bibe.2019.00091]
24. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
25. NVivo 11. QSR International. 2015. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home [accessed 2021-12-05]

26. Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, et al. Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems. Lancet 2019 Feb 02;393(10170):447-492. [doi: 10.1016/S0140-6736(18)31788-4] [Medline: 30660336]
27. Rapp A, Cena F. Personal informatics for everyday life: how users without prior self-tracking experience engage with personal data. Int J Hum Comput Stud 2016 Oct;94:1-17 [FREE Full text] [doi: 10.1016/j.ijhcs.2016.05.006]
28. Minimum dietary diversity for women – a guide to measurement. Food and Agriculture Organization of the United Nations. 2016. URL: https://www.fao.org/3/i5486e/i5486e.pdf [accessed 2021-03-17]
29. Knuppel A, Papier K, Key TJ, Travis RC. EAT-Lancet score and major health outcomes: the EPIC-Oxford study. Lancet 2019 Jul 20;394(10194):213-214. [doi: 10.1016/S0140-6736(19)31236-X] [Medline: 31235280]
30. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
31. Yoshida-Montezuma Y, Ahmed M, Ezezika O. Does gamification improve fruit and vegetable intake in adolescents? A systematic review. Nutr Health 2020 Dec;26(4):347-366. [doi: 10.1177/0260106020936143] [Medline: 32703067]
32. Hakulinen L, Auvinen T, Korhonen A. The effect of achievement badges on students' behavior: an empirical study in a university-level computer science course. Int J Emerg Technol Learn 2015 Feb 21;10(1):18-29 [FREE Full text] [doi: 10.3991/ijet.v10i1.4221]
33. O'Connor EA, McQuigge A. Exploring badging for peer review, extended learning and evaluation, and reflective/critical feedback within an online graduate course. J Educ Technol Syst 2014 Apr 04;42(2):87-105. [doi: 10.2190/et.42.2.b]
34. Zichermann G, Cunningham C. Gamification By Design: Implementing Game Mechanics in Web and Mobile Apps. Sebastopol, CA, USA: O'Reilly Media; 2011.
35. Mann T, de Ridder D, Fujita K. Self-regulation of health behavior: social psychological approaches to goal setting and goal striving. Health Psychol 2013 May;32(5):487-498. [doi: 10.1037/a0028533] [Medline: 23646832]
36. Thompson D. Talk to me, please!: the importance of qualitative research to games for health. Games Health J 2014 Jun;3(3):117-118. [doi: 10.1089/g4h.2014.0023] [Medline: 26196170]
37. Bibbins-Domingo K, Burroughs Peña M. Caring for the "young invincibles". J Gen Intern Med 2010 Jul;25(7):642-643 [FREE Full text] [doi: 10.1007/s11606-010-1388-8] [Medline: 20499197]
38. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
39. Vasalou A, Ingram G, Khaled R. User-centered research in the early stages of a learning game. In: Proceedings of the Designing Interactive Systems Conference. 2012 Presented at: DIS '12; June 11-15, 2012; Newcastle Upon Tyne, UK p. 116-125. [doi: 10.1145/2317956.2317976]
40. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: a systematic review. Health Educ Behav 2011 Feb;38(1):49-62. [doi: 10.1177/1090198110372878] [Medline: 21189422]
41. Busch M, Mattheiss E, Orji R, Marczewski A, Hochleitner W, Lankes M, et al. Personalization in serious and persuasive games and gamified interactions. In: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play. 2015 Presented at: CHI PLAY '15; October 5-7, 2015; London, UK p. 811-816. [doi: 10.1145/2793107.2810260]
42. Dahl RE. Adolescent brain development: a period of vulnerabilities and opportunities. Keynote address. Ann N Y Acad Sci 2004 Jun;1021:1-22. [doi: 10.1196/annals.1308.001] [Medline: 15251869]
43. Alkhaldi G, Hamilton FL, Lau R, Webster R, Michie S, Murray E. The effectiveness of prompts to promote engagement with digital interventions: a systematic review. J Med Internet Res 2016 Jan 08;18(1):e6 [FREE Full text] [doi: 10.2196/jmir.4790] [Medline: 26747176]
44. Gallien T, Oomen-Early J. Personalized versus collective instructor feedback in the online courseroom: does type of feedback affect student satisfaction, academic performance and perceived connectedness with the instructor? Int J E Learn 2008 Jul;7(3):463-476 [FREE Full text]
45. Hermsen S, Frost J, Renes RJ, Kerkhof P. Using feedback through digital technology to disrupt and change habitual behavior: a critical review of current literature. Comput Human Behav 2016 Apr;57:61-74 [FREE Full text] [doi: 10.1016/j.chb.2015.12.023]
46. Gamberini L, Spagnolli A, Corradi N, Jacucci G, Tusa G, Mikkola T, et al. Tailoring feedback to users' actions in a persuasive game for household electricity conservation. In: Proceedings of the 7th International Conference on Persuasive Technology: Design for Health and Safety. 2012 Presented at: PERSUASIVE '12; June 6-8, 2012; Linköping, Sweden p. 100-111. [doi: 10.1007/978-3-642-31037-9_9]
47. Fukuoka Y, Kamitani E, Bonnet K, Lindgren T. Real-time social support through a mobile virtual community to improve healthy behavior in overweight and sedentary adults: a focus group analysis. J Med Internet Res 2011 Jul 14;13(3):e49 [FREE Full text] [doi: 10.2196/jmir.1770] [Medline: 21752785]
48. Nguyen B, Shrewsbury VA, O'Connor J, Steinbeck KS, Lee A, Hill AJ, et al. Twelve-month outcomes of the loozit randomized controlled trial: a community-based healthy lifestyle program for overweight and obese adolescents. Arch Pediatr Adolesc Med 2012 Feb;166(2):170-177. [doi: 10.1001/archpediatrics.2011.841] [Medline: 22312175]

49. Partridge SR, Allman-Farinelli M, McGeechan K, Balestracci K, Wong AT, Hebden L, et al. Process evaluation of TXT2BFiT: a multi-component mHealth randomised controlled trial to prevent weight gain in young adults. Int J Behav Nutr Phys Act 2016 Jan 19;13:7 [FREE Full text] [doi: 10.1186/s12966-016-0329-2] [Medline: 26785637]
50. Stawarz K, Cox AL, Blandford A. Beyond self-tracking and reminders: designing smartphone apps that support habit formation. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Presented at: CHI '15; April 18-23, 2015; Seoul, South Korea p. 2653-2662. [doi: 10.1145/2702123.2702230]
51. Woolford SJ, Barr KL, Derry HA, Jepson CM, Clark SJ, Strecher VJ, et al. OMG do not say LOL: obese adolescents' perspectives on the content of text messages to enhance weight loss efforts. Obesity (Silver Spring) 2011 Dec;19(12):2382-2387 [FREE Full text] [doi: 10.1038/oby.2011.266] [Medline: 21869762]
52. Peterson K, Fuller R. Anorexia nervosa in adolescents: an overview. Nursing 2019 Oct;49(10):24-30. [doi: 10.1097/01.NURSE.0000580640.43071.15] [Medline: 31568077]
53. Volpe U, Tortorella A, Manchia M, Monteleone AM, Albert U, Monteleone P. Eating disorders: what age at onset? Psychiatry Res 2016 Apr 30;238:225-227 [FREE Full text] [doi: 10.1016/j.psychres.2016.02.048] [Medline: 27086237]
54. Diamond J, Gonzalez PC. Digital badges for teacher mastery: an exploratory study of a competency-based professional development badge system: CCT reports. Education Development Center, Center for Children & Technology. 2014 Nov. URL: https://files.eric.ed.gov/fulltext/ED561894.pdf [accessed 2021-08-31]
55. Gamrat C, Zimmerman HT, Dudek J, Peck K. Personalized workplace learning: an exploratory study on digital badging within a teacher professional development program. Br J Educ Technol 2014 Aug 20;45(6):1136-1148. [doi: 10.1111/bjet.12200]
56. Cavusoglu H, Li Z, Kim SH. How do virtual badges incentivize voluntary contributions to online communities? Inf Manag 2021 Jul;58(5):103483. [doi: 10.1016/j.im.2021.103483]
57. Nour MM, Rouf AS, Allman-Farinelli M. Exploring young adult perspectives on the use of gamification and social media in a smartphone platform for improving vegetable intake. Appetite 2018 Jan 01;120:547-556. [doi: 10.1016/j.appet.2017.10.016] [Medline: 29032184]
58. 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;20(1):151-171. [doi: 10.1111/bjhp.12114] [Medline: 25130682]
59. Rapp A, Tirassa M. Know thyself: a theory of the self for personal informatics. Human Comput Interact 2017 Apr 10;32(5-6):335-380. [doi: 10.1080/07370024.2017.1285704]
60. Nour M, Yeung SH, Partridge S, Allman-Farinelli M. A narrative review of social media and game-based nutrition interventions targeted at young adults. J Acad Nutr Diet 2017 May;117(5):735-52.e10. [doi: 10.1016/j.jand.2016.12.014] [Medline: 28238894]
61. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to Web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]

Abbreviations

AI: artificial intelligence
FGD: focus group discussion
FRANI: Food Recognition Assistance and Nudging Insights

Edited by A Mavragani; submitted 25.11.21; peer-reviewed by I Gabashvili, L Weinert, YL Leung; comments to author 21.02.22; revised version received 22.04.22; accepted 05.06.22; published 21.07.22.

Please cite as:
C Braga B, Nguyen PH, Aberman NL, Doyle F, Folson G, Hoang N, Huynh P, Koch B, McCloskey P, Tran L, Hughes D, Gelli A Exploring an Artificial Intelligence–Based, Gamified Phone App Prototype to Track and Improve Food Choices of Adolescent Girls in Vietnam: Acceptability, Usability, and Likeability Study
JMIR Form Res 2022;6(7):e35197
URL: https://formative.jmir.org/2022/7/e35197
doi:10.2196/35197
PMID:35862147

Original Paper

Web-Based Alcohol and Sexual Assault Prevention Program With Tailored Content Based on Gender and Sexual Orientation: Preliminary Outcomes and Usability Study of Positive Change (+Change)

Amanda K Gilmore[1], PhD; Ruschelle M Leone[1], PhD; Daniel W Oesterle[1,2], BS; Kelly Cue Davis[3], PhD; Lindsay M Orchowski[4], PhD; Viswanathan Ramakrishnan[5], PhD; Debra Kaysen[6], PhD

[1]Department of Health Policy & Behavioral Sciences, School of Public Health, Georgia State University, Atlanta, GA, United States
[2]Department of Psychological Sciences, Purdue University, West Lafayette, IN, United States
[3]Edson College of Nursing and Health Innovation, Arizona State University, Tempe, AZ, United States
[4]Department of Psychiatry & Human Behavior, Alpert Medical School of Brown University, Providence, RI, United States
[5]Medical University of South Carolina, Charleston, SC, United States
[6]Department of Psychiatry & Behavioral Sciences, Stanford University, Stanford, CA, United States

Corresponding Author:
Amanda K Gilmore, PhD
Department of Health Policy & Behavioral Sciences
School of Public Health
Georgia State University
PO Box 3995
Atlanta, GA, 30302
United States
Phone: 1 404 413 1416
Email: agilmore12@gsu.edu

Abstract

Background: Alcohol use and sexual assault are common on college campuses in the United States, and the rates of occurrence differ based on gender identity and sexual orientation.

Objective: We aimed to provide an assessment of the usability and preliminary outcomes of *Positive Change* (*+Change)*, a program that provides integrated personalized feedback to target alcohol use, sexual assault victimization, sexual assault perpetration, and bystander intervention among cisgender heterosexual men, cisgender heterosexual women, and sexual minority men and women.

Methods: Participants included 24 undergraduate students from a large university in the Southwestern United States aged between 18 and 25 years who engaged in heavy episodic drinking in the past month. All procedures were conducted on the web, and participants completed a baseline survey, *+Change*, and a follow-up survey immediately after completing *+Change*.

Results: Our findings indicated that *+Change* was acceptable and usable among all participants, despite gender identity or sexual orientation. Furthermore, there were preliminary outcomes indicating the benefit for efficacy testing of *+Change*.

Conclusions: Importantly, *+Change* is the first program to target alcohol use, sexual assault victimization, sexual assault perpetration, and bystander intervention within the same program and to provide personalized content based on gender identity and sexual orientation.

Trial Registration: ClinicalTrials.gov NCT04089137; https://clinicaltrials.gov/ct2/show/NCT04089137

(JMIR Form Res 2022;6(7):e23823) doi:10.2196/23823

KEYWORDS

sexual assault prevention; alcohol; college students; sexual and gender minorities

Introduction

Background

Alcohol use and sexual assault are widespread problems on college campuses in the United States [1,2], and the rates of occurrence differ based on gender identity and sexual orientation [3,4]. A total of 52.5% of college students used alcohol in the past month, and 33% engaged in heavy episodic drinking (>4 drinks for females and >5 drinks for males in <2 hours) [1,5]. Individuals who identify as a sexual or gender minority (lesbian, gay, bisexual, trans, queer or questioning [LGBTQ]) engage in alcohol misuse at higher rates than their heterosexual and cisgender counterparts [3]. Similarly, sexual assault is common among college students, with students who identify as cisgender heterosexual women and LGBTQ both experiencing the highest rates of sexual assault [4]. Despite alcohol interfering with sexual assault victimization risk perception, sexual assault victimization resistance [6] and bystander behavior [7], as well as increasing the risk for sexual assault perpetration [8], no intervention has targeted alcohol use and sexual assault victimization, perpetration, and bystander intervention in a single intervention. This is problematic, given that sexual assault is a multifaceted issue, and providing prevention for one component of sexual assault (victimization, perpetration, or bystander intervention) cannot address the full scope of sexual assault. Furthermore, providing one prevention component to men (ie, perpetration prevention) may send the message that men are not victims of assault or that it is not necessary to address all 3 components. Targeting multiple, related health issues, such as alcohol and sexual assault, is more effective than targeting them separately [9,10]. Furthermore, no program has provided tailored feedback for cisgender heterosexual men, cisgender heterosexual women, and LGBTQ students, despite their differential risk and differential risk factors. Therefore, this study presents the usability of a novel social norms–based intervention, *Positive Change* (*+Change)*, which targets alcohol and sexual assault (victimization, perpetration, and bystander intervention) and is tailored based on gender identity and sexual orientation.

Alcohol Use and Sexual Assault Among LGBTQ Individuals

Individuals who identify as LGBTQ are not a homogeneous group, and the drinking patterns of the LGBTQ community are, on average, equal to or greater than their cisgender heterosexual peers [11-15]. For example, one national study in the United States found that individuals who identify as lesbian or bisexual women and gay or bisexual men were 3.81 and 1.76 times more likely, respectively, to engage in high-intensity drinking (>12 drinks in a single drinking episode), compared with those who identify as heterosexual [3]. Another study indicated that women who identify as lesbian and bisexual were 10.7 times more likely to drink compared with women who identify as heterosexual [16]. This discrepancy in drinking behavior has been replicated in college students who identify as LGBTQ [14,17]. Drinking to cope with minority stress [18,19] and social norms of bar culture among individuals who identify as LGBTQ may explain the disparities in alcohol use among LGBTQ men and women [12,13,20]. According to the minority stress model [18,19], individuals who identify as LGBTQ experience daily heterosexist and transphobic aggressions and microaggressions that cause compounded stress over time. According to the negative reinforcement model, alcohol can be used to cope with stress. Therefore, drinking to cope with minority stress and LGBTQ-specific drinking norms are essential to address in interventions targeting alcohol use for LGBTQ individuals.

Lifetime rates of sexual assault victimization among LGBTQ individuals are higher than their cisgender heterosexual counterparts, with 63% of LGBTQ individuals experiencing sexual assault victimization [21]. Lesbian and bisexual women experience sexual assault victimization at higher rates and experience more mental health symptoms after sexual assault victimization, including higher rates of posttraumatic stress disorder and depression [22,23] compared with their heterosexual counterparts. Gay and bisexual men have also reported high rates of sexual assault victimization. A recent study found that 67% of gay and bisexual men reported an adult sexual assault victimization experience, and 67% of those sexual assaults involved alcohol [24]. This rate of sexual assault victimization is higher than that when examining national data sets of men in the United States. For example, the National Intimate Partner and Sexual Violence Survey found that 24.8% of men experienced sexual assault victimization in their lifetime [25]. LGBTQ populations may be uniquely targeted for sexual assault because of heterosexism. Furthermore, perceived normative behaviors related to resisting sexual assault may differ based on gender and sexual orientation. For example, as LGBTQ populations are disproportionately targeted for violence, they may believe that their peers would not support them if they used active resistance strategies when targeted for sexual assault. Little is known about bystander behaviors among students who identify as LGBTQ, but it is anticipated that there may be unique barriers to engaging in bystander intervention as a member of the LGBTQ community, mostly because LGBTQ populations are disproportionately targeted for violence.

Despite the heightened risk for heavy episodic drinking and sexual assault victimization [24,26-29], men and women who identify as LGBTQ are often overlooked in heavy episodic drinking and sexual assault prevention programs. One exception is an assessment of an in-person bystander intervention, *Green Dot*, which has been tested in high schools, and secondary data analyses examined whether the program was effective for individuals who identified as a sexual minority [30]. This program was not specifically adapted to the unique needs of LGBTQ students; however, it was found that there were reductions in victimization and perpetration among heterosexual youth but not sexual minority youth. Therefore, it is essential to adapt interventions made to prevent sexual assault to LGBTQ populations. Furthermore, it is essential to address this large gap by including alcohol use and sexual assault victimization risk reduction content specific to students who identify as LGBTQ and this study presents an initial step toward this needed effort.

Integrated Personalized Normative Feedback Interventions

The National Institute on Alcohol Abuse and Alcoholism Alcohol Intervention Matrix [31] recommends both personalized normative feedback and skills training as evidence-based interventions with low cost, high effectiveness, and high reach potential. Personalized normative feedback interventions target perceptions of normative drinking, which is the strongest predictor of alcohol use [32,33]. Specifically, personalized normative feedback interventions correct misperceptions of peer alcohol use by comparing one's own use with actual peer use as well as comparing one's perceptions of peer use with actual peer use. In a systematic review, 64% (25/39) of the trials found support for descriptive drinking norms as a mechanism of change in alcohol use interventions [34]. However, current approaches do not account for the unique drinking patterns of cisgender heterosexual women, cisgender heterosexual men, and LGBTQ men and women.

Social norms approaches for prevention allow for the prevention of cross-cutting behavior in an integrated manner as targeting social norms for multiple risk behaviors can be done within one theoretical framework. Furthermore, given that alcohol is associated with an increased risk for sexual assault victimization and perpetration and a decreased likelihood of engaging in a potential sexual assault situation as a bystander [7,8,35], it is essential to target alcohol use, sexual assault victimization, perpetration, and bystander intervention within 1 integrated program. This study included preliminary testing of *Positive Change (+Change)*, a multipronged personalized normative feedback intervention targeting alcohol use and sexual assault (victimization, perpetration, and bystander intervention) within 1 prevention program.

Sexual assault is still widely prevalent, and prevention efforts have not resulted in 0 perpetration rates; thus, feminist scholars emphasize harm reduction behaviors to reduce sexual assault risk while still placing the responsibility of the sexual assault solely with the perpetrator [36,37]. Theoretical models for sexual assault victimization risk reduction focus on providing skills for sexual assault risk perception and empowerment to resist sexual assault [37,38]. As social and psychological barriers can interfere with using active sexual assault resistance strategies, a social norms approach can be used to correct misperceptions of peer disapproval of using active resistance strategies [9]. Similarly, perpetration prevention can use a social norms approach [39,40] by targeting perceptions of peer rape supportive attitudes, beliefs, and behaviors that foster sexual assault perpetration [41]. This approach can also be used to target bystander intervention behaviors. Bystander intervention training encourages bystanders to engage in interventions when witnessing potential sexual assault situations. These programs are based heavily on the Social Norms Theory, as sexual assault perpetrators overestimate supportive peer attitudes toward sexual assault [42]. A recent meta-analysis of sexual assault bystander programs indicated that students who participated in these programs engaged in more bystander behaviors and had more prosocial attitudes, compared with those who did not [42]. Thus, a social norms approach combined with skills training can be a useful approach to address alcohol use and sexual assault prevention (victimization, perpetration, and bystander intervention) within 1 multipronged, comprehensive program.

Researchers have integrated sexual assault victimization risk, perpetration, and bystander intervention programs as a multipronged approach to prevent sexual assault on college campuses [43]; however, this integrated approach has not yet been implemented or tested. Furthermore, no program to date has targeted sexual assault prevention using this multipronged approach with integrated alcohol content. There are several advantages to addressing the needs of all college students within 1 program. First, it is costly for universities and time-consuming for students to provide 4 separate prevention programs. Second, excluding LGBTQ students from prevention programming is a form of heterosexism and may contribute to the continuing higher rates of alcohol and sexual assault among this population. Third, integrated intervention programming is more effective than providing separate interventions for related health behaviors [9].

This Study

This study assessed the usability and acceptability of *+Change*, a multipronged program targeting alcohol misuse, sexual assault victimization risk, sexual assault perpetration, and bystander intervention among cisgender heterosexual men, cisgender heterosexual women, and LGBTQ men and women. It was hypothesized that *+Change* would have high usability and acceptability ratings. Furthermore, although the study was not powered or designed to detect differences in outcomes related to indicators of alcohol use, sexual assault victimization, sexual assault perpetration, and bystander intervention, these indicators were examined to determine whether there was a decrease in alcohol use risk and sexual assault risk after *+Change*.

Methods

Participants

A total of 30 participants consented to and were enrolled in the study, and 24 (80%) participants fully viewed the content and postintervention questionnaire and were included in the analyses for this manuscript. The final sample included 24 undergraduate students aged between 18 and 25 years who engaged in heavy episodic drinking in the past month. Participants were recruited from a large university in the Southwestern United States.

Measures

Demographics

Participants completed items assessing age, race or ethnicity identity, year in college, and relationship status.

Gender Identity and Sexual Orientation

To assess gender identity, participants were asked the following question: "Understanding gender identity can be complex, which one category best describes your gender identity now?" Prior research has validated the use of this item to assess gender identity [44]. Responses included the following options: (1) female, (2) male, (3) transgender (female-to-male), (4) transgender (male-to-female), and (5) other. To assess sexual orientation, participants responded to the following question:

"Understanding that sexual identity can be complex, which one category best describes your sexual identity now?" This item has been used in previous research among sexual minority individuals [45]. Response options included the following: (1) lesbian, (2) gay, (3) bisexual, (4) queer, (5) two-spirit, (6) straight or heterosexual, (7) questioning, (8) other, and (9) prefer not to answer. Participants responding that they identified as male and straight were placed into the cisgender heterosexual male group. Participants who responded that they were female and straight were placed in the cisgender heterosexual female group. Finally, participants who responded that their sexual orientation was any response other than straight or heterosexual or that their gender was anything other than male or female were labeled in this study as LGBTQ (no participants in the sample identified as transgender or other).

Usability and Acceptability

Participants were asked 2 questions about intervention helpfulness for themselves ("How helpful did you find the intervention content?") and their peers ("How helpful do you think the intervention content would be for students at [your university]?") and one question about how distressed they were by the intervention ("How distressing did you find the intervention content?"). Participants rated these items on a scale from 1 (very unhelpful or not at all distressing) to 7 (very helpful or very distressing). Participants also responded to 2 instruments assessing the intervention's usability and functionality. To assess usability, participants responded to the Post-Study System Usability Questionnaire [46]. This 18-item instrument uses a Likert-type scale with response options ranging from 1 (strongly disagree) to 7 (strongly agree), where participants indicated their agreement with items such as "it was comfortable using this web-based intervention" and "it was easy to find the information I needed." The Post-Study System Usability Questionnaire includes subscale scores assessing system usefulness, information quality, and interface quality.

Alcohol Use

The Daily Drinking Questionnaire [47] was used to assess participants' alcohol consumption during a typical week. Participants were asked how many standard drinks they typically consume on each day of the week and then queried on the amount of time (ie, in hours) they typically consume that amount of alcohol. The Drinking Norms Rating Form [48] was used to assess normative perceptions of alcohol use among peers. This instrument assessed the perceived amount to which other students at their university in a particular group consume alcohol during a specific time frame for each day of the typical week. The peer groups assessed in this study included cisgender heterosexual women, cisgender heterosexual men, and LGBTQ, intersex, and asexual (LGBTQIA+) students at their university. The Injunctive Drinking Behaviors Scale [49] is a 15-item scale that was used to assess injunctive norms regarding the acceptability of drinking-related behaviors. Example items include rating how acceptable the typical student thinks it is to "drink shots" or to "drink alcohol every weekend." Participants responded using a Likert-type scale, where they rated each behavior from 1 (unacceptable) to 7 (acceptable), with higher scores indicating increased levels of acceptability for each behavior. The contemplation ladder [50] assesses stages of change for alcohol use on a 0 (I have no thoughts of changing my drinking now) to 10 (I'm taking action to change [ie, cutting down]) scale.

Sexual Assault Victimization Risk

To assess the perceived risk of sexual assault victimization while intoxicated, participants were asked how likely they would be incapacitated by alcohol while engaging in unwanted sex. Participants responded using a 7-point Likert-type scale, with response options ranging from 1 (very unlikely) to 7 (very likely). Participants were also asked to estimate the percentage (0 to 100) of each group (cisgender heterosexual men, cisgender heterosexual women, and LGBTQIA+ students at their university) who have experienced sexual assault victimization since entering college.

Sexual Assault Perpetration Risk

The Stages of Change Scale [51-53] assessed the participants' perceptions of sexual assault prevention efforts on campus. This instrument contains 8 items such as "I don't think sexual assault is a big problem on campus" and "I am actively involved in projects to deal with sexual assault on campus." The response options ranged from 1 (strongly disagree) to 4 (strongly agree), with higher scores indicating greater agreement with each item. The Illinois Rape Myth Acceptance Scale [54] was used to accept participant endorsement of 8 common rape myths on a Likert-type scale, with response options ranging from 1 (strongly disagree) to 4 (strongly agree), with high scores indicating greater agreement for each item. To assess the perceived risk for sexual assault perpetration, participants were asked how likely they would be to ask for verbal consent during sexual activity while drinking. To estimate the percentage of false reports at their university, participants were asked "What percent of sexual assaults are falsely reported at [your university]?" Participants were also asked how likely they would be to decide not to engage in sexual activity with someone who is drunk on a 6-point Likert-type scale, with response options ranging from 1 (not at all likely) to 6 (very likely).

Bystander Intervention

The Bystander Efficacy Scale [53] assessed participants' confidence in performing prosocial behaviors related to the prevention of sexual violence. Specifically, the Bystander Efficacy Scale is an 18-item measure, which contains items where participants rate their confidence in performing behaviors such as "ask a friend/stranger if they need to be walked home from a party" and "speak up to someone who is making excuses for forcing someone to have sex with them." Participants rated each of the 18 behaviors on a 0% (can't do) to 100% (very certain) scale as to their confidence in performing the corresponding behavior, with high scores indicating greater levels of confidence.

Procedure

Participants aged 18 to 25 years, who engaged in heavy episodic drinking in the past month, were recruited from a large university in the Southwestern United States using a random sample of students from the registrar by email. From a list of over 6000 students, 468 (7.8%) prospective participants were

randomly selected to receive an email inviting them to participate in this study. Of those 468 invited to participate in the screening survey, 41 (88%) participants were eligible, and 30 (6.4%) participants were enrolled in the open pilot trial. Of these 30 participants, 24 (80%) completed the study procedures. The participants were capped to ensure equal recruitment of cisgender heterosexual women, cisgender heterosexual men, and LGBTQIA+ students. Participants completed consent procedures, a baseline survey, a social norms–based personalized feedback intervention (*+Change*), and a postintervention survey. All study procedures were performed on the web. They were compensated with US $25 for their participation.

Ethics Approval

All study procedures were approved by the Georgia State University's institutional review board (H2006) and participants consented to all study procedures before participating in the study.

+Change Program Content

+Change included content from an integrated alcohol and sexual assault risk reduction program for women [9] and a web-based adaptation of a brief motivational interviewing personalized feedback protocol integrated with the men's workshop for sexual assault perpetration and bystander intervention for men [41,55]. Given that the previous interventions included separate content for men and women and did not address the needs of LGBTQ men and women, new content was created for men's victimization risk, women's perpetration risk, and women's bystander intervention skills training. Furthermore, content was created for the LGBTQ students. The intervention content underwent a rigorous process of intervention development where mockups were provided to college students (equal numbers of cisgender heterosexual men, cisgender heterosexual women, and sexual or gender minorities) and administrators for extensive feedback in interviews and focus groups, and iterative changes were made based on the feedback.

Participants received personalized feedback based on their answers to the baseline survey compared with a larger sample of college students at their university. Feedback was tailored by gender and sexual orientation (cisgender heterosexual men, cisgender heterosexual women, and LGBTQ individuals). *+Change* targeted heavy episodic drinking and sexual assault by integrating existing, theoretically driven [9,56-58], and evidence-based prevention initiatives delivered via a web-based platform: (1) social norms approach to reduce or prevent alcohol misuse, (2) programming to reduce sexual assault perpetration, (3) bystander intervention to all students tailored by gender and sexual orientation, and (4) sexual assault risk reduction programming tailored by gender and sexual orientation (Multimedia Appendix 1 [9,56-58]).

Analytic Plan

Independent 2-tailed *t* tests were used to examine preintervention and postintervention differences. Given the small sample size, we examined both the trends ($P<.10$ and $P>.05$) and significance ($P<.05$). Separate *t* tests were conducted for variables in the full sample and for each group examined (cisgender heterosexual men, cisgender heterosexual women, and sexual minority men and women).

Results

Demographics

Of the 24 participants, 8 (33%) identified as cisgender heterosexual men, 9 (38%) as cisgender heterosexual women, and 7 (29%) as LGBTQ (1 [4%] identified as gay, 4 [17%] as bisexual, 1 [4%] as queer, and 1 [4%] as questioning). Of the 7 participants who identified as LGBTQ, 6 (86%) identified as female and 1 (14%) identified as male. The mean age of participants was 19.63 (SD 0.97) years, and most were in their second year of school. Most participants identified as White (16/24, 59%), non-Hispanic, or non-Latinx (15/24, 63%), and reported being in a long-term monogamous relationship lasting at least 6 months (10/24, 42%). Participants reported engaging in 2.08 (SD 2.04) episodes of heavy episodic drinking (≥4 drinks for individuals assigned female at birth; ≥5 drinks for individuals assigned male at birth) per month on average. Furthermore, participants reported drinking 4.75 (SD 4.87) drinks per week on average. Sample characteristics are presented in Table 1. Of the 30 initially enrolled, 2 (7%) cisgender heterosexual men, 1 (3%) cisgender heterosexual woman, and 3 (10%) LGBTQ participants did not fully complete the study procedures and were excluded from the analyses.

Table 1. Demographic characteristics of participants in open pilot.

Demographics	Cisgender heterosexual men (n=8)	Cisgender heterosexual women (n=9)	LGBTQ[a] individuals (n=7)	Total sample (N=24)
Racial identity, n (%)				
Asian	0 (0)	0 (0)	2 (29)	2 (8)
Black	1 (13)	0 (0)	1 (14)	2 (8)
White	7 (87)	7 (78)	2 (29)	16 (67)
Multiracial	0 (0)	2 (22)	1 (14)	3 (13)
Other	0 (0)	0 (0)	1 (14)	1 (4)
Ethnicity, n (%)				
Hispanic or Latinx	1 (13)	4 (44)	3 (43)	8 (33)
Non-Hispanic or non-Latinx	7 (88)	5 (56)	3 (43)	15 (63)
In sorority or fraternity	1 (13)	1 (11)	0 (0)	2 (8)
Relationship status, n (%)				
Not dating	3 (38)	2 (22)	0 (0)	5 (21)
Casually dating	2 (25)	2 (22)	5 (71)	9 (38)
Involved in a long-term monogamous relationship	3 (38)	5 (56)	2 (29)	10 (42)
Age (years), mean (SD)	19.88 (1.13)	19.78 (0.97)	19.14 (0.69)	19.63 (0.97)
Years in college, mean (SD)	1.75 (0.87)	1.78 (0.97)	1.49 (0.79)	1.67 (0.87)
Heavy episodic drinking days per month, mean (SD)	3.00 (3.30)	1.56 (0.73)	1.71 (0.95)	2.08 (2.04)
Drinks per week, mean (SD)	7.12 (5.94)	3.77 (4.76)	3.28 (2.87)	4.75 (4.87)

[a]LGBTQ: lesbian, gay, bisexual, trans, queer or questioning.

Usability and Acceptability of +Change

On average, participants took 17.52 (SD 11.75) minutes to complete +*Change*. Overall, the participants were satisfied with the information quality (mean 5.52, SD 1.19), interface quality (mean 5.98, SD 1.05), and system usefulness (mean 5.74, SD 1.05) of +*Change*. In terms of +*Change*'s helpfulness, participants reported they found the intervention content helpful (mean 5.22, SD 1.51) and believed their peers would as well (mean 5.17, SD 1.47).

Alcohol Misuse

Results of *t* test analyses indicated that participants reported significant pre-post decreases in drinking norms, but did not report significant pre-post differences in injunctive drinking norms or in stages of change (Table 2).

Table 2. Changes in pre-post alcohol and sexual assault variables in open pilot.

Variables	Cisgender heterosexual men, mean (SD)		Cisgender heterosexual women, mean (SD)		LGBTQ[a], mean (SD)		Total sample			
	Before *+Change*	After *+Change*	Before *+Change*	After *+Change*	Before *+Change*	After *+Change*	Before *+Change*, mean (SD)	After *+Change*, mean (SD)	*t* test (*df*)	*P* value
Alcohol variables										
Descriptive norms	24.71 (16.50)	7.60 (3.78)	20.50 (9.21)	7.78 (6.16)	9.60 (3.36)	9.00 (0.00)	16.00 (7.01)	7.64 (5.24)	3.79 (20)	.002
Injunctive norms	4.99 (0.80)	5.19 (1.03)	5.62 (0.79)	5.36 (1.14)	5.13 (0.75)	4.93 (0.51)	5.25 (0.79)	5.16 (0.90)	0.45 (20)	.66
Contemplation ladder	.63 (1.06)	.75 (1.03)	4.33 (4.33)	5.13 (4.76)	.71 (1.89)	0.00 (0.00)	2.04 (3.33)	2.14 (3.65)	0.30 (22)	.77
Sexual assault victimization variables										
Likelihood of incapacitated sex	1.13 (0.35)	1.00 (0.00)	3.13 (1.46)	2.43 (2.30)	4.14 (2.04)	3.14 (2.12)	2.76 (1.89)	2.19 (1.94)	1.49 (20)	.15
Men's victimization estimate	5.71 (3.15)	14.29 (6.07)	16.14 (18.43)	16.00 (7.59)	24.00 (16.36)	22.60 (15.93)	14.37 (15.23)	17.17 (10.11)	−.59 (17)	.56
Women's victimization estimate	21.71 (13.38)	49.00 (15.82)	39.43 (30.18)	44.67 (20.22)	45.00 (15.81)	46.20 (11.41)	34.37 (22.85)	46.78 (15.58)	−1.86 (17)	.08
SGM's[b] victimization estimate	12.17 (13.79)	49.71 (12.41)	30.43 (25.88)	51.17 (25.99)	36.00 (14.75)	52.60 (12.40)	25.89 (21.18)	51.00 (17.05)	−3.95 (16)	.001
Sexual assault perpetration variables										
Precontemplation	2.62 (0.59)	2.10 (0.71)	2.19 (0.66)	1.71 (0.52)	2.00 (0.58)	1.47 (0.51)	2.29 (0.64)	1.78 (0.61)	4.40 (19)	>.001
Contemplation	2.19 (0.26)	2.71 (0.62)	2.97 (0.50)	3.00 (0.40)	2.94 (0.65)	3.14 (0.69)	2.59 (0.12)	2.94 (0.12)	−3.10 (19)	.006
Rape myths	1.33 (0.13)	1.19 (0.11)	1.34 (0.33)	1.27 (0.25)	1.11 (0.04)	1.03 (0.02)	1.27 (0.23)	1.18 (0.19)	2.50 (21)	.02
Estimated false reports	18.71 (16.83)	7.43 (7.91)	20.38 (22.17)	11.33 (8.36)	2.75 (1.50)	3.75 (1.50)	16.05 (18.34)	7.94 (7.39)	2.11 (15)	.05
While drinking, decide not to have sex with someone who is drunk	4.89 (1.76)	5.38 1.19)	3.67 (1.97)	5.60 (0.55)	3.67 (1.97)	5.60 (0.55)	4.15 (1.69)	5.00 (1.48)	−2.20 (20)	.04
Bystander variables										
Likelihood	4.07 (0.37)	4.29 (0.31)	4.09 (0.55)	4.23 (0.38)	4.24 (0.20)	4.55 (0.23)	4.12 (0.40)	4.34 (0.33)	−2.45 (21)	.02
Efficacy	83.86 (9.64)	85.89 (15.71)	87.04 (9.15)	88.31 (7.88)	91.47 (6.44)	96.05 (4.33)	87.03 (8.82)	89.40 (11.02)	−1.51 (20)	.15

[a]LGBTQ: lesbian, gay, bisexual, trans, queer or questioning.

[b]SGM: sexual and gender minority.

Sexual Assault Victimization Risk

Results of the *t* test analyses indicated that there was no significant pre-post difference in participants' estimations of the risk of experiencing incapacitated sexual assault victimization themselves while in college (Table 2). However, all participants significantly increased their estimation of how many LGBTQIA+ students had experienced sexual assault since entering college significantly after *+Change* (Table 2).

Sexual Assault Perpetration Risk

Results of the *t* test analyses indicated that after *+Change*, participants were significantly more aware of the problem of sexual assault on their campus (ie, less precontemplative) and

had greater intentions to make changes to prevent sexual assault perpetration (ie, more contemplative; Table 2). Participants reported significant reductions in pre-post rape myths and increases in deciding not to have sex with someone who was drunk (Table 2). In relation to false reports, participants at baseline estimated that, on average, 15.76% (SD 17.83%) of sexual assaults at their university were false reports. After participating in *+Change*, participants estimated that, on average, 7.94% (SD 7.39%) of sexual assaults at their university were false reports. This decrease in estimated false reports was a trend that did not reach significance (P=.052).

Bystander Intentions

Results of the *t* test analyses indicated that participants reported a significant preintervention to postintervention increase in the likelihood to intervene when witnessing sexual assault and nonsignificant increases in bystander efficacy following *+Change* (Table 2).

Discussion

Overview

This is the first program to provide personalized normative feedback to students who identify as LGBTQ and integrate multiple components of alcohol-related sexual assault prevention including victimization risk reduction, perpetration prevention, and bystander intervention training. Although a larger clinical trial to examine *+Change*'s efficacy is needed, these findings provide initial evidence that a comprehensive alcohol and sexual assault prevention program can be used among college students of varied genders and sexual orientations. Given the high rates of alcohol use and sexual assault among women and LGBTQ individuals, the rates of men as victims, and differences in perpetration rates, tailored content based on gender and sexual orientation are needed to move the prevention field forward.

Principal Findings

The results supported the usability and acceptability hypotheses such that *+Change* had high usability and acceptability ratings among cisgender heterosexual women, cisgender heterosexual men, and LGBTQ college students. Furthermore, despite the relatively low power to test for significant differences, there were some significant initial indicators suggesting that *+Change* may be helpful.

The findings from this study suggest that *+Change* may be an acceptable strategy to target alcohol and sexual assault among cisgender heterosexual men, cisgender heterosexual women, and LGBTQ college students. Overall, participants rated *+Change* to be acceptable across several usability domains including information quality, interface quality, and system usefulness. Furthermore, participants indicated that *+Change* was helpful for themselves and believed it would be helpful for their peers. This may be because of the brief duration of the prevention program and the user-friendly format of the web-based personalized feedback intervention.

There was a significant decrease in descriptive drinking norms after participating in *+Change*. Although differences between cisgender heterosexual men, cisgender heterosexual women, and LGBTQ individuals were not examined because of the small sample size, mean values suggest that the largest changes may have occurred within cisgender heterosexual men and cisgender heterosexual women. Although there were no significant changes in injunctive drinking norms or in stages of change in drinking, an examination of the means before and after *+Change* suggests that there were small changes in the direction toward lower injunctive drinking norms and higher motivations for change. These findings are similar to other social norms interventions targeting drinking among college students [34] and suggest that targeting descriptive drinking norms is a viable strategy when targeting both alcohol and sexual assault among college students.

There were significant increases in the awareness of sexual assault victimization risk among LGBTQ students during college. Specifically, participants estimated that more LGBTQ students experienced sexual assault during college at their university after *+Change* than before *+Change*. This is an important finding because awareness of sexual assault perpetrated against LGBTQ students could encourage bystander intervention behavior if a potential sexual assault is witnessed against an LGBTQ peer. This is also important because 17 years is the median age at which LGBTQ individuals begin to identify as LGBTQ (Pew Research Center [59]). Therefore, although individuals are identified as cisgender and heterosexual at the time of the intervention, their identity may change later in life. There was also a similar nonsignificant trend among female students. Although group comparisons were not assessed, an examination of the means suggested that LGBTQ students estimated that all college students experienced sexual assault during college more than other groups. This may be because LGBTQ students themselves have higher rates of sexual assault [21] and therefore, may be more aware of the risk for all college students. Nonetheless, the findings suggest that providing current rates of risk based on gender identity and sexual orientation can change one's perceived risk of experiencing sexual assault on a particular college campus.

Importantly, there were some indicators that *+Change* may have the potential to reduce sexual assault perpetration. Specifically, participants reported significantly less precontemplation and significantly more contemplation in terms of readiness to change sexual assault on their college campuses. In addition, participants reported decreased endorsement of rape myths and decreased estimate of how many sexual assault reports are false reports. Interestingly, before *+Change*, cisgender men and women believed that approximately one-fourth of reported sexual assaults at their university were false reports. This finding suggests that work is needed to change perceptions, given that only 5.9% of assaults are false accusations, the same rate as other crimes [60]. Furthermore, participants indicated increases in deciding not to have sex with someone who is drunk while they are drinking. This is an important behavioral intention indicator as it suggests that *+Change* may be helpful in reducing incapacitated sexual assault perpetration.

In terms of bystander intentions and attitudes, participants reported a significant increase in the likelihood to intervene when witnessing sexual assault and nonsignificant increases in

bystander efficacy. These results are promising, especially in light of the theory suggesting that intentions to perform a behavior are the closest cognitive antecedent of behavioral performance [61,62]. Furthermore, both longitudinal [63] and experimental [64] studies have found that bystander intentions predict subsequent bystander behavior for sexual assault.

Comparison With Prior Work

+Change is the first program to tailor content based on gender identity and sexual orientation. It is also the first program to integrate sexual assault victimization risk reduction, perpetration prevention, and bystander intervention training within one program. Therefore, this work extends previous research indicating that alcohol and sexual assault risks differ based on gender and sexual orientation [1-4], and previous calls for integrated programs for victimization risk reduction, perpetration prevention, and bystander intervention training [43]. Previous work has tested nontailored interventions of one component of sexual assault, such as the *Green Dot* which focuses on bystander intervention training, among sexual minority high-school students and found that the bystander content was not effective at reducing sexual assault among sexual minority youth [30]. Therefore, this study provides promising initial findings for an intervention that may address this large gap in prevention literature.

Limitations

This study had several limitations including the fact that it was an open pilot study of a small sample of college students at one university. Therefore, conclusions on the initial outcomes are only preliminary and a large-scale randomized clinical trial across multiple universities is needed. In addition, as no participants identified as gender-diverse, future research is needed to assess the efficacy of *+Change* among gender-diverse students. As the assessments were conducted on the same day, before and after *+Change*, they were not able to capture any behavioral changes. Future efficacy trials should examine whether *+Change* is effective at reducing alcohol use, sexual assault victimization, perpetration, and increasing bystander intervention behaviors to determine efficacy. This study assessed potential helpfulness and distress in assessing acceptability. Future research should use more in-depth acceptability measures. Future research is needed to understand Black, Indigenous, and People of Color students and LGBTQ Black, Indigenous, and People of Color students, as their experiences may differ and they could benefit from tailored interventions. LGBTQ students were included within one group rather than including different intervention components for the LGBTQ subgroups. This is problematic as LGBTQ students are not a heterogeneous group. However, this is a necessary first step in tailoring the programs for LGBTQ students. Although it appears in this sample that the LGBTQ participants did not engage in more alcohol use than their cisgender heterosexual peers, which is likely because all participants enrolled in this study engaged in heavy episodic drinking at least once in the past month. Research findings that LGBTQ men and women engage in drinking at higher levels than their cisgender heterosexual peers do not include a restricted sample as in this study and this sample likely has a ceiling effect because of the inclusion criteria. Although this study is limited to the United States, research indicates that sexual assault in higher education occurs at high rates across the globe [65]. Furthermore, alcohol-involved sexual assault is a global problem [66]. Future research should focus on the development of integrated prevention programs for alcohol use and sexual assault using culturally appropriate content worldwide.

Conclusions

+Change is the first program to integrate sexual assault victimization risk reduction, perpetration prevention, and bystander intervention training into one program. This is important to reduce both university costs and student time. It is also the first program to provide tailored content for LGBTQ students. This is important to acknowledge and address the unique risks of LGBTQ students as ignoring their needs may be a form of heterosexism embedded within the university prevention programming. Finally, it harnesses the strength of previous works [9,41]. The findings from this study suggest that *+Change* has high acceptability and usability among college students. Furthermore, there were several pre-post differences in outcomes related to alcohol use and sexual assault suggesting the need for a large randomized clinical trial.

Acknowledgments

Data collection was supported by a grant from the National Institute on Alcohol Abuse and Alcoholism (R34AA025691; principal investigator: AKG). Manuscript preparation was supported by grants from the National Institute on Drug Abuse and the National Institute on Alcohol Abuse and Alcoholism (K23DA042935; principal investigator: AKG and K01AA028844; principal investigator: RL).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Example positive change content.
[PDF File (Adobe PDF File), 78 KB - formative_v6i7e23823_app1.pdf]

References

1. 2019 National Survey of Drug Use and Health (NSDUH) releases. Substance Abuse and Mental Health Services Administration. URL: https://www.samhsa.gov/data/release/2019-national-survey-drug-use-and-health-nsduh-releases [accessed 2020-01-08]
2. Fedina L, Holmes JL, Backes BL. Campus sexual assault: a systematic review of prevalence research from 2000 to 2015. Trauma Violence Abuse 2018 Jan;19(1):76-93. [doi: 10.1177/1524838016631129] [Medline: 26906086]
3. Fish JN, Hughes TL, Russell ST. Sexual identity differences in high-intensity binge drinking: findings from a US national sample. Addiction 2018 Apr 02;113(4):749-758 [FREE Full text] [doi: 10.1111/add.14041] [Medline: 28940778]
4. Johnson LM, Matthews TL, Napper SL. Sexual orientation and sexual assault victimization among US college students. Social Sci J 2019 Dec 09;53(2):174-183. [doi: 10.1016/j.soscij.2016.02.007]
5. NIAAA council approves definition of binge drinking. NIAAA Newsletter. URL: https://pubs.niaaa.nih.gov/publications/Newsletter/winter2004/Newsletter_Number3.pdf [accessed 2021-04-30]
6. Norris J, Nurius PS, Dimeff LA. THROUGH HER EYES: factors affecting women's perception of and resestance to acquaintance sexual aggression threat. Psychol Women Q 1996 Mar 24;20(1):123-145 [FREE Full text] [doi: 10.1111/j.1471-6402.1996.tb00668.x] [Medline: 25705073]
7. Leone RM, Haikalis M, Parrott DJ, DiLillo D. Bystander intervention to prevent sexual violence: the overlooked role of bystander alcohol intoxication. Psychol Violence 2018 Oct;8(5):639-647 [FREE Full text] [doi: 10.1037/vio0000155] [Medline: 30505616]
8. Abbey A, Zawacki T, Buck PO, Clinton A, McAuslan P. Sexual assault and alcohol consumption: what do we know about their relationship and what types of research are still needed? Aggression Violent Behav 2004 May;9(3):271-303. [doi: 10.1016/s1359-1789(03)00011-9]
9. Gilmore AK, Lewis MA, George WH. A randomized controlled trial targeting alcohol use and sexual assault risk among college women at high risk for victimization. Behav Res Ther 2015 Nov;74:38-49. [doi: 10.1016/j.brat.2015.08.007]
10. Lewis MA, Patrick ME, Litt DM, Atkins DC, Kim T, Blayney JA, et al. Randomized controlled trial of a web-delivered personalized normative feedback intervention to reduce alcohol-related risky sexual behavior among college students. J Consult Clin Psychol 2014 Jul;82(3):429-440 [FREE Full text] [doi: 10.1037/a0035550] [Medline: 24491076]
11. Day JK, Fish JN, Perez-Brumer A, Hatzenbuehler ML, Russell ST. Transgender youth substance use disparities: results from a population-based sample. J Adolesc Health 2017 Dec;61(6):729-735 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.06.024] [Medline: 28942238]
12. Gilbert PA, Pass LE, Keuroghlian AS, Greenfield TK, Reisner SL. Alcohol research with transgender populations: a systematic review and recommendations to strengthen future studies. Drug Alcohol Depend 2018 May 01;186:138-146 [FREE Full text] [doi: 10.1016/j.drugalcdep.2018.01.016] [Medline: 29571076]
13. Hughes TL, Wilsnack SC, Kantor LW. The influence of gender and sexual orientation on alcohol use and alcohol-related problems: toward a global perspective. Alcohol Res 2016;38(1):121-132 [FREE Full text] [Medline: 27159819]
14. Reed E, Prado G, Matsumoto A, Amaro H. Alcohol and drug use and related consequences among gay, lesbian and bisexual college students: role of experiencing violence, feeling safe on campus, and perceived stress. Addict Behav 2010 Mar;35(2):168-171 [FREE Full text] [doi: 10.1016/j.addbeh.2009.09.005] [Medline: 19796880]
15. Talley AE, Gilbert PA, Mitchell J, Goldbach J, Marshall BD, Kaysen D. Addressing gaps on risk and resilience factors for alcohol use outcomes in sexual and gender minority populations. Drug Alcohol Rev 2016 Jul 13;35(4):484-493 [FREE Full text] [doi: 10.1111/dar.12387] [Medline: 27072658]
16. Ridner SL, Frost K, Lajoie AS. Health information and risk behaviors among lesbian, gay, and bisexual college students. J Am Acad Nurse Pract 2006 Aug;18(8):374-378. [doi: 10.1111/j.1745-7599.2006.00142.x] [Medline: 16907699]
17. Tupler LA, Zapp D, DeJong W, Ali M, O'Rourke S, Looney J, et al. Alcohol-related blackouts, negative alcohol-related consequences, and motivations for drinking reported by newly matriculating transgender college students. Alcohol Clin Exp Res 2017 May 21;41(5):1012-1023 [FREE Full text] [doi: 10.1111/acer.13358] [Medline: 28324915]
18. Hatzenbuehler ML. How does sexual minority stigma "get under the skin"? A psychological mediation framework. Psychol Bull 2009 Oct;135(5):707-730 [FREE Full text] [doi: 10.1037/a0016441] [Medline: 19702379]
19. Meyer IH. Minority stress and mental health in gay men. J Health Social Behav 1995 Mar;36(1):38. [doi: 10.2307/2137286]
20. Ebersole RC, Noble JJ, Madson MB. Drinking motives, negative consequences, and protective behavioral strategies in lesbian, gay, bisexual, and transgender college students. J LGBT Issues Counseling 2012 Oct;6(4):337-352. [doi: 10.1080/15538605.2012.725650]
21. Heidt JM, Marx BP, Gold SD. Sexual revictimization among sexual minorities: a preliminary study. J Trauma Stress 2005 Oct;18(5):533-540. [doi: 10.1002/jts.20061] [Medline: 16281251]
22. Sigurvinsdottir R, Ullman SE. Sexual orientation, race, and trauma as predictors of sexual assault recovery. J Fam Violence 2016 Oct 20;31(7):913-921 [FREE Full text] [doi: 10.1007/s10896-015-9793-8] [Medline: 27713597]
23. Sigurvinsdottir R, Ullman SE. Sexual assault in bisexual and heterosexual women survivors. J Bisex 2016 Mar 16;16(2):163-180 [FREE Full text] [doi: 10.1080/15299716.2015.1136254] [Medline: 27453694]
24. Hequembourg AL, Parks KA, Collins RL, Hughes TL. Sexual assault risks among gay and bisexual men. J Sex Res 2015 Jan 31;52(3):282-295 [FREE Full text] [doi: 10.1080/00224499.2013.856836] [Medline: 24483778]

25. Smith S, Zhang X, Basile K, Merrick M, Wang J, Kresnow M, et al. National intimate partner and sexual violence survey: 2015 data brief – updated release. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/violenceprevention/pdf/2015data-brief508.pdf [accessed 2021-04-30]
26. Hughes TL, Szalacha LA, Johnson TP, Kinnison KE, Wilsnack SC, Cho Y. Sexual victimization and hazardous drinking among heterosexual and sexual minority women. Addict Behav 2010 Dec;35(12):1152-1156 [FREE Full text] [doi: 10.1016/j.addbeh.2010.07.004] [Medline: 20692771]
27. Campus climate survey validation study final technical report. Bureau of Justice Statistics Research and Development Series. URL: https://www.bjs.gov/content/pub/pdf/ccsvsftr.pdf [accessed 2021-04-30]
28. Marshal MP, Friedman MS, Stall R, King KM, Miles J, Gold MA, et al. Sexual orientation and adolescent substance use: a meta-analysis and methodological review. Addiction 2008 May;103(4):546-556 [FREE Full text] [doi: 10.1111/j.1360-0443.2008.02149.x] [Medline: 18339100]
29. Marshal MP, Friedman MS, Stall R, Thompson AL. Individual trajectories of substance use in lesbian, gay and bisexual youth and heterosexual youth. Addiction 2009 Jul;104(6):974-981 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02531.x] [Medline: 19344440]
30. Coker AL, Bush HM, Clear ER, Brancato CJ, McCauley HL. Bystander program effectiveness to reduce violence and violence acceptance within sexual minority male and female high school students using a cluster RCT. Prev Sci 2020 Apr 06;21(3):434-444. [doi: 10.1007/s11121-019-01073-7] [Medline: 31907755]
31. Planning alcohol interventions using NIAAA's CollegeAim Alcohol Intervention Matrix. National Institute on Alcohol Abuse and Alcoholism. URL: https://www.niaaa.nih.gov/sites/default/files/publications/NIAAA_College_Matrix_Booklet.pdf [accessed 2021-04-30]
32. Neighbors C, Lee CM, Lewis MA, Fossos N, Larimer ME. Are social norms the best predictor of outcomes among heavy-drinking college students? J Stud Alcohol Drugs 2007 Jul;68(4):556-565 [FREE Full text] [doi: 10.15288/jsad.2007.68.556] [Medline: 17568961]
33. Perkins HW. Social norms and the prevention of alcohol misuse in collegiate contexts. J Stud Alcohol Suppl 2002 Mar(14):164-172. [doi: 10.15288/jsas.2002.s14.164] [Medline: 12022722]
34. Reid AE, Carey KB. Interventions to reduce college student drinking: state of the evidence for mechanisms of behavior change. Clin Psychol Rev 2015 Aug;40:213-224 [FREE Full text] [doi: 10.1016/j.cpr.2015.06.006] [Medline: 26164065]
35. Norris J, Masters NT, Zawacki T. Cognitive mediation of women's sexual decision making: the influence of alcohol, contextual factors, and background variables. Annu Rev Sex Res 2004;15:258-296. [Medline: 16913281]
36. Gidycz C. Sexual assault risk reduction: Current state and historical underpinnings. In: Sexual Assault Risk Reduction and Resistance Theory, Research, and Practice. San Diego, CA: Academic Press; 2018.
37. Rozee PD, Koss MP. Rape: a century of resistance. Psychol Women Q 2016 Jun 24;25(4):295-311. [doi: 10.1111/1471-6402.00030]
38. Nurius PS, Norris J. A cognitive ecological model of women's response to male sexual coercion in dating. J Psychol Human Sexuality 1996 Jul 11;8(1-2):117-139. [doi: 10.1300/j056v08n01_09]
39. Men and Rape Theory, Research, and Prevention Programs in Higher Education. San Francisco, California, United States: Jossey-Bass; 1994.
40. Berkowitz A. College men as perpetrators of acquaintance rape and sexual assault: a review of recent research. J Am Coll Health 1992 Jan;40(4):175-181. [doi: 10.1080/07448481.1992.9936279] [Medline: 1583239]
41. Orchowski L, Barnett N, Berkowitz A, Borsari B, Oesterle D, Zlotnick C. Sexual assault prevention for heavy drinking college men: development and feasibility of an integrated approach. Violence Against Women 2018 Sep;24(11):1369-1396. [doi: 10.1177/1077801218787928] [Medline: 30078368]
42. Jouriles EN, Krauss A, Vu NL, Banyard VL, McDonald R. Bystander programs addressing sexual violence on college campuses: a systematic review and meta-analysis of program outcomes and delivery methods. J Am Coll Health 2018 Mar 12;66(6):457-466. [doi: 10.1080/07448481.2018.1431906] [Medline: 29405865]
43. Orchowski LM, Edwards KM, Hollander JA, Banyard VL, Senn CY, Gidycz CA. Integrating sexual assault resistance, bystander, and men's social norms strategies to prevent sexual violence on college campuses: a call to action. Trauma Violence Abuse 2020 Oct 11;21(4):811-827. [doi: 10.1177/1524838018789153] [Medline: 30205767]
44. Reisner SL, Conron KJ, Tardiff LA, Jarvi S, Gordon AR, Austin SB. Monitoring the health of transgender and other gender minority populations: validity of natal sex and gender identity survey items in a U.S. national cohort of young adults. BMC Public Health 2014 Dec 26;14(1):1224 [FREE Full text] [doi: 10.1186/1471-2458-14-1224] [Medline: 25427573]
45. Dworkin ER, Kaysen D, Bedard-Gilligan M, Rhew IC, Lee CM. Daily-level associations between PTSD and cannabis use among young sexual minority women. Addict Behav 2017 Nov;74:118-121 [FREE Full text] [doi: 10.1016/j.addbeh.2017.06.007] [Medline: 28618391]
46. Lewis JR. Psychometric evaluation of the post-study system usability questionnaire: the PSSUQ. Proc Human Factors Society Annual Meeting 2016 Aug 06;36(16):1259-1260. [doi: 10.1177/154193129203601617]
47. Collins RL, Parks GA, Marlatt GA. Social determinants of alcohol consumption: the effects of social interaction and model status on the self-administration of alcohol. J Consult Clin Psychol 1985;53(2):189-200. [doi: 10.1037//0022-006x.53.2.189]

48. Lewis MA, Neighbors C. An examination of college student activities and attentiveness during a web-delivered personalized normative feedback intervention. Psychol Addict Behav 2015 Mar;29(1):162-167 [FREE Full text] [doi: 10.1037/adb0000003] [Medline: 25134036]
49. Lewis MA, Neighbors C, Geisner IM, Lee CM, Kilmer JR, Atkins DC. Examining the associations among severity of injunctive drinking norms, alcohol consumption, and alcohol-related negative consequences: the moderating roles of alcohol consumption and identity. Psychol Addict Behav 2010 Jul;24(2):177-189 [FREE Full text] [doi: 10.1037/a0018302] [Medline: 20565144]
50. Biener L, Abrams DB. The contemplation ladder: validation of a measure of readiness to consider smoking cessation. Health Psychol 1991;10(5):360-365. [doi: 10.1037/0278-6133.10.5.360]
51. Potter SJ. Using a multimedia social marketing campaign to increase active bystanders on the college campus. J Am Coll Health 2012 May 04;60(4):282-295. [doi: 10.1080/07448481.2011.599350] [Medline: 22559087]
52. Potter SJ, Moynihan MM, Stapleton JG, Banyard VL. Empowering bystanders to prevent campus violence against women: a preliminary evaluation of a poster campaign. Violence Against Women 2009 Jan 01;15(1):106-121. [doi: 10.1177/1077801208327482] [Medline: 19052283]
53. Banyard VL, Eckstein RP, Moynihan MM. Sexual violence prevention: the role of stages of change. J Interpers Violence 2010 Jan 27;25(1):111-135. [doi: 10.1177/0886260508329123] [Medline: 19252067]
54. Payne DL, Lonsway KA, Fitzgerald LF. Rape myth acceptance: exploration of its structure and its measurement using theillinois rape myth acceptance scale. J Res Personality 1999 Mar;33(1):27-68. [doi: 10.1006/jrpe.1998.2238]
55. Orchowski L, Barnett N, Oesterle D, Zlotnick C, Merrill J, Wood M. Randomized pilot trial of an integrated alcohol and sexual assault for heavy drinking college men: six-month findings. Alcohol Clin Experiment Res 2016;40:253A [FREE Full text]
56. Baer JS, Stacy A, Larimer M. Biases in the perception of drinking norms among college students. J Stud Alcohol 1991 Dec;52(6):580-586. [doi: 10.15288/jsa.1991.52.580] [Medline: 1758185]
57. Steele CM, Josephs RA. Alcohol myopia: its prized and dangerous effects. Am Psychologist 1990;45(8):921-933. [doi: 10.1037/0003-066x.45.8.921]
58. Goldman M, Del Boca F. Alcohol expectancy theory. In: Psychological Theories of Drinking and Alcoholism, 2nd edition. New York: Guilford; 1999.
59. Social trends. Pew Research Center. URL: https://www.pewresearch.org/social-trends/2013/06/13/chapter-3-the-coming-out-experience/%20accessed%2008/01/2020 [accessed 2020-01-08]
60. Lisak D, Gardinier L, Nicksa SC, Cote AM. False allegations of sexual assualt: an analysis of ten years of reported cases. Violence Against Women 2010 Dec 16;16(12):1318-1334. [doi: 10.1177/1077801210387747] [Medline: 21164210]
61. Hill RJ, Fishbein M, Ajzen I. Belief, attitude, intention and behavior: an introduction to theory and research. Contemporary Sociol 1977 Mar;6(2):244. [doi: 10.2307/2065853]
62. Ajzen I. The theory of planned behavior. Org Behav Human Decision Processes 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-t]
63. Banyard VL. Measurement and correlates of prosocial bystander behavior: the case of interpersonal violence. Violence Vict 2008 Feb 01;23(1):83-97. [doi: 10.1891/0886-6708.23.1.83] [Medline: 18396583]
64. Leone RM, Parrott DJ. Acute alcohol intoxication inhibits bystander intervention behavior for sexual aggression among men with high intent to help. Alcohol Clin Exp Res 2019 Jan 30;43(1):170-179 [FREE Full text] [doi: 10.1111/acer.13920] [Medline: 30500086]
65. Sivertsen B, Nielsen MB, Madsen IEH, Knapstad M, Lønning KJ, Hysing M. Sexual harassment and assault among university students in Norway: a cross-sectional prevalence study. BMJ Open 2019 Jun 09;9(6):e026993 [FREE Full text] [doi: 10.1136/bmjopen-2018-026993] [Medline: 31182445]
66. Intimate partner violence and alcohol. World Health Organization. URL: https://www.who.int/violence_injury_prevention/violence/world_report/factsheets/fs_intimate.pdf [accessed 2020-01-08]

Abbreviations

LGBTQ: lesbian, gay, bisexual, trans, queer or questioning
LGBTQIA+: lesbian, gay, bisexual, trans, queer or questioning, intersex, and asexual

Edited by A Mavragani; submitted 24.08.20; peer-reviewed by R Anderson, H Daniela, D Sugarman; comments to author 12.01.21; revised version received 30.04.21; accepted 19.05.22; published 22.07.22.

Please cite as:
Gilmore AK, Leone RM, Oesterle DW, Davis KC, Orchowski LM, Ramakrishnan V, Kaysen D
Web-Based Alcohol and Sexual Assault Prevention Program With Tailored Content Based on Gender and Sexual Orientation: Preliminary Outcomes and Usability Study of Positive Change (+Change)
JMIR Form Res 2022;6(7):e23823
URL: https://formative.jmir.org/2022/7/e23823
doi:10.2196/23823
PMID:35867393

Original Paper

Hemodialysis Record Sharing: Solution for Work Burden Reduction and Disaster Preparedness

Keisuke Ido[1], MPH; Mariko Miyazaki[2], MD, PhD; Masaharu Nakayama[3], MD, PhD

[1]Medical Information Center, Tohoku University Hospital, Sendai, Japan

[2]Division of Nephrology, Endocrinology and Vascular Medicine, Tohoku University Hospital, Sendai, Japan

[3]Department of Medical Informatics, Tohoku University Graduate School of Medicine, Sendai, Japan

Corresponding Author:
Masaharu Nakayama, MD, PhD
Department of Medical Informatics
Tohoku University Graduate School of Medicine
2-1 Seiryo-machi
Aoba-ku
Sendai, 9808575
Japan
Phone: 81 227177572
Fax: 81 227177505
Email: nakayama@cardio.med.tohoku.ac.jp

Abstract

Background: After the Great East Japan Earthquake in 2011, backup systems for clinical information were launched in Japan. The system in Miyagi Prefecture called the Miyagi Medical and Welfare Information Network (MMWIN) is used as a health information exchange network to share clinical information among various medical facilities for patients who have opted in. Hospitals and clinics specializing in chronic renal failure require patients' data and records during hemodialysis to facilitate communication in daily clinical activity and preparedness for disasters.

Objective: This study aimed to facilitate the sharing of clinical data of patients undergoing hemodialysis among different hemodialysis facilities.

Methods: We introduced a document-sharing system to make hemodialysis reports available on the MMWIN. We also recruited hospitals and clinics to share the hemodialysis reports of their patients and promoted the development of a network between emergency and dialysis clinics.

Results: In addition to basic patient information as well as information on diagnosis, prescription, laboratory data, hospitalization, allergy, and image data from different facilities, specific information about hemodialysis is available, as well as a backup of indispensable information in preparation for disasters. As of June 1, 2021, 12 clinics and 10 hospitals of 68 dialysis facilities in Miyagi participated in the MMWIN. The number of patients who underwent hemodialysis in Miyagi increased by more than 40%.

Conclusions: Our backup system successfully developed a network of hemodialysis facilities. We have accumulated data that are beneficial to prevent the fragmentation of patient information and would be helpful in transferring patients efficiently during unpredictable disasters.

(JMIR Form Res 2022;6(7):e32925) doi:10.2196/32925

KEYWORDS

hemodialysis; electronic health record; EHR; information sharing; information exchange; data sharing; document sharing; health information exchange; disaster; work burden; clinical information; clinical data; clinical report; medical report; information network; medical informatics; renal failure; kidney; renal; clinical record; medical record; backup; data security; data protection; data recovery

Introduction

Patients undergoing hemodialysis (HD) require collaborative team care, not only because their renal function is poor but also because they are vulnerable to various comorbidities [1,2]. Since an increase in the number of comorbidities makes care for patients more complex, multiple medical facilities may need to share patient information, and diverse specialist expertise may be required for patient management. However, sharing clinical information among different facilities is not an easy task because most dialysis clinics and hospitals do not have the same electronic health record (EHR) or do not have a health information exchange (HIE). Furthermore, the EHR and HD systems are rarely integrated. Such fragmentation of clinical information becomes a more serious concern during natural disasters such as earthquakes, tsunamis, and hurricanes. In 2011, the Great East Japan Earthquake (GEJE) hit Miyagi Prefecture [3,4] and damaged many facilities, including HD clinics and hospitals, resulting in the transfer of patients undergoing HD from Miyagi to other distant places [5]. At that time, the importance of patient information was recognized through experience. In addition, Katrina, a hurricane that hit the United States in 2005, caused similar situations in which many patients could not receive optimal dialysis treatments at alternate dialysis institutions because of a lack of medical records and information about their dialysis regimen [6]. Thus, sharing patient information is critical for disaster preparedness.

After the GEJE, backup systems of clinical data were developed in many places in Japan, and Miyagi Prefecture launched the Miyagi Medical and Welfare Information Network (MMWIN), which saves patient clinical information obtained from medical facilities, including hospitals, clinics, pharmacies, and nursing homes [7]. Moreover, the MMWIN is now used as an HIE to share clinical information, such as patients' basic information, history of diagnosis, prescription data, laboratory test data, and hospitalization data, among more than 800 facilities in Miyagi Prefecture [8]. This study aimed to facilitate the sharing of clinical data of patients undergoing HD among HD facilities.

Methods

The MMWIN System and Data Storage

The system used in MMWIN was explained in detail in our previous paper [7]. Briefly, the main system of the MMWIN consists of backup storage, a portal server to manage patients' ID registration information, the gateway that receives clinical data from facilities and stores them in the storage, and viewer applications. A standard storage format, Standardized Structured Medical Information eXchange version 2 (SS-MIX2), which is authorized by the Ministry of Health, Labour and Welfare of the Japanese government, enables the collection of clinical information from different vendor systems and is commonly used in several national projects to store and use large amounts of clinical data in Japan [9-12]. The storage format is divided into two categories. The first is the standardized storage, including standard clinical data in a standard form (Health Level 7 [HL7] v2.5), such as basic patient data, prescriptions, and laboratory data. The second is the extension storage that includes the remaining data not stored in the standard storage [12].

Computers in the MMWIN-affiliated facilities are connected to the MMWIN server via a virtual private network (MMWIN secure network). Clinical data from the hospital information system (HIS; or other information systems available in clinics, pharmacies, or nursing care homes) are transferred to SS-MIX2–formatted XML/HL7 files via gateways.

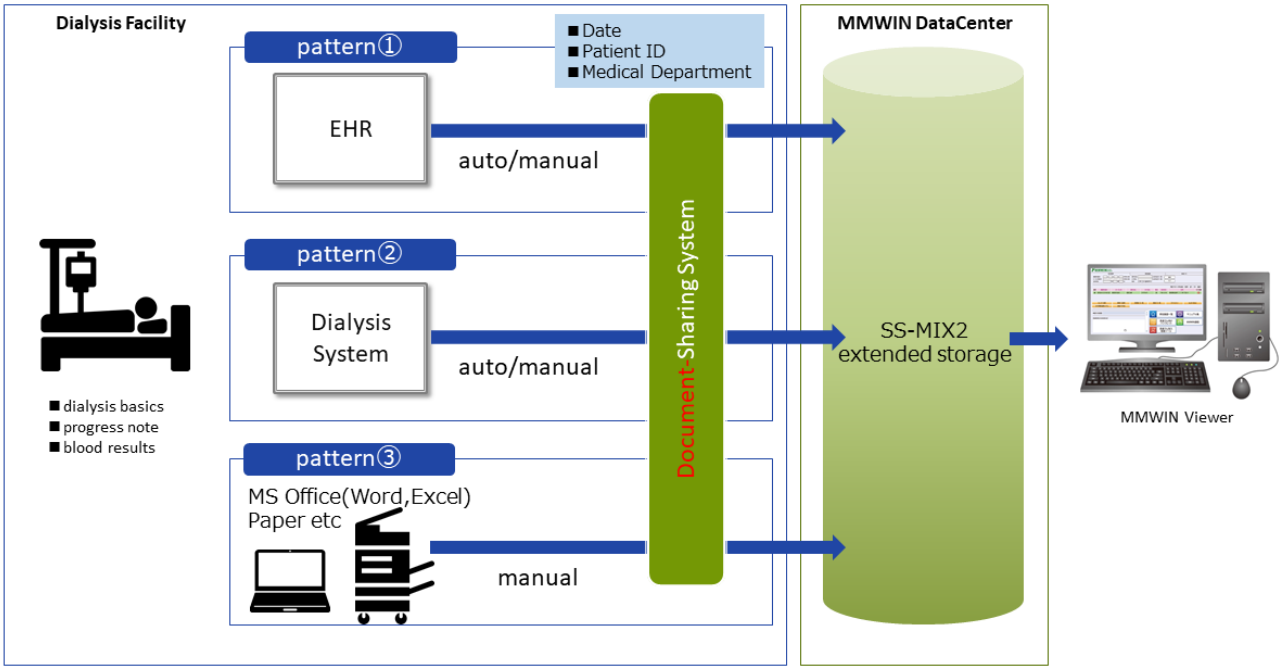
Opt-in for patients to register with the MMWIN includes a unique MMWIN ID number, which is connected to a different ID number in each facility, thereby permitting the sharing of patient information. Finally, these data are available through a browser-based SS-MIX2 viewer application, Human Bridge (Fujitsu). Thus, in each facility, the clinical information of all patients in the facility can be seen on MMWIN, which allows clinicians to check clinical data even when their own information system is out of order.

Data Transfer of HD Reports

Since there is no standard format for HD reports, we started storing the reports in extension storage using Docuworks (FUJIFILM Business Innovation Corp.), which enables the transfer of HD reports from the facilities to the MMWIN center. Figure 1 shows the scheme of the system for sharing HD reports between the MMWIN and facilities.

There are three ways to transfer HD information from facilities. First, HD reports could be integrated into the HIS and automatically or manually transferred to the SS-MIX2 center in the MMWIN. Second, HD reports could be released from the dialysis system and automatically or manually transferred to the SS-MIX2 center in the MMWIN. Third, a facility that still uses a paper record of HD or the original format using Word or Excel (Microsoft Corporation) can transfer the data manually via our document-sharing system. Through Docuworks, these files with additional necessary patient information, such as patient ID and labels as HD reports, are converted to SS-MIX2 format and are transferred to the SS-MIX2 extension storage in MMWIN through the facility's gateway and MMWIN secure network.

Figure 1. A scheme of connection between the MMWIN data center and the hemodialysis (HD) report in each facility. Three sources of HD information from facilities are shown: HD reports integrated with the hospital information system and transferred to the SS-MIX2 center in the MMWIN; HD reports released from the dialysis system and transferred to the SS-MIX2 center in the MMWIN; and a manual transfer of HD data via our document-sharing system. EHR: electronic health record; MMWIN: Miyagi Medical and Welfare Information Network; MS: Microsoft; SS-MIX2: Standardized Structured Medical Information eXchange version 2.



Ethics Approval

The project was approved by the administrative boards of MMWIN on December 25, 2015.

System Promotion

To promote the use of the system, we tried to encourage the development of a network within the clinical communities of HD clinics and hospitals. We asked nephrologists in Miyagi Prefecture to use this system and explained the benefits, such as being paperless and reducing the work burden of medical staff in HD facilities, through information brochures, personal meetings, and seminars since 2016. We also encouraged patients undergoing HD to provide consent to share their clinical information including prescription, laboratory data, and HD reports with the other HD facilities.

Statistical Analysis

A chi-square test and the Welch *t* test were used for statistical analysis using R 3.6.0 (R Foundation for Statistical Computing). Two-sided *P* values <.05 were considered significant.

Results

We introduced a document-sharing system to make HD reports available on the MMWIN. Figure 2 shows a screenshot of the HD report on the MMWIN. MMWIN users can refer to data regarding HD as well as a variety of information, such as basic patient information, diagnosis, prescription, laboratory data, hospitalization, allergy information, and image data from different facilities. When a user clicks the icon, detailed information is obtained. Next, we compared the amount of information available on the MMWIN to the items that were required during a disaster, as recommended in previous reports [13,14]. Table 1 shows that 39 items are available in the MMWIN, and these covered 92% (23/25) of recommendation 1 [13] and 100% (19/19) of recommendation 2 [14].

According to the network linkage of patient transfer between core hospitals and HD facilities, we recruited hospitals and clinics sharing HD reports and promoted the development of a network between emergency hospitals and dialysis clinics. As of June 1, 2021, 12 clinics and 10 hospitals of the 68 dialysis facilities in Miyagi participated in the MMWIN. The basic characteristics of the facilities with or without HD are shown in Table 2. There were no significant differences among them, except for the number of beds in clinics, with a smaller number of beds in the facilities using the MMWIN-HD system (*P*=.02). We also asked patients for their consent to share clinical information among facilities. Figure 3 shows the rate of opt-in patients (line chart) and the total number of all patients undergoing HD in Miyagi Prefecture (bar chart). The number of opt-in patients increased to more than 40% of all patients who underwent HD in Miyagi Prefecture after we began recruitment.

Figure 2. A screenshot of the Miyagi Medical and Welfare Information Network viewer and hemodialysis report. Users can access information on dialysis regimen, the latest body weight, and vital signs by this view.

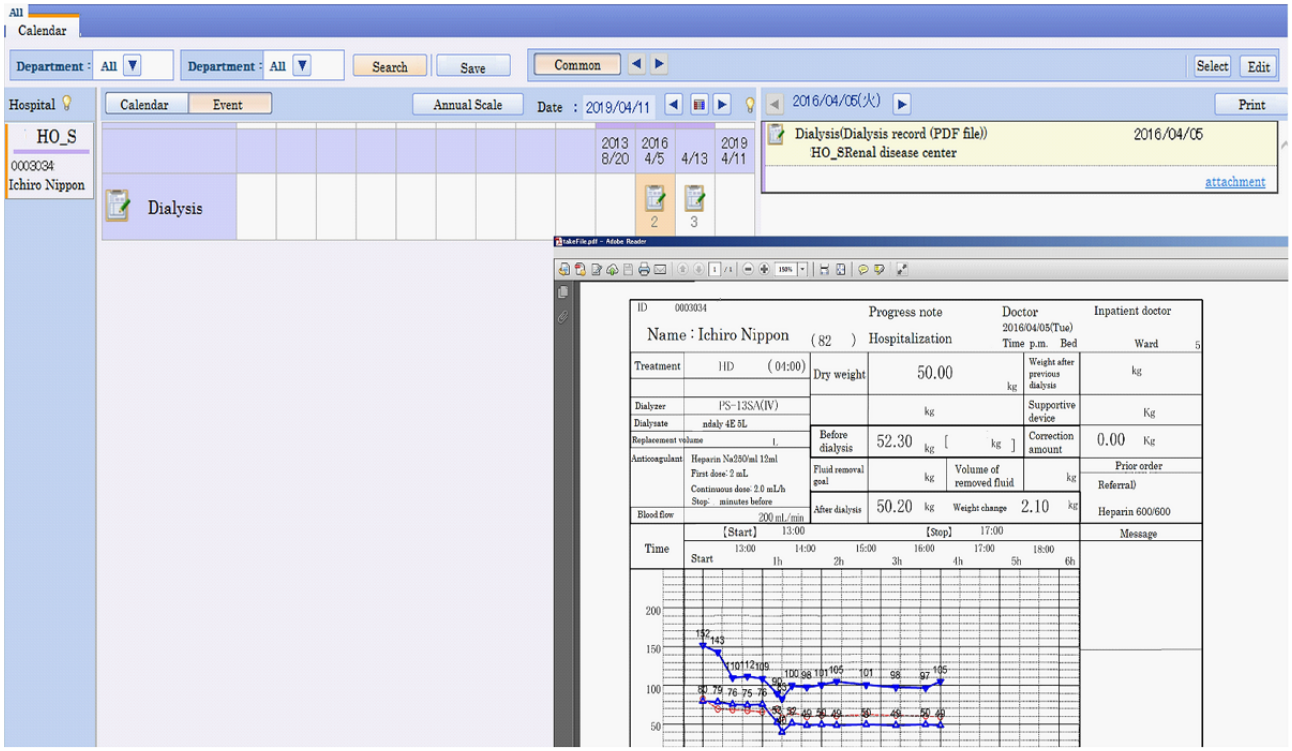


Table 1. Available information about dialysis in the MMWIN and recommended data items for disaster preparation.

	Recommendation 1 [13]	Recommendation 2 [14]	MMWIN[a]
Institution		✓[b]	✓
Cause of renal failure		✓	✓
Diabetic yes/no		✓	✓
Dialyzer	✓	✓	✓
Anticoagulant		✓	✓
Hours of dialysis	✓	✓	✓
Times per week	✓		✓
Date of previous HD[c]	✓	✓	✓
Primary diagnosis	✓		✓
Allergies	✓	✓	✓
Medications	✓	✓	✓
Past medical history	✓	✓	✓
Center HD	✓		✓
Home HD	✓		
CAPD[d]	✓		✓
CCPD[e]	✓		✓
HDF[f] no/pre/post		✓	✓
Dialysis prescription	✓	✓	✓
Usual dialysis machine		✓	✓
Vascular access	✓	✓	✓
Dialysate	✓		✓
Start time			✓
End time			✓
BP[g] before HD			✓
BP after HD			✓
Weight before HD			✓
Weight after HD			✓
Dry weight	✓	✓	✓
Body temperature			✓
Average weight gain	✓	✓	✓
Comments	✓	✓	✓
Heparinization	✓		✓
Needle size	✓	✓	✓
Blood flow rate	✓		✓
Reuse[h]	✓		
Lidocaine	✓		✓
HBsAg[i]	✓		✓
Blood type	✓		✓
Laboratory data		✓	✓

[a]MMWIN: Miyagi Medical and Welfare Information Network.

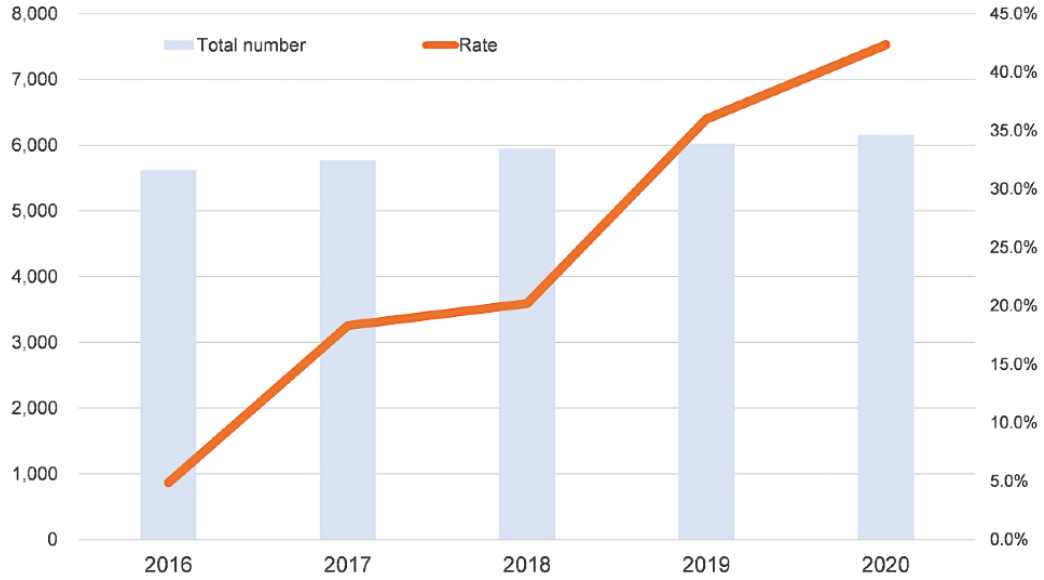
[b]Indicates this item is present in the recommendations or MMWIN.
[c]HD: hemodialysis.
[d]CAPD: continuous ambulatory peritoneal dialysis.
[e]CCPD: continuous cycling peritoneal dialysis.
[f]HDF: hemodiafiltration.
[g]BP: blood pressure.
[h]Reuse is not allowed in Japan
[i]HBsAg: hepatitis B surface antigen.

Table 2. The characteristics of hemodialysis facilities participating in the Miyagi Medical and Welfare Information Network.

Facility	Attending (n=22), n		Not attending (n=46), n	
	Clinic	Hospital	Clinic	Hospital
Fund				
Public	0	5	0	9
Private	12	5	26	11
Area				
Coast	1	2	3	3
North	3	1	6	7
South	4	1	3	2
Central	4	6	14	8
Beds, mean (SD)	1.6 (5.3)	279.3 (320.4)	7.7 (9.1)	306.8 (182.3)
Beds for HD[a], mean (SD)	51.8 (22.1)	40.1 (19.2)	37.4 (13.0)	28.8 (17.3)
Doctors, mean (SD)	2.1 (1.1)	155.1 (396.1)	2.0 (1.1)	67.8 (68.7)
HD patients, mean (SD)	118.0 (66.5)	95.9 (55.7)	87.9 (51.1)	70.8 (55.5)

[a]HD: hemodialysis.

Figure 3. The rate of opt-in patients undergoing hemodialysis in the Miyagi Medical and Welfare Information Network and the number of patients undergoing hemodialysis in Miyagi Prefecture. The bars show the total number of patients undergoing hemodialysis in Miyagi. The line chart represents the ratio of opt-in patients to all patients undergoing hemodialysis in Miyagi.



Discussion

Principal Results

We developed a system for sharing HD reports and other clinical information among medical facilities and enhancing their connection networks. This would be helpful when the condition of patients undergoing HD is aggravated and when they need to be transferred to other hospitals. It would also be effective for the improvement of medical staff members' work burden in the view of a paperless environment where they do not have to send HD reports by fax or mail to other facilities. Another merit is that patient information can be made available during a disaster, resulting in prompt extraction of patient information even in a chaotic situation. Lastly, sharing patient information would be very helpful for collaborative care, especially among general care practitioners and specialists, including nephrologists. An HIE can be a good tool to support clinical decision-making and ensure continuity of care, especially for complex medical problems.

In catastrophic situations, like the GEJE in 2011, the clinical information on MMWIN can be accessed from different facilities to acquire patient clinical data when needed. Since the MMWIN system is independent of the HIS in each facility, the information can be accessed even if the HIS is unavailable. Medical staff members can refer to medication and laboratory results and input clinical notes on the MMWIN system. Therefore, even when an HIS is down, medical staff can conduct HD as usual according to patient records on MMWIN.

Comparison With Prior Work

Access to prior clinical data from different facilities is important for continuing appropriate treatment in the community. A fully interoperable EHR or HIE would also be helpful in maintaining the flow of critical patient information. In particular, for the care of patients undergoing HD, the information would reduce the risks of HD and prevent aggravation of renal function. In our system, we introduced a document management system for several measures to transfer HD reports from an EHR, dialysis systems, or original documents. Although specific information systems were launched elsewhere [15,16], we successfully integrated the sharing of HD reports into the main HIE system without the high cost of additional implementation. This system is based on document sharing. As the next step, digital data transfer is desirable for using clinical data, such as copying and pasting to avoid tedious record keeping by medical staff in daily clinical activity or for research purposes when researchers obtain informed consent from patients for secondary use. In addition to HD, since our document-sharing system is used in various situations, such as reference letters when a patient is transferred, notes to doctors from pharmacists when patients have questions regarding medication at the pharmacy, and sharing documents between care workers and doctors, this system can be used in different fields as a digital transformation tool.

The future of interoperability may lie in using novel standards to transfer clinical data more easily between an HIS and mobile devices, such as personal health records (PHRs). One such standard may be based on the Fast Healthcare Interoperability Resource (FHIR) [17]. We have already attempted to transmit clinical data using the FHIR for PHRs [18]. It could also be applied to HD reports soon. In addition, this system is used for care coordination between primary care physicians (PCPs) and nephrologists to manage the associated complex chronic conditions of patients undergoing HD. This could improve the management of patients with end-stage renal disease [19]. Our previous study revealed that an HIE is a good tool to improve patient prognosis, with collaboration between PCPs and specialists [8].

One of our major concerns in patient management is preparedness for disasters. During a disaster, HD facilities sometimes cannot supply the usual HD environment, resulting in HD at alternate dialysis locations without patients' medical records or information about their dialysis regimen being available. Therefore, we checked whether the available information on the MMWIN could cover the necessary data items during a disaster for dialysis-specific disaster preparedness [13,14,20]. Thus, as shown in Table 1, we covered almost all the recommendations. Available data provided in MMWIN include the latest body weight and vital sign on regular HD, which are essential information for optimal and secure dialysis management. Although it is recommended that the patients themselves always have basic information of their own for disaster preparedness, this is not easy because older people comprise two-thirds of patients undergoing dialysis in Japan [21].

Limitations

This study had some limitations. First, we did not include all HD facilities in Miyagi. Although we managed to enroll 22 HD facilities joining the MMWIN and more than 40% of patients undergoing HD in Miyagi were registered, HD records from dialysis facilities cannot be shared with other dialysis facilities that do not have access to the MMWIN. It would be difficult for small facilities to implement the system without the support of public budgets. Encouragement by the local government through stronger incentives to both facilities and EHR vendors is needed to implement the MMWIN system. There were no significant differences among facilities, except for the number of beds in clinics, with fewer beds in the facilities within the MMWIN system (P=.02; Table 2). In the survey, most clinics with fewer beds were HD-specific facilities that did not have admission beds for inpatients, only for outpatients with HD. Table 2 shows that the HD-specific facilities favored MMWIN systems over other facilities, mainly because of our document-sharing system for patients undergoing HD. This indicates that the HD-specific clinics for only HD outpatients favored using the MMWIN-HD system to share patient information with hospitals, which would accept patients when their condition worsened. In addition to HD reports, the document systems are available for sharing reference letters or notes among hospitals, between hospitals and clinics, between doctors and pharmacists, and between hospitals and nursing care homes. As well as HD facilities, we believe that all health care facilities should make an effort to share patient information across all care settings. Another potential benefit for such a network may be its use to guide medical decisions based on individual patient characteristics (personalized medicine) rather than averages over a whole population. Personalized medicine

in patients with kidney disease is lagging behind the other medical disciplines as most randomized controlled trials are currently excluding patients with kidney disease. Second, we have not encountered a disaster that requires HD records to be shared among facilities. This is fortunate, but there is a requirement to test the system with more practical simulations for possible disasters.

Additionally, we did not have a wide range of training exercises among different facilities to simulate large disasters. Support from patients and many facilities throughout the prefecture would be practical. A plan for a wide range of training exercises is needed in the near future. However, the staff members in facilities attending MMWIN have already implemented daily use of this system in clinical activities. This habit can be helpful, even during disaster situations.

Conclusions

In conclusion, our backup system successfully developed a network of HD facilities. We have accumulated data that are beneficial to prevent the fragmentation of patient information and would be useful to transfer patients efficiently as preparedness for unpredictable disasters.

Acknowledgments

We thank Ms Mikiko Sato, Ms Misaki Arakawa, and Miyagi Medical and Welfare Information Network members for their technical support.

Conflicts of Interest

None declared

References

1. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu C. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. N Engl J Med 2004 Sep 23;351(13):1296-1305. [doi: 10.1056/NEJMoa041031] [Medline: 15385656]
2. Tonelli M, Wiebe N, Guthrie B, James MT, Quan H, Fortin M, et al. Comorbidity as a driver of adverse outcomes in people with chronic kidney disease. Kidney Int 2015 Oct;88(4):859-866 [FREE Full text] [doi: 10.1038/ki.2015.228] [Medline: 26221754]
3. Shibahara S. The 2011 Tohoku earthquake and devastating tsunami. Tohoku J Exp Med 2011 Apr;223(4):305-307 [FREE Full text] [doi: 10.1620/tjem.223.305] [Medline: 21478655]
4. Hao K, Takahashi J, Ito K, Miyata S, Sakata Y, Nihei T, Miyagi AMI Registry Study Investigators. Emergency care of acute myocardial infarction and the great East Japan earthquake disaster. Circ J 2014;78(3):634-643 [FREE Full text] [doi: 10.1253/circj.cj-13-1286] [Medline: 24451649]
5. Masakane I, Akatsuka T, Yamakawa T, Tsubakihara Y, Ando R, Akizawa T, et al. Survey of dialysis therapy during the Great East Japan Earthquake Disaster and recommendations for dialysis therapy preparation in case of future disasters. Renal Replacement Ther 2016 Aug 25;2(1):48. [doi: 10.1186/s41100-016-0060-0]
6. Anderson AH, Cohen AJ, Kutner NG, Kopp JB, Kimmel PL, Muntner P. Missed dialysis sessions and hospitalization in hemodialysis patients after Hurricane Katrina. Kidney Int 2009 Jun;75(11):1202-1208 [FREE Full text] [doi: 10.1038/ki.2009.5] [Medline: 19212421]
7. Ido K, Nakamura N, Nakayama M. Miyagi Medical and Welfare Information Network: a backup system for patient clinical information after the Great East Japan Earthquake and Tsunami. Tohoku J Exp Med 2019 May;248(1):19-25. [doi: 10.1620/tjem.248.19] [Medline: 31080195]
8. Nakayama M, Inoue R, Miyata S, Shimizu H. Health information exchange between specialists and general practitioners benefits rural patients. Appl Clin Inform 2021 May;12(3):564-572 [FREE Full text] [doi: 10.1055/s-0041-1731287] [Medline: 34107543]
9. Yamaguchi M, Inomata S, Harada S, Matsuzaki Y, Kawaguchi M, Ujibe M, et al. Establishment of the MID-NET medical information database network as a reliable and valuable database for drug safety assessments in Japan. Pharmacoepidemiol Drug Saf 2019 Oct;28(10):1395-1404 [FREE Full text] [doi: 10.1002/pds.4879] [Medline: 31464008]
10. Nakagawa N, Sofue T, Kanda E, Nagasu H, Matsushita K, Nangaku M, et al. J-CKD-DB: a nationwide multicentre electronic health record-based chronic kidney disease database in Japan. Sci Rep 2020 Apr 30;10(1):7351. [doi: 10.1038/s41598-020-64123-z] [Medline: 32355258]
11. Sugiyama T, Miyo K, Tsujimoto T, Kominami R, Ohtsu H, Ohsugi M, et al. Design of and rationale for the Japan Diabetes compREhensive database project based on an Advanced electronic Medical record System (J-DREAMS). Diabetol Int 2017 Nov;8(4):375-382 [FREE Full text] [doi: 10.1007/s13340-017-0326-y] [Medline: 30603343]
12. Nakayama M, Takehana K, Kohro T, Matoba T, Tsutsui H, Nagai R, IHE Cardiology Team and SEAMAT Committee. Standard export data format for extension storage of standardized structured medical information exchange. Circ Rep 2020 Sep 19;2(10):587-616 [FREE Full text] [doi: 10.1253/circrep.CR-20-0077] [Medline: 33693184]
13. Planning for emergencies: a guide for people with chronic kidney disease. National Kidney Foundation. URL: https://www.kidney.org/sites/default/files/11-10-0807_IBD_disasterbrochure.pdf [accessed 2022-07-11]

14. Johnson DW, Hayes B, Gray NA, Hawley C, Hole J, Mantha M. Renal services disaster planning: lessons learnt from the 2011 Queensland floods and North Queensland cyclone experiences. Nephrology (Carlton) 2013 Jan;18(1):41-46. [doi: 10.1111/nep.12008] [Medline: 23252802]
15. Vito D, Casagrande G, Bianchi C, Costantino M. An interoperable common storage system for shared dialysis clinical data. 2016 Presented at: 2016 IEEE EMBS International Student Conference; May 29-31, 2016; Ottawa, ON p. 1-4. [doi: 10.1109/embsisc.2016.7508626]
16. Vandenberg AE, Jaar BG, James KP, Lea J, O'Donnell C, Masud T, et al. Making sense of DialysisConnect: a qualitative analysis of stakeholder viewpoints on a web-based information exchange platform to improve care transitions between dialysis clinics and hospitals. BMC Med Inform Decis Mak 2021 Feb 09;21(1):47 [FREE Full text] [doi: 10.1186/s12911-021-01415-y] [Medline: 33563290]
17. Welcome to FHIR®. Health Level Seven International. URL: https://www.hl7.org/fhir/ [accessed 2021-08-04]
18. Xiao D, Song C, Nakamura N, Nakayama M. Development of an application concerning fast healthcare interoperability resources based on standardized structured medical information exchange version 2 data. Comput Methods Programs Biomed 2021 Sep;208:106232 [FREE Full text] [doi: 10.1016/j.cmpb.2021.106232] [Medline: 34174764]
19. Wang V, Diamantidis CJ, Wylie J, Greer RC. Minding the gap and overlap: a literature review of fragmentation of primary care for chronic dialysis patients. BMC Nephrol 2017 Aug 29;18(1):274 [FREE Full text] [doi: 10.1186/s12882-017-0689-0] [Medline: 28851313]
20. Murakami N, Siktel HB, Lucido D, Winchester JF, Harbord NB. Disaster preparedness and awareness of patients on hemodialysis after Hurricane Sandy. Clin J Am Soc Nephrol 2015 Aug 07;10(8):1389-1396 [FREE Full text] [doi: 10.2215/CJN.10181014] [Medline: 26220814]
21. Takeuchi H, Uchida HA, Kakio Y, Okuyama Y, Okuyama M, Umebayashi R, et al. The prevalence of frailty and its associated factors in Japanese hemodialysis patients. Aging Dis 2018 Apr;9(2):192-207 [FREE Full text] [doi: 10.14336/AD.2017.0429] [Medline: 29896410]

Abbreviations

EHR: electronic health record
FHIR: Fast Healthcare Interoperability Resource
GEJE: Great East Japan Earthquake
HD: hemodialysis
HIE: health information exchange
HIS: hospital information system
HL7: Health Level 7
MMWIN: Miyagi Medical and Welfare Information Network
PCP: primary care physician
PHR: personal health record
SS-MIX2: Standardized Structured Medical Information eXchange version 2

Edited by A Mavragani; submitted 15.08.21; peer-reviewed by H Matsui, W Huang, T Ashihara; comments to author 15.05.22; revised version received 06.07.22; accepted 09.07.22; published 22.07.22.

Please cite as:
Ido K, Miyazaki M, Nakayama M
Hemodialysis Record Sharing: Solution for Work Burden Reduction and Disaster Preparedness
JMIR Form Res 2022;6(7):e32925
URL: https://formative.jmir.org/2022/7/e32925
doi:10.2196/32925
PMID:35867394

Original Paper

Association Between Care Modality and Use With Treatment Response Among Members Accessing Virtual Mental Health Services: Real-world Observational Study

Emily Shih[1], PhD; Brandon S Aylward[1], PhD; Sarah Kunkle[1], MPH; Grant Graziani[1], PhD

Ginger, San Francisco, CA, United States

Corresponding Author:
Emily Shih, PhD
Ginger
116 New Montgomery St. Suite 500
San Francisco, CA, 94105
United States
Phone: 1 855 446 4374
Email: eshih@ginger.io

Abstract

Background: There is a growing bottleneck in mental health care, as the demand for services has outpaced the availability of mental health professionals. Consequently, many health systems have shifted to teletherapy as a scalable approach to increasing accessibility to care. Within these care models, various treatment modalities (eg, coaching and clinical care) are used to deliver support for anxiety and depression. However, more research is needed to better understand the differences in treatment responses.

Objective: The purpose of this study was to examine the association between different care modalities and the levels of use with symptom score changes for members seeking virtual care services.

Methods: We conducted an observational study of 4219 members who accessed Ginger, an on-demand mental health service, between September 2020 and September 2021. Using a mobile app, members can access text-based behavioral health coaching and virtual clinical services. This study focused on members with clinically elevated depression or anxiety levels at baseline. Logistic regressions were used to assess the association between care modalities and the levels of use with treatment response in depression and anxiety, using the Patient Health Questionnaire and Generalized Anxiety Disorder Assessment, respectively.

Results: Of the 4219 members, 1623 (38.47%) demonstrated a full response to depression, and 1684 (39.91%) demonstrated a full response to anxiety. Members who completed care (ie, text-based coaching, virtual clinical therapy, hybrid of coaching, and clinical care) beyond the introductory session showed significantly increased odds of a full response compared with those who completed only limited care. Members who completed a hybrid of care had the highest odds of improvement; the odds of showing a full response in depression were 2.31 times higher (95% CI 1.91-2.80; $P<.001$) and in anxiety were 2.23 times higher (95% CI 1.84-2.70; $P<.001$) compared with members who completed limited care. For members who completed only coaching or clinical care, the largest effects were observed among those with high use. For members who completed a hybrid care program, we observed similar treatment responses across all levels of use.

Conclusions: Our real-world study found that members who completed text-based coaching achieved full treatment responses at similar rates compared with members who completed virtual clinical care and members who completed a hybrid of care. There were no significant differences in the predicted probabilities of full treatment response between coaching and clinical care. Generally, the odds for a full response were highest among members with high use within each care modality; however, there were no differences in full-response treatment odds across levels of use with hybrid care. The results support the utility of digital behavioral health interventions and further highlight text-based coaching protocols as an accessible and suitable option when considering virtual care for treating anxiety and depression.

(JMIR Form Res 2022;6(7):e36956) doi:10.2196/36956

KEYWORDS

behavioral coaching; engagement; mental health; telehealth; treatment response

Introduction

Background

Before the current pandemic, the prevalence of mental health illness in the United States was more than 1 in 5 adults (20.6%) [1]. Anxiety and depression result in an estimated global economic cost of US $1 trillion each year for lost productivity, absenteeism, and medical costs [2]. The confluence of physical health risks, financial stressors, social isolation, and disruption of daily activities during the COVID-19 pandemic has had a profound impact on the number of individuals with clinical depression and anxiety [3]. In fact, the Centers for Disease Control and Prevention reported that between August 2020 and February 2021, the percentage of adults with recent symptoms of anxiety or depression increased from 36% to 42%, and the percentage of those reporting an unmet mental health care need increased from 9% to 12% [4]. Despite the urgent need for mental health care, individuals still face many barriers in accessing effective treatment services, including high out-of-pocket expenses and transportation challenges [5,6]. In addition, there has been a long-standing care delivery bottleneck for mental health care as the demand for care has outpaced the availability of qualified mental health professionals. Long waitlists have only worsened during the pandemic and underscore the urgency to identify innovative new models to increase access to effective mental health care.

To address the pressing need for services, many health systems and organizations have shifted to teletherapy and other digital interventions as scalable approaches to increasing accessibility to mental health care [7]. Within these care models, various treatment modalities (eg, self-guided content, coaching, and teletherapy) are used to deliver support for anxiety and depression. The efficacy of virtual teletherapy for the treatment of mood and anxiety disorders has been well established, with outcomes similar to face-to-face visits and greater efficacy compared with treatment as usual or placebo [8-11].

Although virtual care can increase access and reach to care, some approaches still rely on the limited supply of highly trained mental health specialists (eg, clinical psychologists). A promising and increasingly popular method of care is behavioral health coaching, which can serve as a lower-intensity alternative to care. Although this type of care may not be suitable for all types of patients (eg, those with suicidal ideation, substance use disorder, or repeated hospitalizations), behavioral health coaching has demonstrated improvements in both the physical and mental health status of patients [12-14]. Care can be delivered by bachelor's or master's level providers and represents a more scalable solution that does not overly rely on the limited supply of mental health specialists with advanced training (eg, doctoral degree). Coaching uses methods similar to traditional psychotherapy and can address anxiety and depression through techniques derived from interventions such as mindfulness, solution-oriented focus, and positive psychology [15]. With the rapidly shifting landscape of care models, traditional coaching models have adapted to digital methods, such as text-, video-, or telephone-based coaching. One of the benefits of text-based coaching in particular is that it requires less coaching time and allows for both synchronous and asynchronous support for members, which may be more suitable for the on-the-go lifestyle of those seeking care [16]. Several recent studies using text-based coaching have demonstrated treatment outcomes equivalent to those of in-person and telephone-based care [16-19]. These treatment modalities can serve as scalable solutions to address the growing demand for mental health services; however, more research is needed to better understand the differences in treatment responses for depression and anxiety.

This Study

The purpose of this study was to examine the association between different care modalities and different levels of use with clinical symptom score changes in members seeking services in a virtual care model. As such, we have two hypotheses: (1) members who completed text-based coaching, clinical sessions, or a hybrid of coaching and clinical care would demonstrate higher odds of a full treatment response (≥50% reduction in symptom scores) for anxiety and depression compared with members who only completed limited care and (2) members who used more sessions would demonstrate higher odds of a full treatment response.

Methods

Overview

We conducted a retrospective observational study of members who accessed Ginger, an on-demand mental health service, between September 2020 and September 2021. This is a secondary analysis of pre-existing deidentified data. The study team did not have access to participants' identifying information and did not intend to recontact the participants.

Ethics Approval

Ginger's research protocols and supporting policies were reviewed and approved by Advarra's institutional review board (Pro00046797) in accordance with the US Department of Health and Human Services regulations at Title 45 Code of Federal Regulations Part 46 [20].

Participants

Study participants had access to Ginger services as part of their employment or health plan benefits. Internal clinical protocols include exclusionary criteria where self-directed telehealth is likely not appropriate and where more specialized and urgent psychiatric services are required (eg, active suicide ideation and active high-risk self-harm behavior) [21]. This study included Ginger members aged ≥18 years who screened positive on either the Patient Health Questionnaire-9 (PHQ-9) or Generalized Anxiety Disorder-7 (GAD-7) at baseline (ie, score≥10).

Procedures

The Ginger platform provides members with access to text-based behavioral health coaching, virtual clinical services, and self-guided content and assessments primarily via a mobile app platform. Examples of self-guided content include mindfulness meditation activity cards and stress-management exercises. After downloading the mobile app, members can start texting

with a behavioral health coach within minutes of requesting connection. During the coaching sessions, members and coaches work together to set goals and work plans to achieve those goals. Goals can range anywhere, from career goals and relationship goals to other personal goals that the member or coach identifies as a source of anxiety or depression. Members typically begin with text-based coaching sessions and many members remain solely at this level of care. Some members will request clinical care (teletherapy or telepsychiatry), and some will require treatment escalation if the coaches identify a clinical need. Clinical severity was ultimately determined by clinicians using the PHQ-9 and GAD-7, in addition to other assessments that gauge the risk and urgency for clinical care. Examples of situations that require escalation include individuals with chronic mental illness and severe trauma, the potential to harm oneself or others, and significant mental instability (eg, hallucinations, delusions, and extreme mood swings). Members who met certain risk thresholds were advised to escalate to therapy or psychiatry. When members are escalated to therapy or psychiatry, they may continue working with a coach, provided they seek additional specialized care concurrently. Members who did not meet these thresholds were recommended to continue coaching unless they had a specific preference for clinical care.

Ginger coaches are full-time employees who have an advanced degree in a field related to mental health or have accredited coach certification (as approved by the National Board for Health and Wellness Coaching). Coaches are also required to have at least 2 years of relevant experience, of which 6 months must have occurred with direct supervision under a qualified, credentialed, or licensed supervisor. To ensure ongoing quality in the delivery of care, Ginger coaches are trained for at least 200 hours each year on up-to-date effective methodologies (eg, motivational interviewing and goal setting). Ginger clinicians are full-time employees who have completed a minimum master's degree in psychology, social work, counseling, marriage and family therapy, or a related field. They are licensed to practice (eg, licensed clinical social worker, licensed marriage and family therapist, or licensed psychologist) and undergo quarterly training on protocols, evidence-based care, and best practices in telehealth. Additional details regarding the Ginger care model and providers can be found in previous publications [3,21].

To help providers assist with personalized care, members can track changes in depression and anxiety symptoms during their care journey. Members were prompted to complete the PHQ-9 and GAD-7 through the platform when they began using Ginger services to measure their baseline symptom scores. Symptom surveys were administered every 2 weeks to members who scored above the clinical threshold (≥10) at baseline. The most recent survey response that fell within a 6- to 16-week window following a member's baseline was considered their follow-up and used for analyses in this study.

Measures

The PHQ-9 is a 9-item self-report questionnaire based on the Diagnostic and Statistical Manual of Mental Disorders, 4th edition, that assesses the frequency and severity of depression symptoms over the previous 2 weeks [22]. Each of the 9 items is scored on a scale from 0 (not at all) to 3 (nearly every day). Total scores range from 0 to 27, with higher scores indicating more depressive symptoms. A score of 10 is often used as the clinical threshold [22]. Scores of 10 to 14, 15 to 19, and >20 represent moderate, moderately severe, and severe depression, respectively [22]. On the basis of the literature, a reduction in PHQ-9 score of >50% is considered a *full response* [23]. This approach to calculate treatment response better accounts for the variation in baseline severity and, as such, enables a more unbiased measure of symptom improvement.

The GAD-7 is a self-report questionnaire based on the Diagnostic and Statistical Manual of Mental Disorders, 4th edition, diagnostic criteria to assess the frequency and severity of anxious thoughts and behaviors over the past 2 weeks [24]. Each of the 7 items is scored on a scale from 0 (not at all) to 3 (nearly every day), with total scores ranging from 0 to 21. Consistent with the existing literature, a score of 10 was used as the clinical threshold for this study [25]. Scores of 10 to 14 and >15 represent moderate and severe anxiety, respectively [24]. Similar to the threshold used for the PHQ-9, a reduction in the GAD-7 score of >50% was considered a *full response* to treatment.

Care Modality and Levels of Use

We calculated the use based on product user behavior data. Coaching sessions were operationalized as the number of unique days on which members and coaches exchanged at least five text messages. We decided on this threshold because we learned from the internal provider feedback that the first few messages are generally characterized as introductory and administrative-related messages with minimal therapeutic intervention. By definition, it is not possible to have more than one coach session per day. Clinical sessions were operationalized as the number of completed video sessions with a clinician and scheduled on a need basis. Each video session was typically an hour long. We further categorized use based on member use of different platform care modalities (ie, text-based coaching and clinical sessions) and the number of times they used these modalities (ie, minimum, low, moderate, and high). This study considered 4 different care modalities. Members who did not complete any coaching or clinical sessions and members who only completed one coaching or one clinical session or one of both were categorized in the *Limited Care* cohort. We included members who completed one coaching or one clinical session or one of both in the *Limited Care* cohort, because, generally, these first sessions are considered introductory, where providers introduce themselves and delineate the structure and plan for the member's care journey. Members who completed more than one text-based coaching session and either none or one clinical session were categorized as the *Coaching Only* cohort. Members who completed more than one clinical teletherapy session and either none or one coaching session were categorized as the *Clinical Only* cohort. Finally, members who completed more than one coaching and more than one clinical session were categorized as the *Hybrid Care* cohort.

Previous literature has not yet established an optimal threshold for on-demand text-based care. As such, we adopted a

data-driven approach to operationalize the levels of use. Different levels of use were determined for the care modalities, Coaching Only, Clinical Only, and Hybrid Care, based on quartiles of completed sessions within each care modality cohort (25th, 50th, and 75th percentiles). Members who fell below the 25th percentile for their respective cohort were categorized as *Minimal Use*. Members who fell between the 25th percentile and the 50th percentile median were categorized as *Low Use.* Members who fell between the 50th and 75th percentiles were categorized as *Moderate Use* and members who scored above the 75th percentile were categorized as *High Use*. The exact number of sessions for each group is shown in Table 1.

Table 1. Description of levels of use by care modality cohort (N=4219; Limited Care n=1072).

Level of use	Definition by quarter percentiles, sessions	Participants, n (%)
Coaching Only (n=1354)		
Minimal	≤2	329 (7.79)
Low	3 to 4	345 (8.18)
Moderate	5 to 7	332 (7.86)
High	≥8	348 (8.25)
Clinical Only (n=941)		
Minimal	≤3	363 (8.6)
Low	4	169 (4.01)
Moderate	5 to 6	212 (5.03)
High	≥7	197 (4.67)
Hybrid Care (n=852)		
Minimal	≤7	260 (6.16)
Low	8 to 9	173 (4.1)
Moderate	10 to 13	231 (5.48)
High	≥14	188 (4.46)

Data Management and Analysis

Analyses were conducted using R Studio (version 1.4.1717; RStudio). Data were first screened for outliers and normality. Separate ANOVA models were used to evaluate whether members varied in their baseline depression and anxiety scores. Given the binary nature of our dependent outcome variable (*full response* vs *no full response*), we used logistic regression modeling, a common statistical method for quantifying the relationship between various factors and a binary clinical outcome. In our first set of analyses, 2 binary logistic regression models were used to explore the association between care modality and the likelihood of demonstrating a full response in depressive and anxiety symptoms. These logistic regression models produce estimates of the probability of demonstrating reductions in depressive and anxiety symptoms when a member is in a particular group that represents a modality of care. The likelihood of each modality group demonstrating a full response relative to the reference group is shown by odds ratios. In addition, 6 logistic regression models were used to explore the association between levels of use within each care modality and treatment response for depression and anxiety symptoms. Members' depression and anxiety symptom scores at baseline were included as covariates in all models to account for variations in the baseline symptom scores. The performance of the logistic regression models was evaluated using the Hosmer-Lemeshow test to measure model fit, and odds ratios and 95% CIs were calculated to show associations with improvements in depression and anxiety. The models were constructed using complete case analyses.

Results

Participant Demographics and Characteristics

A total of 6466 members participated in this study. Of those 6466 members, 4219 (65.25%) screened positive for depression or anxiety (ie, PHQ-9 or GAD-7 symptom scores ≥10). This study focused only on members who screened positive at baseline (N=4219), which will be subsequently referred to as the analytical sample. The complete descriptive statistics are reported in Table 2. Of the 4219 members in the analytical sample, 1645 (38.99%) members were identified as female, 689 (16.33%) as male, 142 (3.37%) as other, and 1743 (41.31%) did not have gender identity information available. Of the 4219 members, a total of 1613 (38.23%) members were aged <35 years, 1264 (29.96%) members were ≥35 years, and 1342 (31.81%) members did not have their age reported.

Table 2. Characteristics of the analytical sample (N=4219)[a,b].

Characteristics	Analytical sample
Age (years), n (%)	
18 to 34	1613 (38.23)
>35	1264 (29.96)
No response	1342 (31.81)
Gender, n (%)	
Female	1645 (38.99)
Male	689 (16.33)
Other	142 (3.37)
No response	1743 (41.31)
Care modality, n (%)	
Limited Care	1072 (25.41)
Coaching Only	1354 (32.09)
Clinical Only	941 (22.3)
Hybrid Care	852 (20.19)
GAD-7[c] baseline score, mean (SD)	12.7 (4.54)
PHQ-9[d] baseline score, mean (SD)	13.7 (5.04)

[a]Counts and percentages were reported for categorical variables.

[b]Mean and SDs were reported for continuous variables.

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

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

On average, members completed 3.53 (SD 3.81) text-based coaching sessions and 2.32 (SD 2.85) clinical sessions. Of the total number of members in the analytical sample, 25.41% (1072/4219) received limited care, 32.09% (1354/4219) completed only text-based coaching, 22.3% (941/4219) completed only clinical care, and 20.19% (852/4219) completed a hybrid of coaching and clinical care.

The average baseline scores on the PHQ-9 and GAD-7 were 13.7 (SD 5.04) and 12.7 (SD 4.54), respectively. A total of 75.07% (3167/4219) of the participants demonstrated at least one unit of improvement in their PHQ-9 score from baseline to follow-up, with an average unit decrease in symptom scores of 4.43 units (SD 6.27). On the GAD-7, 72.05% (3040/4219) of members demonstrated at least one unit of improvement in anxiety scores from baseline to follow-up and the average unit decrease in scores was 4.08 (SD 5.82). With regard to treatment response, 38.47% (1623/4219) demonstrated a full response on the PHQ-9 and 39.91% (1684/4219) demonstrated a full response on the GAD-7.

Separate ANOVA models were used to evaluate the differences in PHQ-9 and GAD-7 baseline scores between the care modality cohorts. The results revealed a significant difference between cohorts for both baseline PHQ-9 scores ($F_{3,4215}$=12.5; P<.001) and baseline GAD-7 scores ($F_{3,4215}$=2.99; P=.03). Tukey post hoc tests revealed that members who completed a hybrid of care had significantly higher baseline PHQ-9 scores than those who completed limited care ($Mean_{Hybrid}$=14.5, $Mean_{Limited}$=13.5; P<.001). Members who completed only text-based coaching sessions had significantly lower baseline PHQ-9 scores than those who completed only clinical sessions ($Mean_{Coaching}$=13.2, $Mean_{Clinical}$=14.0; P<.001) or members who completed a hybrid of care ($Mean_{Hybrid}$=14.5; P<.001). In addition, members who completed only coaching sessions had significantly lower baseline GAD-7 scores than those who completed only clinical sessions ($Mean_{Coaching}$=12.5, $Mean_{Clinical}$=12.8; P=.02). These differences were expected because of the nature of our triaging system used to direct members to the correct level of care. It is also important to point out that the magnitude of these mean differences is small (≤1 point; Table 3). However, owing to significant differences in baseline scores, as indicated by F tests, we included both PHQ-9 and GAD-7 baseline scores as covariates in all regression models.

Table 3. Descriptive statistics for the Patient Health Questionnaire-9 (PHQ-9) and Generalized Anxiety Disorder-9 (GAD-7) scores at baseline and follow-up, by care modality (N=4219).

	F statistic (*df*)	*P* value	Limited care (n=1072), mean (SD)	Coaching Only (n=1354), mean (SD)	Clinical Only (n=941), mean (SD)	Hybrid Care (n=852), mean (SD)
Baseline PHQ-9	12.51 (3,4215)	<.001	13.5 (4.89)	13.2 (4.97)	14.0 (5.14)	14.5 (5.14)
Follow-up PHQ-9	33.5 (3,4215)	<.001	10.8 (6.26)	9.25 (6.28)	8.54 (5.31)	8.39 (5.52)
Baseline GAD-7	2.99 (3,4215)	.03	12.7 (4.61)	12.5 (4.43)	13.0 (4.48)	12.8 (4.65)
Follow-up GAD-7	39.2 (3,4215)	<.001	10.1 (5.84)	8.56 (5.71)	7.97 (5.10)	7.67 (5.37)

Depression: Full Response by Care Modality

We used a logistic regression model to investigate the association between care modality and full response in depression symptom scores. A Hosmer-Lemeshow test failed to reject the null hypothesis, indicating goodness of fit (χ^2_8=6.8; *P*=.56). All modalities showed increased odds of symptom improvement compared with members who completed limited care, but the strongest odds were observed for members who engaged with a hybrid of coaching and clinical care; the odds of showing a full response were 2.31 times higher (95% CI 1.91-2.80; *P*<.001) for those who engaged with hybrid care compared with members who completed limited care. Of note, overlapping CIs among Coaching Only (95% CI 1.41-1.99), Clinical Only (95% CI 1.64-2.38), and Hybrid Care (95% CI 1.91-2.79) cohorts suggest that members in these groups did not differ significantly from one another when predicting depression symptom improvement. Full model coefficients are presented in Table 4. These results are further shown graphically as the predicted probability of a full response in depression scores by care modality and levels of baseline severity (Figure 1).

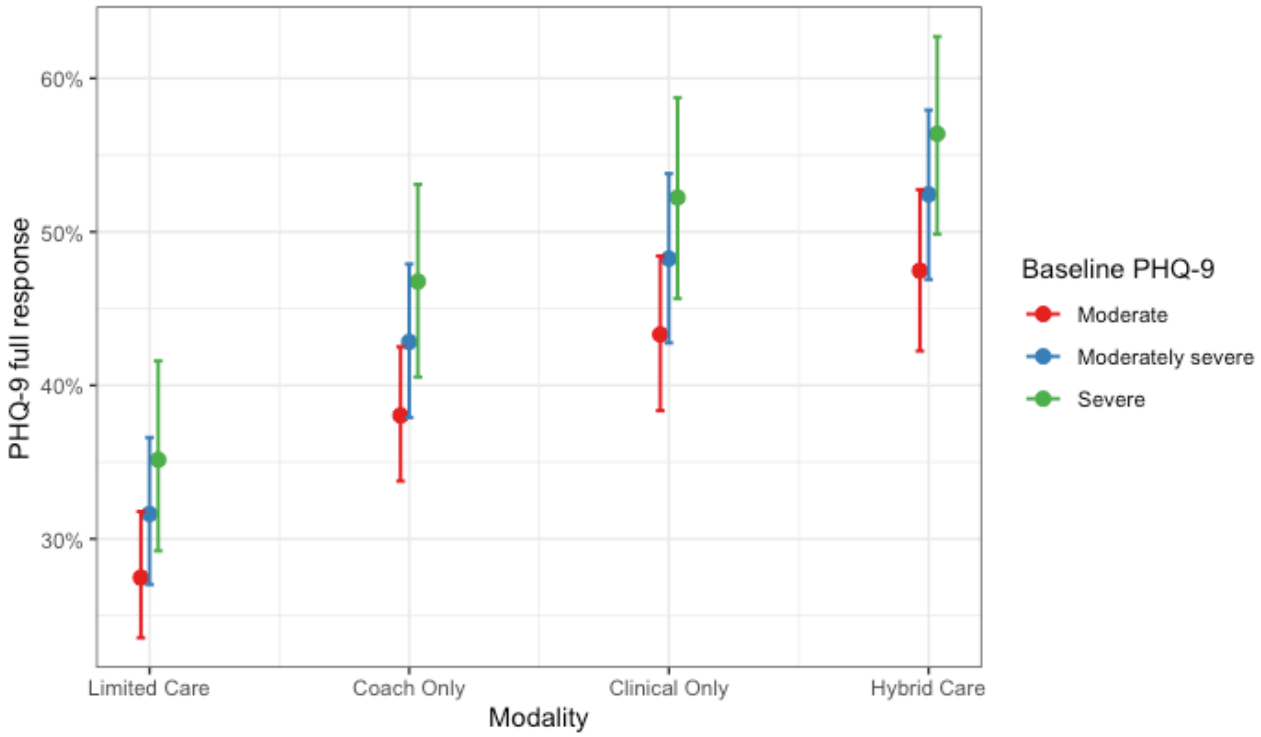
Table 4. Coefficients from the logistic regression model predicting a full response in depression (Patient Health Questionnaire-9 [PHQ-9]) and a full response in anxiety (Generalized Anxiety Disorder-7 [GAD-7]) across care modalities[a], while controlling for scores at baseline (N=4219).

	Odds ratio (95% CI)	Percentage change (%)	*P* value
Dependent variable:[b] full response in PHQ-9			
Intercept	0.29 (0.23-0.37)	−243	<.001
Coaching Only	1.67 (1.41-1.99)	67	<.001
Clinical Only	1.98 (1.64-2.38)	97	<.001
Hybrid Care	2.31 (1.91-2.80)	131	<.001
PHQ-9 score at baseline	1.05 (1.03-1.06)	5	<.001
GAD-7 score at baseline	0.97 (0.96-0.99)	−3	<.001
Dependent variable:[b] full response in GAD-7			
Intercept	0.38 (0.29-0.48)	−166	<.001
Coaching Only	1.69 (1.42-2.01)	69	<.001
Clinical Only	2.20 (1.83-2.66)	120	<.001
Hybrid Care	2.23 (1.84-2.70)	123	<.001
PHQ-9 score at baseline	0.95 (0.94-0.96)	−5	<.001
GAD-7 score at baseline	1.06 (1.05-1.08)	6	<.001

[a]Reference group for care modality: limited care.

[b]Dependent variable: coded full response=1; no full response=0.

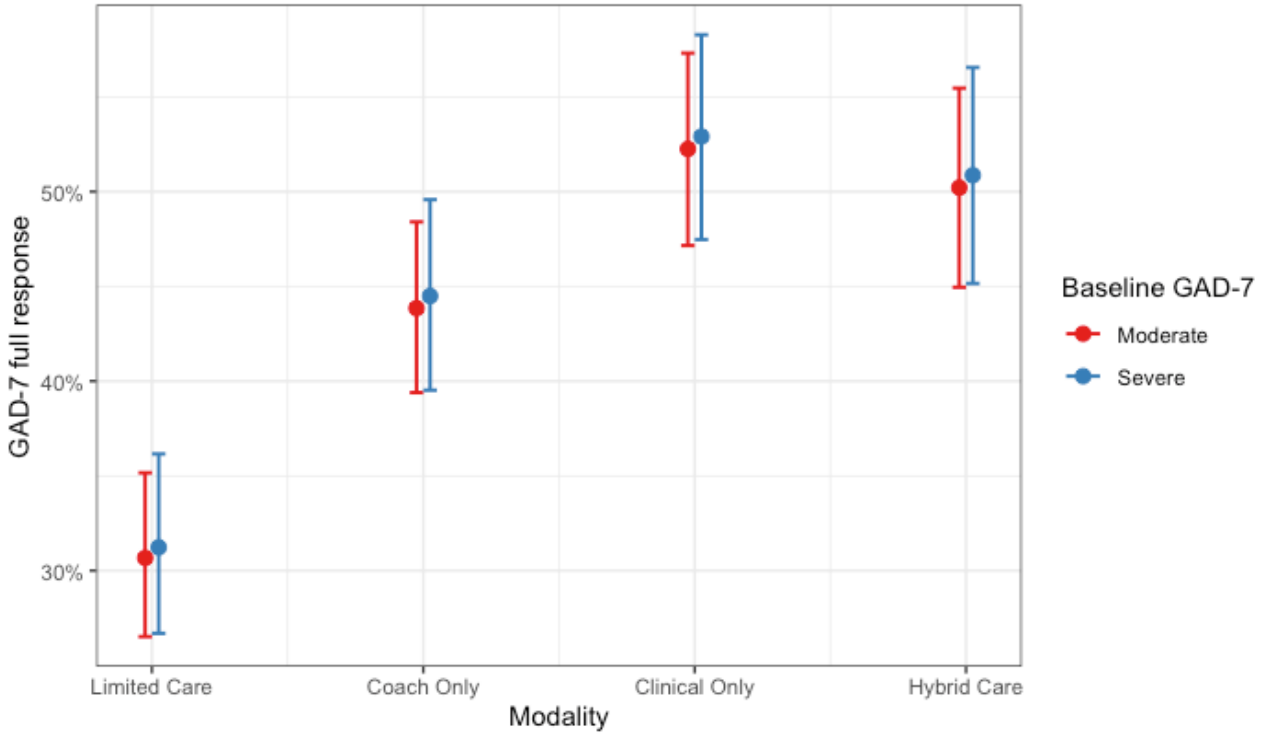
Figure 1. Probability of a full treatment response for depression by care modality and level of baseline severity. PHQ-9: Patient Health Questionnaire-9.



Anxiety: Full Response by Care Modality

We used another logistic regression model to investigate the association between care modality and full response in anxiety symptom scores. A Hosmer-Lemeshow test failed to reject the null hypothesis, indicating goodness of fit (χ^2_8=4.4; *P*=.82). All care modalities showed increased odds of symptom improvement compared with limited care, but the strongest odds were observed for members who completed a hybrid of coaching and clinical care; the odds of showing a full response in anxiety symptom scores were 2.23 times higher (95% CI 1.84-2.70; *P*<.001) for members who completed a hybrid of both clinical and coaching care compared with members who completed limited care. Of note, overlapping CIs among Coaching Only (95% CI 1.42-2.01), Clinical Only (95% CI 1.83-2.66), and Hybrid Care (95% CI 1.84-2.70) cohorts suggest that members in these groups did not differ significantly from one another when predicting full treatment response in anxiety symptoms. Full model coefficients are presented in Table 4. These results are shown graphically as the predicted probability of a full response in anxiety by care modality and levels of baseline severity (Figure 2).

Figure 2. Probability of a full treatment response for anxiety by care modality and level of baseline severity. GAD-7: Generalized Anxiety Disorder-7.



Coaching Only Cohort: Full Response by Levels of Use

Members in the Coaching Only cohort completed, on average, 5.53 (SD 3.65) coaching sessions. Within this modality, we estimated two logistic regression models examining (1) the association between levels of use of text-based coaching and a full response on the PHQ-9 and (2) the association between levels of use of text-based coaching and a full response on the GAD-7. Hosmer-Lemeshow tests failed to reject the null hypothesis, indicating goodness of fit (PHQ-9: χ^2_8=9.7; *P*=.29 and GAD-7: χ^2_8=5.9; *P*=.66). Full coefficients for both models are presented in Table 5. Compared with members with minimal use, members with moderate and high levels of use had significantly increased odds of treatment response in depression. Specifically, the odds of showing a full response in depression were 2.44 times higher (95% CI 1.77-3.37; *P*<.001) for members with high use compared with members with minimal use. Similar patterns were observed for the model predicting a full response in anxiety. Compared with members with minimal use, all other levels of use had significantly increased odds of treatment response for anxiety. The odds of showing a full response in anxiety were 1.99 times higher (95% CI 1.44-2.74; *P*<.001) for members with high use compared with members with minimal use. The association remained after adjusting for baseline anxiety and depression symptom scores. These results are shown graphically as the predicted probability of a full response by coaching use and baseline severity in Figure 3 and Figure 4 for depression and anxiety, respectively.

Table 5. Coefficients from the logistic regression model with the Coaching Only cohort predicting a full response in depression (Patient Health Questionnaire-9 [PHQ-9]) and a full response in anxiety (Generalized Anxiety Disorder-7 [GAD-7]) across levels of use[a], while controlling for scores at baseline (N=1354).

	Odds ratio (95% CI)	Percentage change (%)	*P* value
Dependent variable:[b] full response in PHQ-9			
Intercept	0.39 (0.29-0.67)	−159	<.001
Low use	1.40 (0.95-1.75)	40	.04
Moderate use	1.63 (1.05-1.93)	63	.003
High use	2.44 (1.91-3.53)	144	<.001
PHQ-9 score at baseline	1.04 (1.02-1.07)	4	.001
GAD-7 score at baseline	0.96 (0.93-0.99)	−4	.003
Dependent variable:[b] full response in GAD-7			
Intercept	0.46 (0.29-0.72)	−118	<.001
Low use	1.48 (1.08-2.05)	48	.02
Moderate use	1.84 (1.33-2.55)	84	<.001
High use	1.99 (1.44-2.74)	99	<.001
PHQ-9 score at baseline	0.94 (0.92-0.96)	−7	<.001
GAD-7 score at baseline	1.07 (1.04-1.10)	7	<.001

[a]Reference group for use: minimal use.

[b]Dependent variable coded full response=1; no full response=0.

Figure 3. Probability of a full treatment response for depression by levels of use and level of baseline severity for members in the Coaching Only cohort. PHQ-9: Patient Health Questionnaire-9.

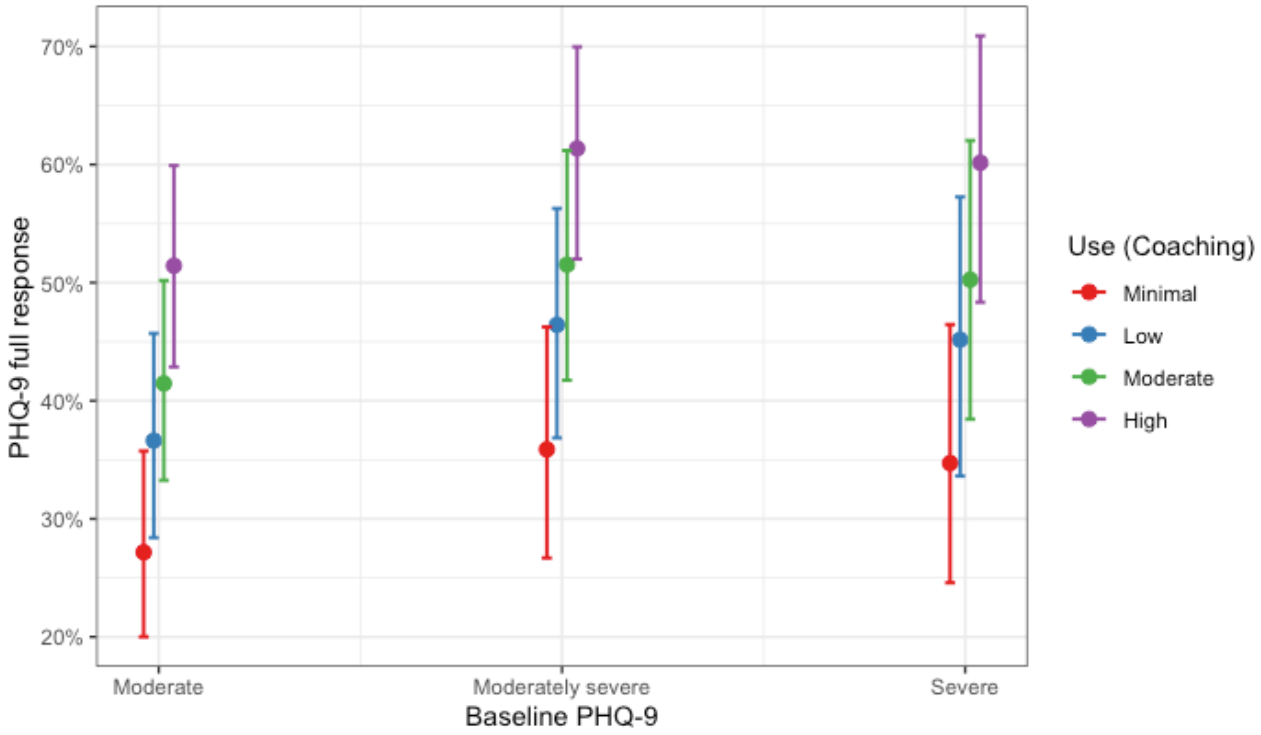
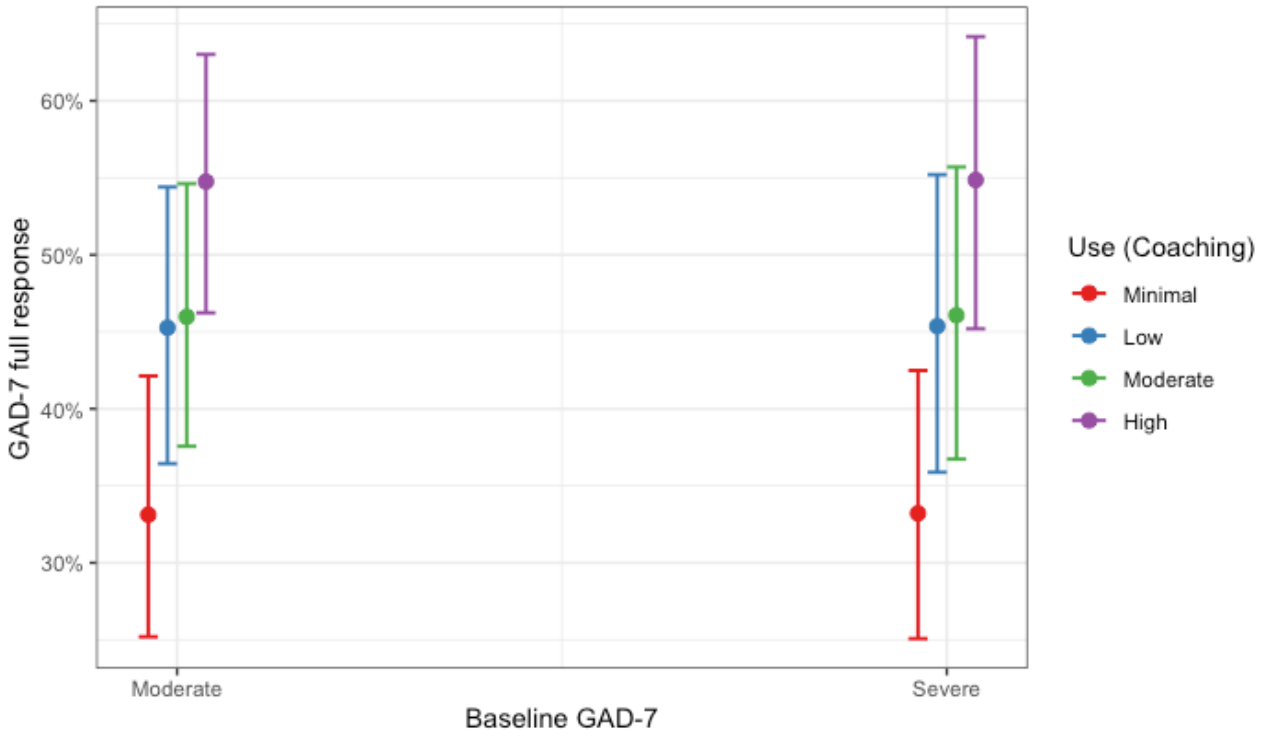


Figure 4. Probability of a full treatment response for anxiety by levels of use and level of baseline severity for members in the Coaching Only cohort. GAD-7: Generalized Anxiety Disorder-7.



Clinical Only Cohort: Full Response by Levels of Use

The Clinical Only cohort of members completed, on average, 4.69 (SD 2.43) clinical sessions. We estimated two logistic regression models examining (1) the association between levels of use in clinical care and full response on the PHQ-9 and (2) the association between levels of use in clinical care and full response on the GAD-7. Hosmer-Lemeshow tests failed to reject the null hypothesis, indicating goodness of fit (PHQ-9: χ^2_8=4.4; *P*=.82 and GAD-7: χ^2_8=6.3; *P*=.62). Full coefficients for both models are presented in Table 6. Compared with members with minimal use, all other levels of use had significantly increased odds of treatment response for depression. The odds of showing a full response in depression were 2.06 times higher (95% CI 1.44-2.95; *P*<.001) for members with high use compared with members with minimal use. Different patterns were observed for the model predicting a full response in anxiety. Compared with members with minimal use, only members who had high levels of use had significantly increased odds of treatment response for anxiety. The odds of showing a full response in anxiety were 1.43 times higher (95% CI 1.00-2.04; *P*<.001) for members with high use compared with members with minimal use. The association remained after adjusting for baseline depression and anxiety scores. These results are shown graphically as the probability of a full response by clinical use and baseline severity in Figure 5 and Figure 6 for depression and anxiety, respectively.

Table 6. Coefficients from the logistic regression model with the Clinical Only cohort predicting a full response in depression (Patient Health Questionnaire-9 [PHQ-9]) and a full response in anxiety (Generalized Anxiety Disorder-7 [GAD-7]) across levels of use[a], while controlling for scores at baseline (N=941).

	Odds ratio (95% CI)	Percentage change (%)	*P* value
Dependent variable:[b] full response in PHQ-9			
Intercept	0.29 (0.17-0.47)	−250	<.001
Low use	1.84 (1.26-2.68)	84	.001
Moderate use	1.61 (1.13-2.28)	61	.008
High use	2.06 (1.44-2.95)	106	<.001
PHQ-9 score at baseline	1.06 (1.03-1.09)	−2	<.001
GAD-7 score at baseline	0.98 (0.95-1.01)	6	.26
Dependent variable:[b] full response in GAD-7			
Intercept	0.53 (0.32-0.87)	−90	.01
Low use	1.27 (0.87-1.84)	27	.21
Moderate use	1.25 (0.88-1.76)	25	.21
High use	1.43 (1.00-2.04)	43	.05
PHQ-9 score at baseline	0.96 (0.93-0.98)	−4	.002
GAD-7 score at baseline	1.07 (1.04-1.11)	7	<.001

[a]Reference group for use: minimal use.

[b]Dependent variable coded full response=1; no full response=0.

Figure 5. Probability of a full treatment response for depression by levels of use and level of baseline severity for members in the Clinical Only cohort. PHQ-9: Patient Health Questionnaire-9.

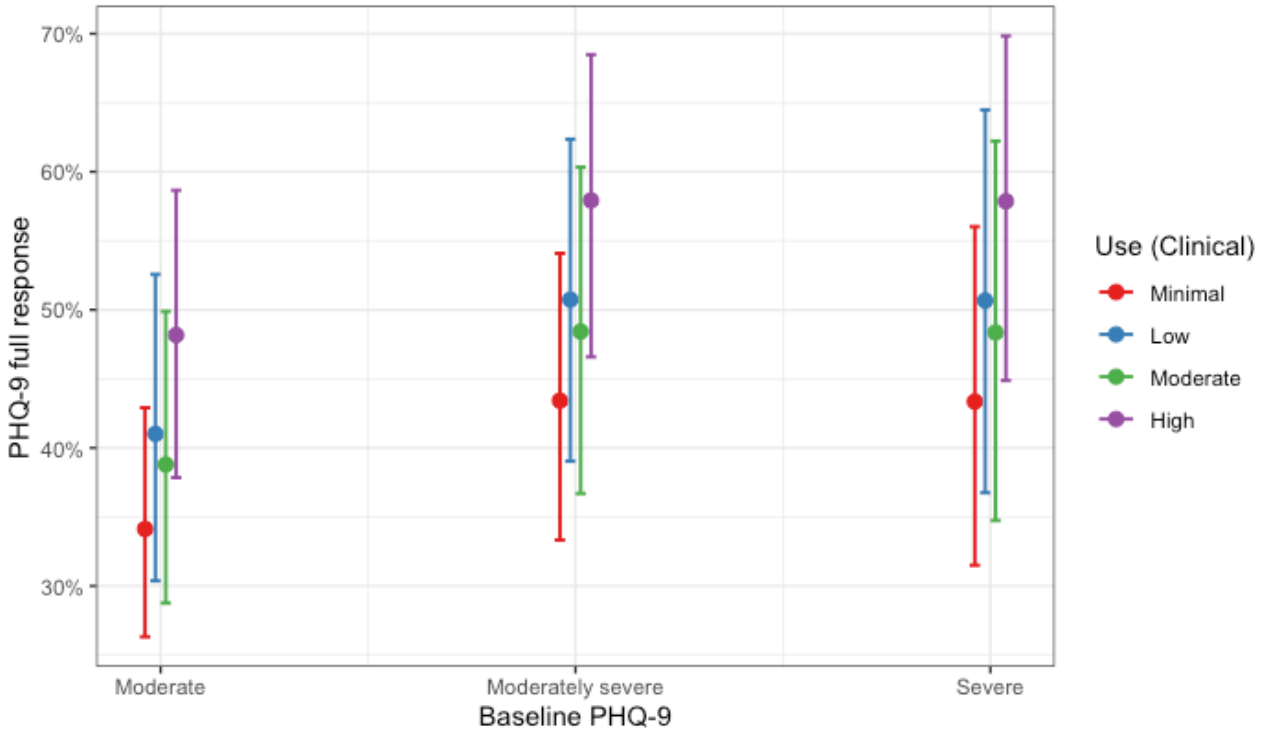
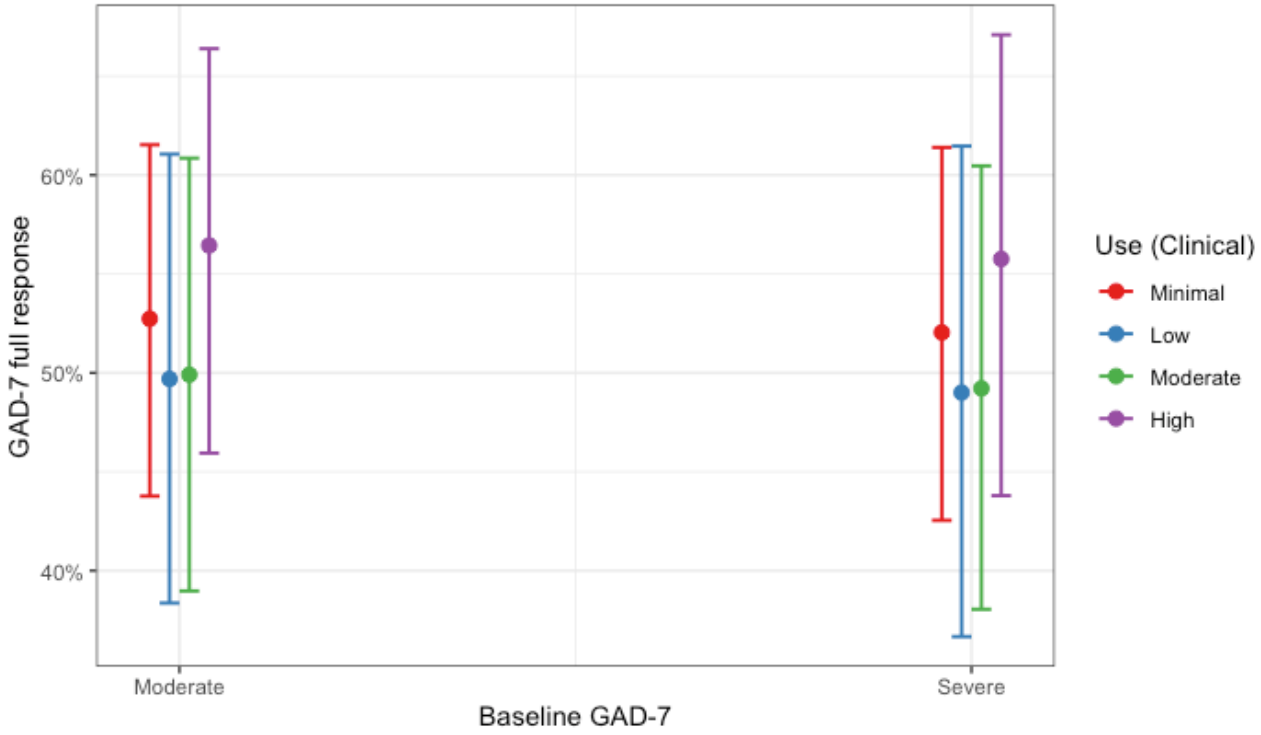


Figure 6. Probability of a full treatment response for anxiety by levels of use and level of baseline severity for members in the Clinical Only cohort. GAD-7: Generalized Anxiety Disorder-7.



Hybrid Care (Coaching+Clinical) Cohort: Full Response by Levels of Use

The Hybrid Care cohort of members completed, on average, 10.5 sessions (5.70 coaching sessions; 4.79 clinical sessions). We estimated two logistic regression models examining (1) the association between levels of use and full response on the PHQ-9 and (2) the association between levels of use and full response on the GAD-7 for members who completed a hybrid of coaching and clinical care. Hosmer-Lemeshow tests failed to reject the null hypothesis, indicating goodness of fit (PHQ-9: χ^2_8=5.6; *P*=.70 and GAD-7: χ^2_8=6.1; *P*=.64). Full coefficients for the models are presented in Table 7. Members with minimal use had similar odds of improvement compared with all other levels of use for both depression and anxiety. The association remained after adjusting for depression and anxiety baseline scores. These results are shown graphically as the probability of a full response by Hybrid Care use and baseline severity in Figure 7 and Figure 8 for depression and anxiety, respectively.

Table 7. Coefficients from the logistic regression model with the Hybrid Care cohort predicting a full response in depression (Patient Health Questionnaire-9 [PHQ-9]) and a full response in anxiety (Generalized Anxiety Disorder-7 [GAD-7]) across levels of use[a], while controlling for scores at baseline (N=852).

	Odds ratio (95% CI)	Percentage change (%)	*P* value
Dependent variable:[b] full response in PHQ-9			
Intercept	0.61 (0.37-1.03)	−63	.06
Low use	1.01 (0.69-1.49)	1	.96
Moderate use	1.07 (0.75-1.53)	7	.72
High use	0.86 (0.59-1.25)	−17	.43
PHQ-9 score at baseline	1.05 (1.02-1.08)	5	.003
GAD-7 score at baseline	0.98 (0.95-1.01)	−2	.23
Dependent variable:[b] full response in GAD-7			
Intercept	0.61 (0.36-1.02)	−65	.06
Low use	1.23 (0.83-1.82)	23	.30
Moderate use	1.27 (0.89-1.83)	27	.19
High use	1.28 (0.87-1.88)	28	.21
PHQ-9 score at baseline	0.96 (0.93-0.98)	−5	.002
GAD-7 score at baseline	1.07 (1.03-1.10)	7	<.001

[a]Reference group for use: minimal use.

[b]Dependent variable coded full response=1; no full response=0.

Figure 7. Probability of a full treatment response for depression by levels of use and level of baseline severity for members in the Hybrid Care cohort. PHQ-9: Patient Health Questionnaire-9.

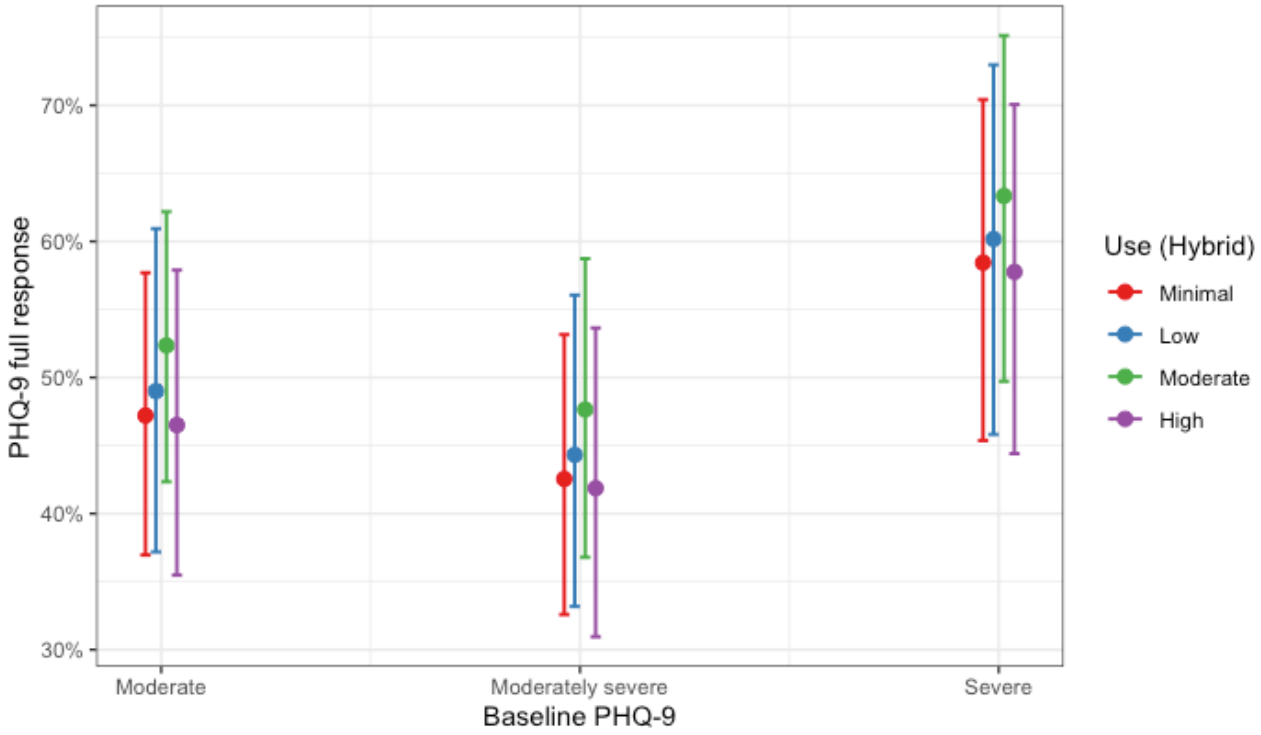
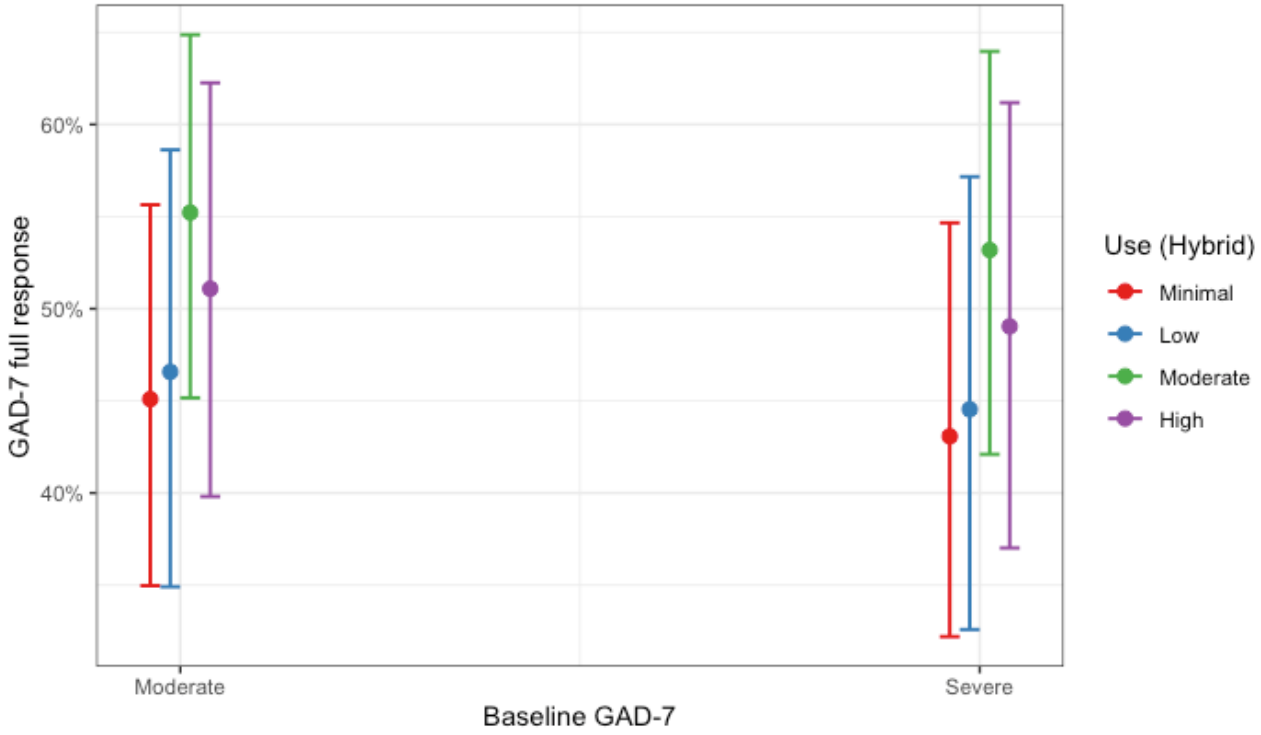


Figure 8. Probability of a full treatment response for anxiety by levels of use and level of baseline severity for members in the Hybrid Care cohort. GAD-7: Generalized Anxiety Disorder-7.



Discussion

Principal Findings

This study examined differences in depression and anxiety treatment response by care modality (ie, limited care, text-based Coaching Only, Clinical Only, and Hybrid Care), as well as by levels of use within each care modality (ie, minimal, low, moderate, and high) in members with moderate to severe baseline anxiety or depression. Overall, nearly 2 out of 5 members demonstrated a full response on the PHQ-9 (1623/4219, 38.47%) and GAD-7 (1684/4219, 39.91%). Our logistic regression models examining the association between care modalities and treatment response found significantly greater odds of full-treatment response in depression and anxiety for members who completed text-based coaching, clinical care, and hybrid care compared with members who completed limited care. The probability of treatment response did not differ significantly between text-based coaching, clinical care, and hybrid care, as indexed by the overlapping 95% CIs, with all 3 modalities estimating a probability of response >35%. In addition to differences by care modality, we also examined treatment responses by levels of use within each care modality cohort. Our data suggest that although all care modalities (ie, Coaching Only, Clinical Only, and Hybrid Care) appeared to offer comparable benefits in managing depression and anxiety above limited care, the treatment responses within each care modality cohort differed among levels of use. The possible explanations for these findings are discussed further.

Members who completed limited care had significantly lower odds of demonstrating a full response on both the PHQ-9 and GAD-7. Given that our analytical sample included only members with a baseline score on the GAD-7 or PHQ-9 ≥10, the established threshold for clinical severity, members likely needed more than limited care (0-1 coaching, 0-1 clinical sessions) to reduce their symptomatology. Members in the Limited Care cohort did not sufficiently interact with a provider, but some could have used self-guided content, although the latter was not explicitly investigated in this study. Our results suggest that interaction with a provider, above and beyond the initial introduction session, offers added benefits in addressing depression and anxiety symptoms. A systematic review of internet interventions for depression found a linear effect on the role of clinician contact, such that between-group Cohen *d* effect size was 0.21, if there was no contact with a clinician either before or during treatment, an effect size of Cohen *d*=0.58 if there was contact with a clinician during treatment, and an effect size of Cohen *d*=0.76 if there was therapist contact before and during treatment [26]. Trends from this study further highlight the importance of offering convenient and accessible telehealth care. In addition, the Ginger model is based on a continuum of care that provides seamless integration and escalation across different levels of care. As such, our findings offer early evidence that members are being properly directed to the correct type of care based on their mental health needs and goals. However, this is speculative and outside the scope of this study; therefore, additional research is required.

We observed similar outcomes in those who completed text-based coaching sessions and clinical care. This result supports prior work suggesting that coaching can be a suitable option for addressing anxiety and depression [12-15], which supplements the literature by investigating text messaging as a viable delivery medium for coaching. The principles of text-based coaching are consistent with how individuals engage with text messaging in their daily routines, and on-demand text-based coaching offers added convenience in terms of not

requiring scheduled appointments [16]. Given the scalability of text-based coaching, in which coaches can have simultaneous conversations, this type of care may be a more cost-effective alternative to traditional psychotherapy.

When focusing on the levels of use within care modality cohorts, our results revealed nuanced patterns. For members who completed only coaching or clinical care, the largest effects were observed among those with high use (above the 75th percentile), estimating a probability of response of >45% for both depression and anxiety. Interestingly, for members who completed a hybrid of coaching and clinical care, we saw similar treatment responses across all levels of use (ie, minimal, low, moderate, and high use), with all levels estimating a probability of response similar to the probability of response for members with high use in the Coaching Only and Clinical Only cohorts (approximately 45% for depression and anxiety). Taken together, our results suggest that typically higher use with care would yield better outcomes; however, this was not true for the members in our Hybrid Care cohort, where all levels of use had similar odds of a full treatment response for both depression and anxiety. This is likely owing to the members of the Hybrid Care cohort that used more sessions. On average, members in the Hybrid Care cohort (mean 10.5) completed approximately twice the number of sessions than members in either the Clinical Only cohort (mean 4.69 sessions) or in the Coaching Only cohort (mean 5.53 sessions) within the study period. In addition, members with minimal use in the Hybrid Care cohort completed a similar number of sessions compared with members with high use in the Coaching Only and Clinical Only cohorts. However, the use of more sessions within the Hybrid Care cohort did not demonstrate increased odds of a full treatment response. These findings suggest that more sessions do not always yield better treatment responses, and future research evaluating care should consider nonlinear patterns.

In addition, across all care modality cohorts, the odds of demonstrating a full response were higher for depression than for anxiety. We have considered a couple of reasons to explain these patterns: (1) the Ginger care model demonstrates more support for treating depression within the study time frame than for treating anxiety and (2) the study time frame we assessed (6-16 weeks following baseline) might be a sufficient amount of time to address depressive symptoms, but more time is needed to address anxiety symptoms. These explanations are speculative, and additional research is needed to specifically test these hypotheses.

Limitations and Future Studies

This study has several limitations. The data set used for this study was limited to people who had access to Ginger services, who had completed the PHQ-9 and GAD-7 survey measures, and whose baseline scores were ≥10. As such, the results are not generalizable to those who do not have access to the system or to those who may have discontinued treatment. A large percentage of members did not have age (1342/4219, 31.81%) or gender (1743/4219, 41.31%) reported, and we also had limited access to other demographic information due to lack of reporting by employers (eg, race and ethnicity, socioeconomic status, and previous mental health treatment). Thus, we were not able to examine the association between these factors and treatment response or stratify the analyses by age or gender. These additional analyses can provide valuable insights into those who may best benefit from virtual care. However, because the Ginger platform is offered through employers, we know that the survey respondents are working-age adults, suggesting that these findings may be generalized to the professional workforce and those enrolled in health benefits through their employer. In addition, we acknowledge that some members could be taking medication, which might affect their responses to care. We do not currently have medication use reported by our members. Thus, our results cannot conclude that the responses to care were not driven by the medication used in combination with therapy. Finally, because this was a retrospective observational study, we lacked a control group to infer the causality of the levels of engagement in treatment response.

The findings of this study generate additional research topics for future studies. Given that the aim of this study was to examine the impact of provider care, we did not specifically evaluate how different levels of use of our self-guided content would impact outcomes. Future research should focus on the impact of self-guided content independent of provider care. In addition, text-based coaching protocols can vary significantly, and some use highly templated messages that facilitate a more efficient coach workflow [16]. A content analysis of coaching messages could provide additional insights into the underlying mechanisms driving symptom improvement in on-demand care.

Conclusions

Our real-world observational study found that members who completed text-based coaching achieved full treatment responses at similar rates compared with members who completed clinical care and members who completed a hybrid of coaching and clinical care. There were no significant differences in the predicted probabilities of a full treatment response for both anxiety and depression between text-based coaching and clinical care. The highest level of use within each care modality cohort generally had increased odds of treatment response compared with minimal use, with the exception of the Hybrid Care cohort. The results support the utility of digital behavioral health interventions and further highlight text-based coaching protocols as an accessible and suitable option when considering virtual care for treating anxiety and depressive symptoms. Future studies should investigate optimal levels of use.

Acknowledgments

The authors are grateful for the participation of the members who completed the surveys, as well as the hard work and dedication of Ginger coaches and clinicians.

Conflicts of Interest

ES, SK, and GG were paid employees of Ginger/Headspace Health.

References

1. Mental Illness. National Institute of Mental Health. URL: https://www.nimh.nih.gov/health/statistics/mental-illness [accessed 2021-12-11]
2. Control of Neglected Tropical Diseases. World Health Organization. URL: https://www.who.int/teams/control-of-neglected-tropical-diseases/yaws/diagnosis-and-treatment/mental-health-and-substances-use [accessed 2021-12-07]
3. Kunkle S, Yip M, Hunt J, Ξ W, Udall D, Arean P, et al. Association between care utilization and anxiety outcomes in an on-demand mental health system: retrospective observational study. JMIR Form Res 2021 Jan 26;5(1):e24662 [FREE Full text] [doi: 10.2196/24662] [Medline: 33496679]
4. Vahratian A, Blumberg SJ, Terlizzi EP, Schiller JS. Symptoms of anxiety or depressive disorder and use of mental health care among adults during the COVID-19 pandemic - United States, August 2020-February 2021. MMWR Morb Mortal Wkly Rep 2021 Apr 02;70(13):490-494 [FREE Full text] [doi: 10.15585/mmwr.mm7013e2] [Medline: 33793459]
5. Bishop TF, Seirup JK, Pincus HA, Ross JS. Population Of US practicing psychiatrists declined, 2003-13, which may help explain poor access to mental health care. Health Aff (Millwood) 2016 Jul 01;35(7):1271-1277. [doi: 10.1377/hlthaff.2015.1643] [Medline: 27385244]
6. 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]
7. Madhusudhan DK, Watts SA, Lord DJ, Ding F, Lawrence DC, Sheldon A, et al. Employer-sponsored health centers provide access to integrated care via a hybrid of virtual and in-person visits. Telemed Rep 2021 Nov 2;2(1):247-257 [FREE Full text] [doi: 10.1089/tmr.2021.0027] [Medline: 35720749]
8. Dorstyn DS, Saniotis A, Sobhanian F. A systematic review of telecounselling and its effectiveness in managing depression amongst minority ethnic communities. J Telemed Telecare 2013 Sep;19(6):338-346. [doi: 10.1177/1357633X13501767] [Medline: 24163298]
9. Ruskin PE, Silver-Aylaian M, Kling MA, Reed SA, Bradham DD, Hebel JR, et al. Treatment outcomes in depression: comparison of remote treatment through telepsychiatry to in-person treatment. Am J Psychiatry 2004 Aug;161(8):1471-1476. [doi: 10.1176/appi.ajp.161.8.1471] [Medline: 15285975]
10. Linde K, Sigterman K, Kriston L, Rücker G, Jamil S, Meissner K, et al. Effectiveness of psychological treatments for depressive disorders in primary care: systematic review and meta-analysis. Ann Fam Med 2015;13(1):56-68 [FREE Full text] [doi: 10.1370/afm.1719] [Medline: 25583894]
11. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
12. Kivelä K, Elo S, Kyngäs H, Kääriäinen M. The effects of health coaching on adult patients with chronic diseases: a systematic review. Patient Educ Couns 2014 Nov;97(2):147-157. [doi: 10.1016/j.pec.2014.07.026] [Medline: 25127667]
13. Butterworth S, Linden A, McClay W, Leo MC. Effect of motivational interviewing-based health coaching on employees' physical and mental health status. J Occup Health Psychol 2006 Oct;11(4):358-365. [doi: 10.1037/1076-8998.11.4.358] [Medline: 17059299]
14. Grant AM. The impact of life coaching on goal attainment, metacognition and mental health. Soc Behav Pers 2003 Jan 01;31(3):253-263. [doi: 10.2224/sbp.2003.31.3.253]
15. Hill B, Richardson B, Skouteris H. Do we know how to design effective health coaching interventions: a systematic review of the state of the literature. Am J Health Promot 2015;29(5):e158-e168. [doi: 10.4278/ajhp.130510-LIT-238] [Medline: 24720388]
16. Lattie EG, Graham AK, Hadjistavropoulos HD, Dear BF, Titov N, Mohr DC. Guidance on defining the scope and development of text-based coaching protocols for digital mental health interventions. Digit Health 2019 Dec 16;5:2055207619896145 [FREE Full text] [doi: 10.1177/2055207619896145] [Medline: 31897306]
17. Hoermann S, McCabe KL, Milne DN, Calvo RA. Application of synchronous text-based dialogue systems in mental health interventions: systematic review. J Med Internet Res 2017 Jul 21;19(8):e267 [FREE Full text] [doi: 10.2196/jmir.7023] [Medline: 28784594]
18. Lindner P, Olsson EL, Johnsson A, Dahlin M, Andersson G, Carlbring P. The impact of telephone versus e-mail therapist guidance on treatment outcomes, therapeutic alliance and treatment engagement in Internet-delivered CBT for depression: a randomised pilot trial. Internet Interv 2014 Oct;1(4):182-187. [doi: 10.1016/j.invent.2014.09.001]
19. Berger T, Caspar F, Richardson R, Kneubühler B, Sutter D, Andersson G. Internet-based treatment of social phobia: a randomized controlled trial comparing unguided with two types of guided self-help. Behav Res Ther 2011 Mar;49(3):158-169. [doi: 10.1016/j.brat.2010.12.007] [Medline: 21255767]
20. 45 CFR 46. Office for Human Research Protections (OHRP). 2016 Feb 16. URL: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html [accessed 2022-01-27]

21. Kunkle S, Yip M, Ξ W, Hunt J. Evaluation of an on-demand mental health system for depression symptoms: retrospective observational study. J Med Internet Res 2020 Jun 18;22(6):e17902 [FREE Full text] [doi: 10.2196/17902] [Medline: 32554387]
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. Coley RY, Boggs JM, Beck A, Hartzler AL, Simon GE. Defining success in measurement-based care for depression: a comparison of common metrics. Psychiatr Serv 2020 Apr 01;71(4):312-318. [doi: 10.1176/appi.ps.201900295] [Medline: 31847739]
24. 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]
25. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. Gen Hosp Psychiatry 2016;39:24-31. [doi: 10.1016/j.genhosppsych.2015.11.005] [Medline: 26719105]
26. Johansson R, Andersson G. Internet-based psychological treatments for depression. Expert Rev Neurother 2012 Jul;12(7):861-870. [doi: 10.1586/ern.12.63] [Medline: 22853793]

Abbreviations

GAD-7: Generalized Anxiety Disorder-7
PHQ-9: Patient Health Questionnaire-9

Edited by A Mavragani; submitted 01.02.22; peer-reviewed by B Kaveladze, J D'Arcey, P Yellowlees; comments to author 25.02.22; revised version received 14.04.22; accepted 19.04.22; published 22.07.22.

Please cite as:
Shih E, Aylward BS, Kunkle S, Graziani G
Association Between Care Modality and Use With Treatment Response Among Members Accessing Virtual Mental Health Services: Real-world Observational Study
JMIR Form Res 2022;6(7):e36956
URL: https://formative.jmir.org/2022/7/e36956
doi:10.2196/36956
PMID:35867401

Original Paper

Translating Promoting Factors and Behavior Change Principles Into a Blended and Technology-Supported Intervention to Stimulate Physical Activity in Children With Asthma (Foxfit): Design Study

Annette Brons[1,2], MSc; Katja Braam[3,4], PhD; Aline Broekema[3], MSc; Annieck Timmerman[5], PhD; Karel Millenaar[6], MA; Raoul Engelbert[3,7], PhD; Ben Kröse[1,8], PhD; Bart Visser[3], PhD

[1]Digital Life Center, Amsterdam University of Applied Sciences, Amsterdam, Netherlands
[2]Department of Information and Computing Sciences, Utrecht University, Utrecht, Netherlands
[3]Centre of Expertise Urban Vitality, University of Applied Sciences Amsterdam, Amsterdam, Netherlands
[4]Faculty of Health, Sports and Social Work, Inholland University of Applied Sciences, Haarlem, Netherlands
[5]Topsport Amsterdam, Amsterdam, Netherlands
[6]Play and Civid Media, University of Applied Sciences Amsterdam, Amsterdam, Netherlands
[7]Department of Rehabilitation, Amsterdam University Medical Center, University of Amsterdam, Amsterdam, Netherlands
[8]Department of Computer Science, University of Amsterdam, Amsterdam, Netherlands

Corresponding Author:
Annette Brons, MSc
Digital Life Center, Amsterdam University of Applied Sciences
Wibautstraat 2-4
Amsterdam, 1091 GM
Netherlands
Phone: 31 621156976
Email: a.e.brons@hva.nl

Abstract

Background: Children with asthma can decrease the impact of their disease by improving their physical activity (PA). However, health care providers lack interventions for children with asthma that effectively increase their PA levels and achieve behavior change. A technology-supported approach can positively influence PA and physical functioning in children.

Objective: The aims of this study were to develop a technology-supported intervention that facilitates health care providers in promoting PA for children (aged 8 to 12 years) with asthma and to systematically describe this developmental process.

Methods: Intervention mapping (IM) was applied to develop a blended and technology-supported intervention in cocreation with children with asthma, their parents, and health care providers. In accordance with the IM framework, the following steps were performed: conduct a needs assessment; define the intervention outcome, performance objectives, and change objectives; select theory-based intervention methods and strategies; create components of the intervention and conduct pilot tests; create an implementation plan; and create an evaluation plan.

Results: We developed the blended intervention *Foxfit* that consists of an app with a PA monitor for children (aged 8 to 12 years) with asthma and a web-based dashboard for their health care provider. The intervention focuses on PA in everyday life to improve social participation. *Foxfit* contains components based on behavior change principles and gamification, including goal setting, rewards, action planning, monitoring, shaping knowledge, a gamified story, personal coaching and feedback, and a tailored approach. An evaluation plan was created to assess the intervention's usability and feasibility for both children and health care providers.

Conclusions: The IM framework was very useful for systematically developing a technology-supported intervention and for describing the translational process from scientific evidence, the needs and wishes of future users, and behavior change principles into this intervention. This has led to the technology-supported intervention *Foxfit* that facilitates health care providers in promoting PA in children with asthma. The structured description of the development process and functional components shows the way behavior change techniques are incorporated in the intervention.

Trial Registration: International Clinical Trial Registry Platform NTR6658; https://tinyurl.com/3rxejksf

(JMIR Form Res 2022;6(7):e34121) doi:10.2196/34121

KEYWORDS

intervention mapping; technology-supported intervention; mobile health; mHealth; tailoring; exercise; cocreation; social participation; gamification; mobile app; web-based dashboard; chronic disease; mobile phone

Introduction

Background

Physical activity (PA) is important for children with a chronic disease [1,2]. Asthma is the most frequently diagnosed medical condition in children [3]. For children with asthma, PA can positively affect their asthma control by improving their physical fitness. This can reduce the threshold of triggers causing asthma symptoms, which in turn may lead to decreased use of medication and increased quality of life [1,4-6]. Besides the positive effects of PA regarding physical fitness, stimulation of PA in children with asthma is also important for their psychosocial functioning, because PA is often intertwined with social life. Children with physical disabilities are less involved in social activities compared to their healthy peers, and children with asthma often feel left out of the group [7-9]. Despite all benefits of PA, children with asthma—especially girls—seem to be less physically active than their healthy peers [10-13].

In the Netherlands, children are treated for asthma according to the *Dutch Pediatric Society* guidelines for asthma in children [14]. This treatment includes taking medication, receiving information, and having regular check-ups. Children with asthma who are inactive are referred to health care providers, such as pulmonary nurse practitioners and pediatric physical therapists. They provide health education, encourage children to increase their PA levels, and pay attention to social participation and children's self-confidence regarding PA. Moreover, they educate them on how to cope with exercise barriers and provide exercise training. Such training interventions, which include, for instance, swimming and aerobic exercises, are shown to be effective in increasing cardiovascular fitness and in improving pulmonary function of children with asthma [6,15-19]. However, health care providers indicate that they lack interventions that result in maintenance of improved PA levels. This is because existing training programs do not aim to achieve behavior change in children with asthma [15,17,20].

This raises the question as to how PA levels can be both increased and maintained in children with asthma who are inactive. Various studies and reviews show that effectiveness of PA stimulating programs can be increased by technology-supported interventions [21-29]. Although most studies focus on adults, several studies show that a technology-supported approach can have positive effects on PA and physical functioning in children and adolescents as well [30,31].

Features of technology-supported interventions, especially behavior change interventions, are often inadequately described [23,24]. Therefore, it is difficult to compare interventions and their effects. To ease and improve this comparison, it is important to describe the design process and the developed components of an intervention [32,33]. Several frameworks are already used to describe interventions that stimulate PA, such as the MRC (Medical Research Council) framework to stimulate PA in older adults, the IDEAS (Integrate, Design, Assess, and Share) framework to promote PA in adults, the CeHReS (Centre for eHealth Research) framework to promote PA in people with depression, and the intervention mapping (IM) model to stimulate PA in patients with heart failure [33-40]. For our intervention, we chose the IM model, because it explicitly differentiates between personal and environmental factors affecting PA, and it takes into account creating an implementation and evaluation plan. Moreover, IM was already successfully applied for developing health interventions for children [41-43]. The framework enables systematic development of evidence-based interventions and enhances adoption, implementation, and maintenance of the developed intervention [39].

Objectives

As part of a project that aims to promote PA in children (aged 8 to 12 years) with asthma, we attempted to develop an intervention that supports health care providers in stimulating PA in children with asthma using the IM model. In a preliminary study, we explored promoting factors for PA in children with asthma according to children with asthma themselves, their parents, and their health care providers [44]. In this study, we applied the IM model to develop a technology-supported intervention that incorporates both these promoting factors and behavior change principles that were discovered. We aimed to transparently report the development process and the final components of the intervention using IM to meet the demand of sufficiently describing technology-supported interventions [32,33]. We answered the following research question: "How can scientific evidence on stimulation of PA, the needs and wishes of future users, and behavior change principles be translated into features of an intervention that facilitates health care providers in promoting PA in children with asthma?"

By systematically describing the translation from requirements to functional components and our specific design and implementation choices, we give insight into the background of our intervention and ease comparisons with other interventions. In addition, we report the contribution of future users, because the intervention was developed in cocreation with children with asthma, their parents, and health care providers.

Methods

IM Model

The IM protocol includes six steps: (1) conduct a needs assessment in which a distinction is made between personal and environmental determinants; (2) define the intervention outcome, performance objectives, and change objectives; (3) select theory-based intervention methods and strategies; (4) create components of the intervention and conduct pilot tests; (5) create an implementation plan; and (6) create an evaluation plan.

Step 1: Needs Assessment

To assess the health problem, background information and scientific evidence regarding children with asthma, the impact of low levels of PA, and available interventions had to be gathered according to the IM model. In our preliminary study, 3 stakeholder groups were questioned to explore promoting factors of PA in children with asthma. To do so, concept-mapping sessions were held with 25 children with asthma (aged 8 to 12 years), 17 parents of children with asthma, and 21 health care providers of children with asthma. The health care providers were lung nurse practitioners and pediatric physiotherapists, experienced in behavior change and supporting PA in children with asthma. The resulting factors were labeled as either personal or environmental factors according to the Physical Activity for People with a Disability model [45]. More details regarding the needs assessment can be found in our preliminary study [44]. On the basis of this needs assessment, we defined the aim of the intervention. The reported factors will be used in subsequent steps of the IM framework, such as designing the program's requirements and functional components.

Step 2: Intervention Outcome, Performance Objectives, and Change Objectives

To examine the target behavior, we created a logic model of change according to the IM framework. First, we defined the program outcomes based on the stakeholders' reported clusters in the needs assessment. They describe the desired changes in the behavior and the environmental conditions that are required to reach the goal of the intervention [39]. Second, we translated the program outcomes in performance objects based on the stakeholders' reported supporting ideas in the needs assessment. Those indicate the required behavior to reach the goal of the intervention. On the basis of these performance objectives, change objectives were defined which indicate the necessary behavior to accomplish the required behavior change and environmental conditions.

Step 3: Selecting Theory-Based Intervention Methods and Strategies

On the basis of the needs assessment and the logic model of change thus created, we selected several methods and strategies that were translated into design considerations for the intervention. This was performed in a stakeholder session with a pediatric pulmonologist, a pediatric physiotherapist, and researchers specialized in exercise behavior in children, behavioral change, health technology, and game development. The final selection was a product of mutual agreement.

Step 4: Creating and Pilot-testing the Intervention's Components

On the basis of the needs assessment, logic model of change, and the design considerations, we formulated several requirements for the intervention to be developed. With a multidisciplinary team involving health care providers, data scientists, game designers, and software developers, we translated the requirements into functional components. This translation was performed in 2 sessions and the final list of requirements was a product of mutual agreement. To develop those components, we performed an agile method consisting of several design and develop iterations. Each iteration lasted for 2 weeks, and there were 10 iterations in total.

In the first 3 iterations, we created story lines, mock-ups, and user journeys resulting in a prototype. In a stakeholder session, this product was tested by children with asthma and their health care providers. Children who participated in the needs assessment were asked whether they wanted to participate in this stakeholder session as well. Children who were diagnosed with asthma and aged between 8 and 12 years could be included. The 3 participating children were aged, on average, 9.3 (SD 1.5) years, and 2 of them were female. One of the children was diagnosed with severe asthma, one of them had mild asthma symptoms, and one of them had moderate asthma symptoms. The 4 participating health care providers were all pediatric physiotherapists with >5 years of work experience with children with asthma. In the stakeholder session, the participants individually tested the prototype and shared their experiences with the researchers. The children were asked to focus on the app's functionalities, the story line, and the usability of the PA monitor, whereas the health care providers were asked to focus on the dashboard's functionalities and the clarity of the graphs showing the children's PA. Afterward, the researchers compared the experiences of the participants and translated them into requirements for improvement of the prototype. In the remaining 7 iterations, these improvements were made, and the prototype was further developed into a usable product.

Step 5: Creating an Implementation Plan

In a stakeholder session with a pediatric physiotherapist, a pediatric pulmonologist, and a lung nurse practitioner, we examined the health care provider's daily practice and the current health care situation for children with asthma who are inactive receiving care regarding PA. We explored the available time and skills of the health care providers, to support children with asthma who might benefit from PA with an intervention.

Step 6: Creating an Evaluation Plan

For a blended intervention, it is important to evaluate the usability and feasibility in children's everyday life and in their health care provider's daily practice. Therefore, we selected methods to evaluate user experiences, adherence, and usability of both children and health care providers. In addition, to evaluate the possible impact of the intervention on children, we selected methods to look into the self-perception of PA among children, their social participation, and their enjoyment of PA.

Ethics Approval

The Amsterdam University Medical Center Medical Ethical Committee approved the research protocol with trial number METC 2017 191. The study was registered in the International Clinical Trial Registry Platform with trial number NTR6658 on August 21, 2017.

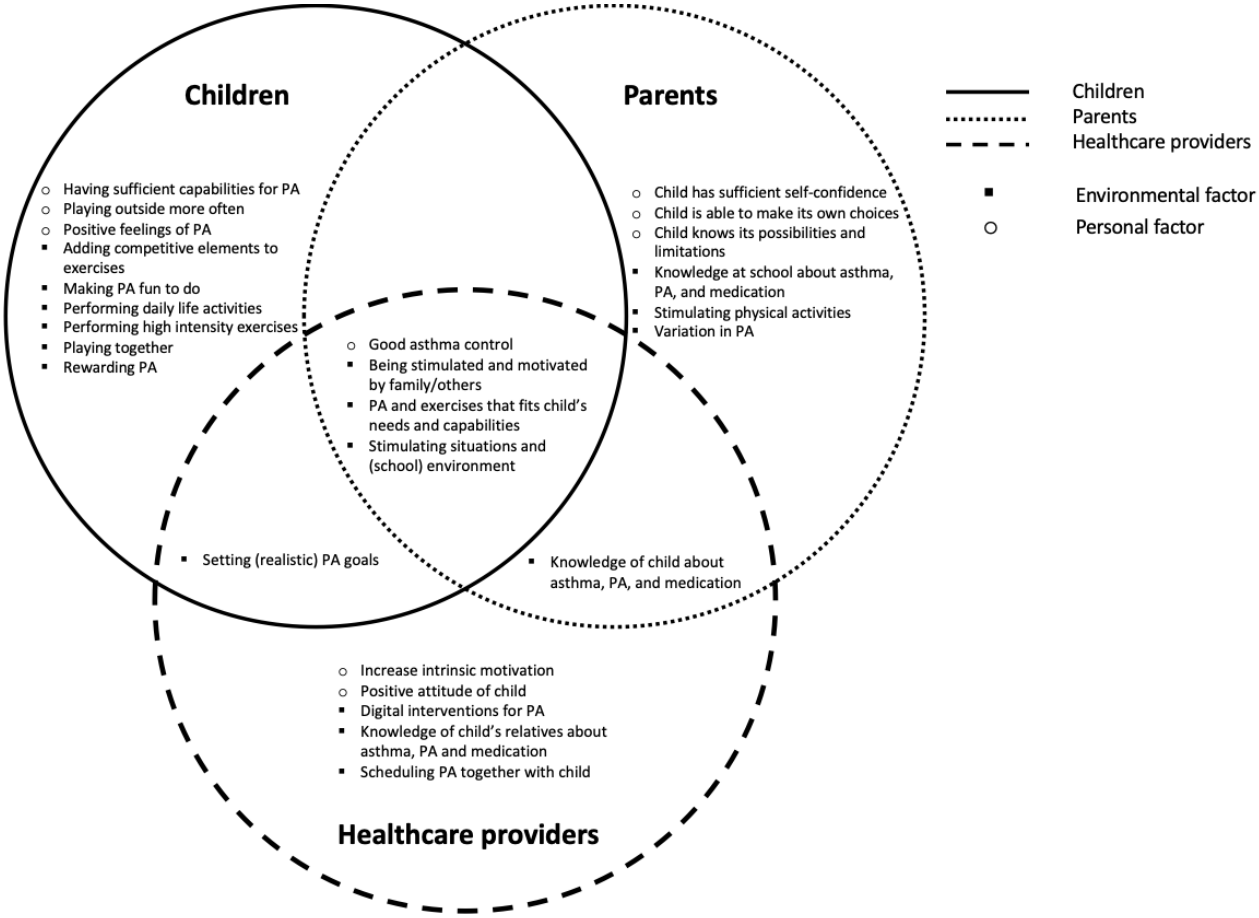
Results

Step 1: Needs Assessment

Several studies show that children with asthma benefit from PA. Increased levels of PA are related to both improved cardiovascular fitness and quality of life [4-6,16,19]. However, children with asthma experience limitations in PA and social activities because of their disease [7-10,46]. Reported explanations for these experienced limitations are, for instance, inaccurate symptom perception of airflow limitation, misinterpreting healthy exercise-induced shortness of breath as an asthma attack, and low self-efficacy in relation to PA [47,48]. Parents endorse the limitations that children experience and report parental fear for exercise-induced asthma, challenges with asthma management, and lack of trust in teachers or sport coaches as parental barriers [49,50]. Increasing self-esteem and improving knowledge about PA in relation to asthma and medication are reported as strategies to reduce children's perceived limitations in PA [51,52].

Figure 1 shows differences and similarities among all reported factors of the 3 stakeholder groups. Factors that were rated as very important were getting positive feelings of PA, playing together, being stimulated and motivated by relatives, good asthma control, making PA enjoyable, rewarding PA, tailored PA, playing outside more often, setting realistic goals, relatives of child have sufficient knowledge about asthma in relation to PA, child itself has sufficient knowledge about asthma in relation to PA, digital interventions for PA, increasing extrinsic motivation, positive attitude of child, scheduling PA together with child, and variation in PA.

Figure 1. Venn diagram in which the reported factors of all stakeholders are combined. The circles indicate the stakeholder groups. The factors are listed in random order (reprinted from Brons et al [44] with permission from the author). PA: physical activity.



The aim of the intervention was defined as follows: the intervention needs to support health care providers in improving PA levels in children with asthma who might benefit from increased levels of PA. The intervention will be developed for children who are inactive, with mild to moderate asthma symptoms, who are referred to a health care provider, either a lung nurse practitioner or a pediatric physiotherapist, to improve their PA levels.

Step 2: Intervention Outcome, Performance Objectives, and Change Objectives

All program outcomes, along with their corresponding performance and change objectives, are shown in Table 1.

Table 1. Logic model of change: program outcomes with their corresponding performance objectives and change objectives[a].

Program outcomes and performance objectives	Change objectives
Increase PA[b] participation	
Use active transport instead of motorized transport	• Walk or cycle more often
Participate in sports activities at school	• Communicate about willing to join sports activities • Inform school about the disease to create a safe environment
Participate in social activities with PA elements	• Register for social PAs, even when having some doubts • Find a trusted buddy that can help when needed
Participate in sports activities	• Register for sport activity that fits capabilities • Inform the coach about the disease to create a safe environment
Resist sedentary behavior	• Search for enjoyable alternatives for sedentary time • Screen time should be reduced (by parents)[c]
Improve knowledge about asthma and medication in relation to PA	
Gain knowledge about impact of PA and medication on asthma control	• Information about asthma and medication in relation to PA should be provided by health care provider or parents[c]
Gain knowledge about PA possibilities despite having asthma	• Information about PA possibilities when having asthma should be provided by health care provider[c]
Increase asthma control	
Good medication adherence	• Take medication as prescribed
Increase self-esteem regarding PA	
Experience positive PA feelings	• Learn individual capabilities and limitations
Increase perceived joy during PA	
Play together	• Ask friends or family to participate in PA together • Schedule PA in family agenda
Include competition element	• Add competitive elements to PA • PAs with competitive elements should be available[c]
Introduce variation in PA	• A variety of PAs should be accessible[c]
Earn rewards	• Rewards should be given for performed PA[c]
Provide a stimulating environment	
Receive support from parents	• PA should be actively supported by parents[c]
Create a supportive school environment	• PA should be actively supported by school[c] • Access to PAs should be given by school[c]
Provide tailored PA	
Provide an adaptive tool	• A flexible tool that adapts to the child's specific needs and wishes should be available[c]
Receive tailored feedback	• Individual feedback based on monitored PA behavior should be given by health care provider[c]
Learn to set realistic goals	
Set PA goals	• Think about own PA behavior and formulate a goal to improve the current behavior • Obtain insights into one's own PA behavior • Coaching regarding setting realistic goals should be provided by health care provider or parents[c].

[a]Change objectives regarding personal factors (shown in all cells unless otherwise indicated by footnote c).

[b]PA: physical activity.

[c]Change objectives regarding environmental conditions.

Step 3: Selecting Theory-Based Intervention Methods and Strategies

The following three design considerations were formulated: (1) support behavior change, (2) blended technology, and (3) integration in everyday life.

To start with, the intervention should support behavior change regarding PA. Gamification is frequently used in the improvement of health and well-being and has shown to be particularly effective in behavior change regarding PA [53]. Gamification is defined as "the use of game design elements in non-game contexts" [54]. Conceptually, gamification combines design elements for behavior change techniques from persuasive technologies, intrinsically motivating qualities from serious games, and methods for tracking individual behavior from personal informatics [53,55-59]. However, in practice, most gamified eHealth applications merely implement short-term engagement through extrinsic rewards [60]. For children with a chronic disease specifically, the nonsignificant results of several serious games aiming to increase PA might be explained by the lack of incorporating behavior change theories in the game [61]. To reach the full potential of gamification, it is necessary to design tools based on theories that support the overall goal, behavior change, and the effects of game mechanics [60].

To maintain improved levels of PA, children must develop new habits. An important factor in habit formation is rewarding the required behavior [62]. Rewarding increases positive feelings in relation to the behavior, in our case performing PA [63]. To develop intrinsic motivation, such positive feelings are important [64]. Increasing self-efficacy is also shown to be effective, for instance, through action planning and shaping knowledge [65-67]. Moreover, goal setting, monitoring and feedback are shown to be effective persuasive methods for improving PA behavior [67-69]. Personal coaching from health care providers, for instance, with the widely used motivational interviewing technique, might enhance the behavior change process [70,71].

Second, the intervention should use blended technology, meaning that individual use of digital technology is combined with face-to-face interaction with a health care provider. Digital technology was shown to be valuable for improving health in children and supporting behavior change [30,72]. However, the results of such digital technologies can be optimized by offering them in combination with personal coaching [30]. Personal coaching is important to provide feedback and support goal setting, which are both identified as success factors when behavior change is aimed [67-69]. Moreover, personal coaching by health care providers reinforces a tailored approach, which was also identified as a success factor of eHealth interventions that induce health behavior change [73-76]. Both a tailored approach and support of a health care provider are important supporting factors of shaping knowledge, which is a frequently used technique in supporting healthy behaviors [67]. Children with asthma specifically need to learn more about the positive relation between PA and asthma control, how to properly take asthma medication, and the difference between healthy shortness of breath and an asthma attack [47]. Moreover, feedback and knowledge regarding accurate symptom perception is important for perceived exercise limitations. Some children report substantial discomfort when there is limited bronchoconstriction, whereas others report no symptoms even if severe obstruction is present [48].

Finally, it is important that use of the intervention and performing PA could be integrated in everyday life of the users. Although participation in everyday activities is a vital part of children's development, children with physical disabilities or chronic diseases are less involved in such social activities than their healthy peers [8,58,77,78]. Children with asthma report that they often feel left out of the group [9]. Integration in everyday life is therefore important from 2 perspectives. On the one hand, we have to prevent that using the intervention creates a special situation as this might reinforce children's feeling of being different and excluded. On the other hand, PA is often intertwined with social life. By focusing on PA in everyday life situations instead of sports specifically, children might experience better social participation.

Step 4: Creating and Pilot-testing the Intervention's Components

Requirements of the Intervention

The requirements of the intervention are as follows:

1. Facilitate behavior change regarding PA by supporting goal setting, action planning, behavior execution, self-monitoring, and evaluation.
2. Monitor children's PA so that both children and their parents can obtain insights into their PA behavior.
3. Connect the technology used by the children with the technology used by the health care providers. That way, health care providers can obtain insights into children's PA behavior and can provide feedback. However, participating children should not see PA behavior of other children to prevent them from comparing themselves with each other. Children's PA possibilities vary by the severity of their asthma symptoms, and their motivation might be lowered when they see other children achieving PA goals that are not achievable for themselves.
4. Promote PA in everyday life to increase social participation instead of focusing on sports activities.
5. Include learning elements to improve children's knowledge about asthma in relation to PA. This should be presented in an attractive and stimulating way, because children with asthma themselves did not report improving their knowledge as an important stimulating factor.
6. Reward both exhibited PA and the effort undertaken to extrinsically motivate children. Although intrinsic motivation is preferred over extrinsic motivation to achieve behavior change, extrinsically motivated PA can yield positive feelings such as experiencing fun and feeling

competent. Those positive experiences are in itself important factors to become intrinsically motivated.

7. Include adaptive components for a tailored approach such as personalized goal setting, planning, and education. Personalized goals are important to adapt to children's specific situation.
8. Minimize required screen time for the intervention. There should only be functional screen time, because screen time induces sedentary behavior and therefore defeats the purpose of the intervention.

Functional Components of the Intervention

Structure

The final intervention, named *Foxfit*, consists of 3 technical elements that meet all the requirements: (1) a PA tracker that monitors children's performed PA time and intensity, (2) the *Foxfit* app on the child's smartphone, and (3) the *Foxfit* web-based dashboard for health care providers. Because the intervention is blended, children receive personal coaching and feedback from their health care provider every week during a coaching moment. In the gamified story, the children are a fox that wants to travel from the moon to a planet. The child can do so by gaining points by performing PA.

In the following description of functional components, it is explicitly stated when components were added or changed because of feedback on the tested prototypes. To emphasize the evidence-based components, Table 2 shows the mapping from methods regarding the design consideration, such as behavior change principles, to the final functional components.

Table 2. Translation from methods in the 3 design considerations to functional components of the intervention.

Design consideration and method	Functional component of the intervention
Support behavior change	
Monitoring	• Activity monitor that shows PA[a] • Points in the app that represent PA levels and bonus activities • Overview of performed activities over the last week
Gamified story	• PA behavior, goals, and rewards are translated to an attractive story
Rewards	• Points for PA behavior based on activity tracker • Bonus points for being aware of PA behavior • Trophies from the health care provider for effort and positive attitude
Action planning	• Personal daily and weekly schedule with PAs
Shaping knowledge	• Tips and information as attractive drawings with supporting text • Tips and information from the health care provider during the weekly meetings
Goal setting	• Personal daily and weekly goals in the form of PA points • Personal PA goal for the entire duration of the intervention
Blended technology	
Personal coaching and feedback	• Weekly meeting with health care provider
Tailored approach	• Goals, activity schedules, suggestions for PA, and tips and information are tailored
Integration in everyday life	
Include all PAs	• Suggestions for PAs • Activity tracker monitors both general PA and specific sports activities
No special situation	• Usable without having the app with you (synchronize afterward)

[a]PA: physical activity.

Goal Setting

A total of 4 PA goals can be listed up on the dashboard and the most important one can be marked as the main goal, such as being able to play as a hockey field player instead of goal keeper for half a match. Moreover, a weekly goal can be set. Children can see both their main and weekly goals on the home screen of the app anytime.

Rewards

On the basis of the measurements of the activity tracker, children receive points for their performed PA. Children also receive rewards for their effort, because children that tested a pilot version indicated that they did not only want to be rewarded for actual performed and measured PA but also for a positive attitude and for trying without sufficient results. Therefore, their shown effort is rewarded by their health care providers with either a bronze, silver, or golden trophy.

Self-monitoring

The activity tracker facilitates self-monitoring of PA. The app shows the received points visualized as the location of the fox on its way to the planet. In addition to self-monitoring, the child's PA behavior is visualized in graphs on the health care provider's dashboard. As requested by health care providers who tested the prototype, these graphs are easy to interpret, because they do not have much time to prepare. Moreover, at the request of health care providers, the graphs also show the relation between PA behavior and the asthma symptoms that children report every day. Because the graphs are too complicated for children aged <12 years to interpret on their own, they are shown only on the health care providers' dashboard and not in children's smartphone app.

Awareness of PA and Symptoms

Children are rewarded for the combination of filling out that they are going to be physically active and actually doing so. Moreover, the activities that are registered by the children become visible in an overview of performed activities in their app. Health care providers who tested a prototype indicated that children must learn to recognize their asthma symptoms and the relation among their asthma symptoms, PA, and medication. Therefore, the children's asthma symptoms are monitored by reporting how they feel every morning and evening. On the health care provider's dashboard, these scores are combined with the children's PA behavior.

Suggestions for PA

In the needs assessment, children indicated that they often have motivation to become physically active, but they simply do not know what to do. Therefore, the app has a button that randomly shows suggestions from a list of PAs. The list contains both personalized and general ideas based on input of children participating in the needs assessment and testing the prototype. The personalized items can be added and adjusted every coaching moment with the health care provider.

Tips and Information

According to the needs assessment, health care providers and parents find it important that children learn more about the relation among their asthma symptoms, PA, and medication. Information regarding these topics are shown in informative and attractive drawings together with short supporting texts with practical tips for everyday life. These drawings pop up after children fill out that they are going to perform a PA in their app.

Schedule PA

A weekly schedule is visible in the app and the current day's schedule is shown every morning.

Personal Coaching of Health Care Provider

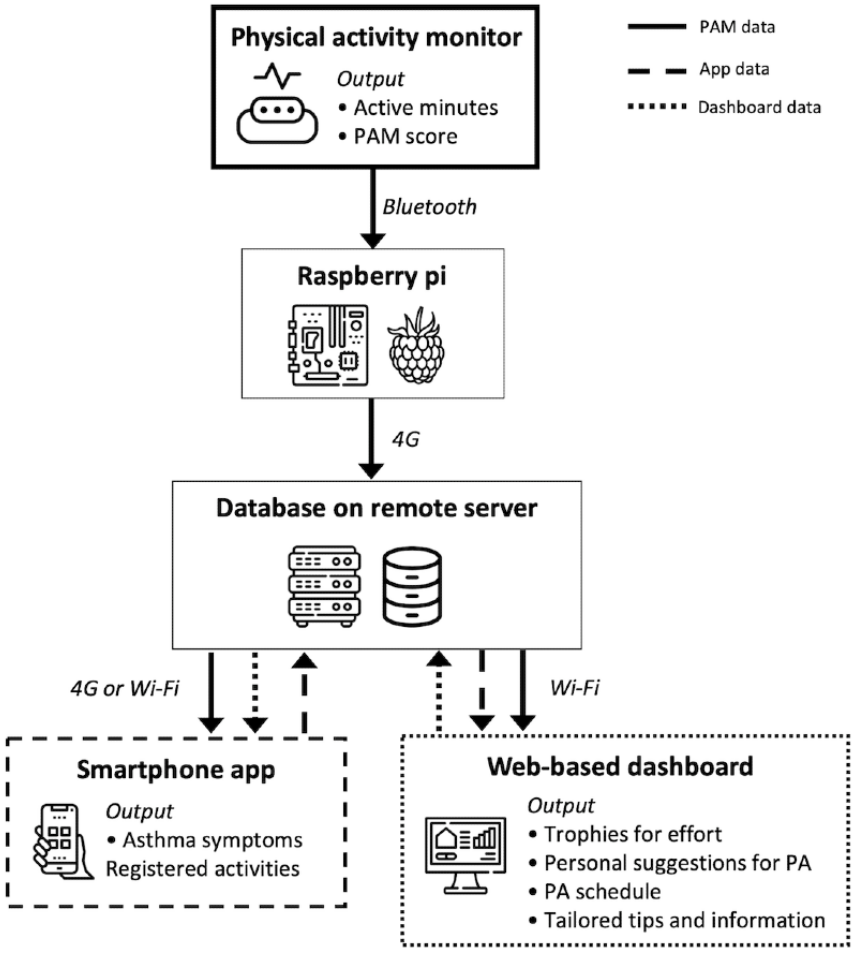
Every week, children speak to their health care provider either in person or via a video call. On the basis of the visualizations on the health care providers' dashboard, they discuss the child's PA behavior of the past week, including the relation between the performed PA and the registered asthma symptoms. Moreover, the health care provider can adjust the personalized items in the dashboard such as the child's PA goals, the informational topics, and the PA schedule.

Step 5: Creating an Implementation Plan

Technical Structure

The three elements of *Foxfit* (the activity tracker, smartphone app for children, and web-based dashboard for health care providers) must communicate with each other to synchronize. Figure 2 shows an overview of the connections among all technical elements.

Figure 2. Overview of all technical elements. The arrows indicate data traffic between elements. PA: physical activity; PAM: physical activity monitor.



PA Tracker

Children wear an activity tracker that monitors their PA levels. After comparing several activity trackers, we chose the PA monitor (PAM) "PAM AM300" [79]. This instrument is shown to be valid for adults [80]. Although there is no information on the validity of the PAM for children, we expected the tracker to be applicable because in our intervention the PAM is primarily used to measure progress and relative differences. Moreover, the instrument met several important criteria for our intervention. The instrument gave the opportunity to send data to our private research database instead of a commercial organization, had positive reviews from young study participants, is reasonably priced, and is extremely easy to use, because it does not require charging and synchronizes automatically [81,82].

Children wear the PAM on their hip, where it is clipped on their waistband. The PAM measures their activity in 2 ways. First, their activity is represented as a PAM score indicating the ratio of energy expended through PA to resting metabolism [83]. Second, their daily activity is represented as active minutes per day. Those active minutes are classified based on the metabolic equivalent of task (MET) as either low intense (MET: 1.8-3), middle intense (MET: 3-7), or high intense (MET: >7). On the basis of this, bonus points are given: (1) 5 points for either 60 minutes of low intense activity or 90 minutes of middle activity (the health care provider chooses which rule fits the child best), (2) 5 points for at least 15 minutes of high intense activity, and (3) 5 points in case children registered an activity and their PAM tracks at least five middle or high intense activity minutes. Children's total daily score is the sum of this PAM score and the gained bonus points.

To synchronize the smartphone app for children and the web-based dashboard for health care providers, the data collected from the PAM must be stored in a database. Therefore, children have a Raspberry Pi (Raspberry Pi Foundation) with a Bluetooth and 4G dongle at home. When the PAM is near the Raspberry Pi, it automatically sends the collected data to a database on a secured remote server via a secured Secure Shell (SSH) connection. The collected data are not shared with other parties, no personal data are stored, and the collected activity data are stored pseudonymized.

Web-Based Dashboard for Health Care Providers

Health care providers visit a web-based dashboard via their internet browser to see visualizations of the collected activity data and to personalize the intervention. Every health care provider has their own ID number and password to access the dashboard. These ID numbers and corresponding hashed passwords are stored in the database on the remote server. ID numbers of health care providers are only matched to ID numbers of children that they treat at the moment to ensure they do not have access to children who are treated by other health

care providers. Figure 3 shows the pages of the dashboard where health care providers log in and select the child's ID.

When a health care provider signs in to the dashboard and selects a child's ID, the child's activity data are sent via a secured SSH connection from the database on the remote server to the web-based dashboard. There, visualizations of the child's PA behavior are shown. Figure 4 shows such visualizations on the health care provider's dashboard. Health care providers also use the dashboard to tailor the intervention. Figure 5 shows pages of the dashboard in which the health care provider can fill out personalized data. When they add information or fill out choices, this is again sent via a secured SSH connection from the web-based dashboard to the database on the remote server.

Figure 3. Start pages of the health care providers' web-based dashboard. Log in to dashboard (top); select the ID of the child (bottom).

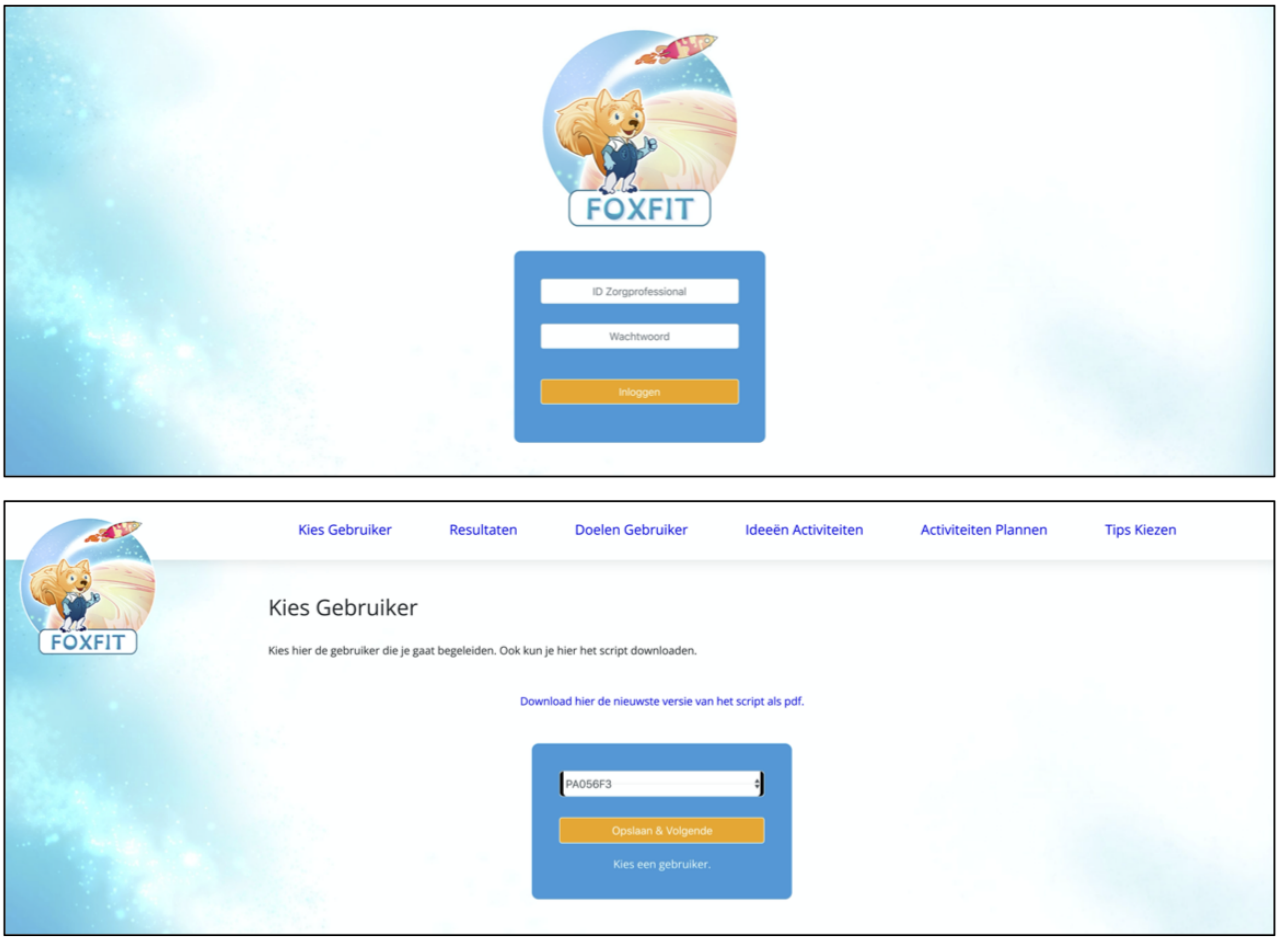
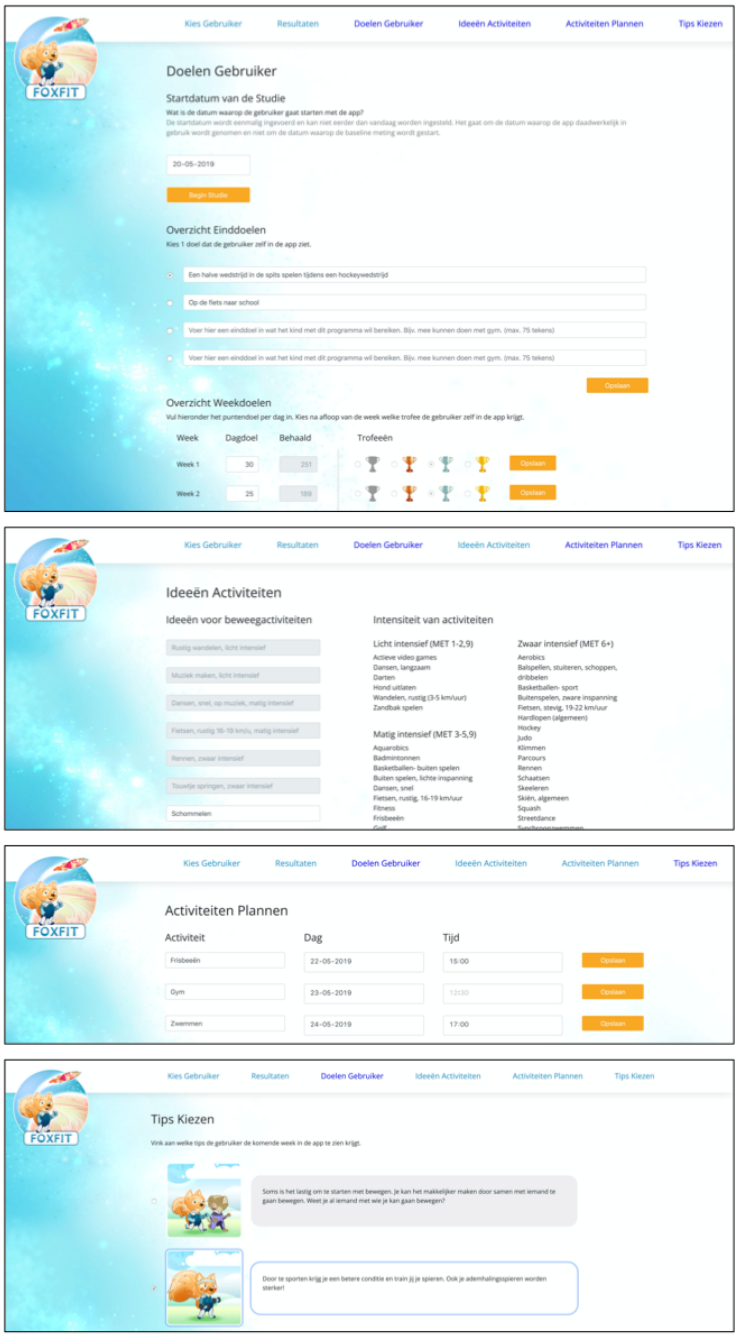


Figure 4. Visualization pages of the health care providers' web-based dashboard. Graphs regarding the physical activity (PA) behavior of the child are shown. From top to bottom: PA behavior of the baseline week (“Week 0”); PA behavior of the fifth week using the app (“Week 5”). In both graphs, the date is shown on the x-axis (“datum”), the y-axis represents the amount of PA minutes (“minuten”), and the intensity of PA (“intensiteit”) is represented by the color—low (“licht”); middle (“middel”); high (“zwaar”). In the bottom graph, the smiley faces indicate whether the child has achieved their daily goal, and the number between 1 and 10 indicates the experienced asthma symptoms in the morning (1=feeling very bad; 10=feeling very well).

Figure 5. Personalization pages of the health care providers' web-based dashboard. From top to bottom: personal goals and trophies; physical activity (PA) suggestions; scheduling PAs; choosing relevant tips.



Smartphone App for Children

Because the *Foxfit* app is built only for Android, and because we wanted to prevent from privacy and installation problems, children receive a smartphone on which the *Foxfit* app is already installed. To see the collected activity data as a gamified story, children use the app. Figure 6 shows the pages of the child's smartphone app. Children have their own ID number and password to access the app and the app is connected with the internet through Wi-Fi or 4G. On the remote server, no personal data of the child are stored. Only the ID number and corresponding hashed password are stored. When a child opens the app, activity data from the PAM that were automatically synchronized via the Raspberry Pi at home are sent from the remote server to the smartphone app through a secured SSH connection. When children add information or fill out choices in the app, this is also sent via a secured SSH connection from the smartphone app to the remote server. The home screen contains a help button that provides contact information to handle technical problems.

Figure 6. Pages of the smartphone app for the children. From left to right and from top to bottom: home screen; starting the day and filling out symptoms experienced in the morning; filling out physical activity (PA); overview of PAs of the current week; ending the day and filling out symptoms experienced during the day; earned trophies.

Implementation in the Health Care System

The intervention takes 7 weeks in total. Health care providers have a coaching moment with the child every week. The first 2 meetings, the midterm meeting, and the last meeting are in-person sessions and take 30 minutes. The remaining meetings are digital, to diminish the time investment and take 15 minutes. Before starting the intervention, children wear the PAM for a week without having access to the app. In this “week 0,” their baseline PA behavior is discovered. In the following 6 weeks, the full intervention including the smartphone app is used by the child. The health care provider can access the web-based dashboard at any time in the 7 weeks. At the start of a coaching meeting, the health care provider signs in into the web-based dashboard with their own credentials and selects the ID number of the child that the meeting is with. The dashboard is designed in such a way that the health care provider is guided through all steps to be taken. All steps and tasks to be performed by the

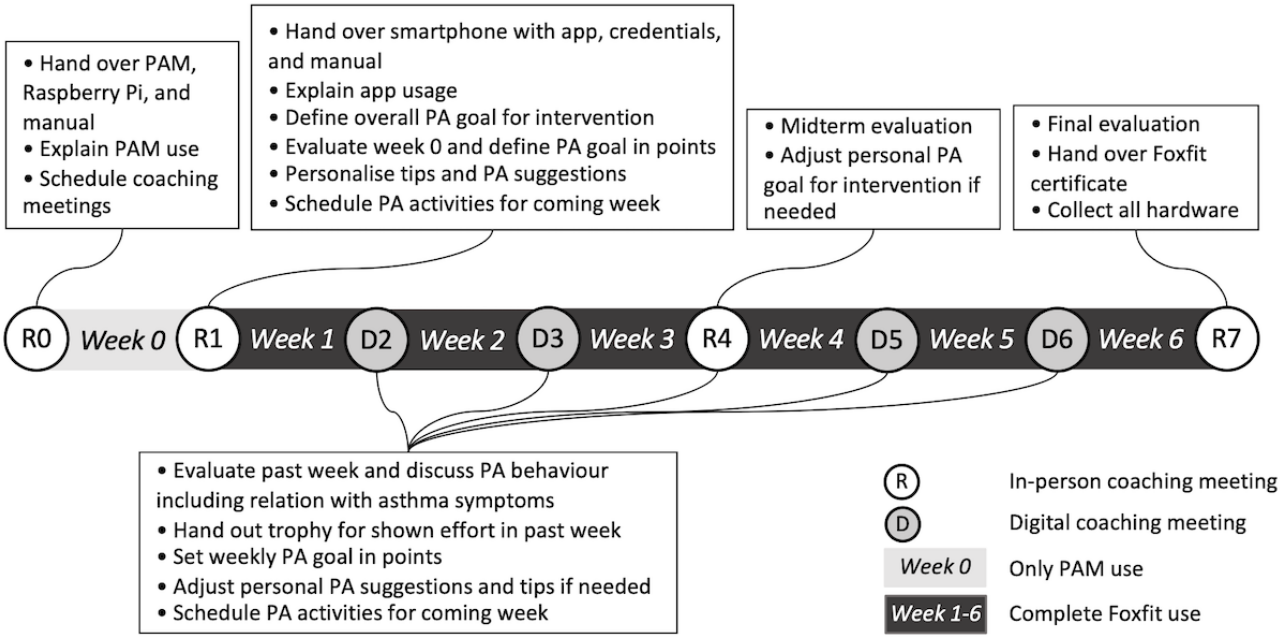
health care provider are described in a manual as well. Figure 7 shows a timeline of the *Foxfit* intervention from the health care provider's perspective.

In the first meeting, health care providers hand over the PAM, Raspberry Pi, and a manual to the child. They introduce the intervention and explain the PAM use to the child and their parents. Moreover, they schedule coaching moments for the coming weeks. After the baseline week is complete, health care providers hand over a smartphone on which the *Foxfit* app is installed, along with the child's credentials and a manual. They explain the intervention to the child and their parents. Together with the child, a personal PA goal for the entire intervention is set as well as a weekly goal in the form of activity points. These goals are filled out in the dashboard. Moreover, personal tips and suggestions for PAs are chosen, and PAs for the coming week are scheduled.

In the digital coaching meetings, health care providers evaluate the PA behavior of the child over the past week, based on the PA graphs on their dashboard. They also discuss the relation between the child's PA levels and their asthma symptoms. For the child's shown effort in the past week, they hand out a trophy. Together with the child, the health care providers set a new weekly PA goal formulated as activity points and schedule PAs for the coming week. If needed, PA suggestions and personal tips can be adjusted in the dashboard as well.

The coaching meeting after 3 weeks of using the app is conducted in real life, instead of being conducted digitally. Then, health care providers conduct a midterm evaluation in addition to the general items covered during a coaching moment. If needed, they can adjust the child's PA goal. In the last meeting, health care providers conduct a final evaluation in which they discuss the child's progress achieved during the intervention. In addition, they hand out a *Foxfit* certificate to the child and they collect all hardware.

Figure 7. Timeline of the *Foxfit* intervention in the daily practice of the health care providers. PA: physical activity; PAM: physical activity monitor.



Implementation in Children's Everyday Life

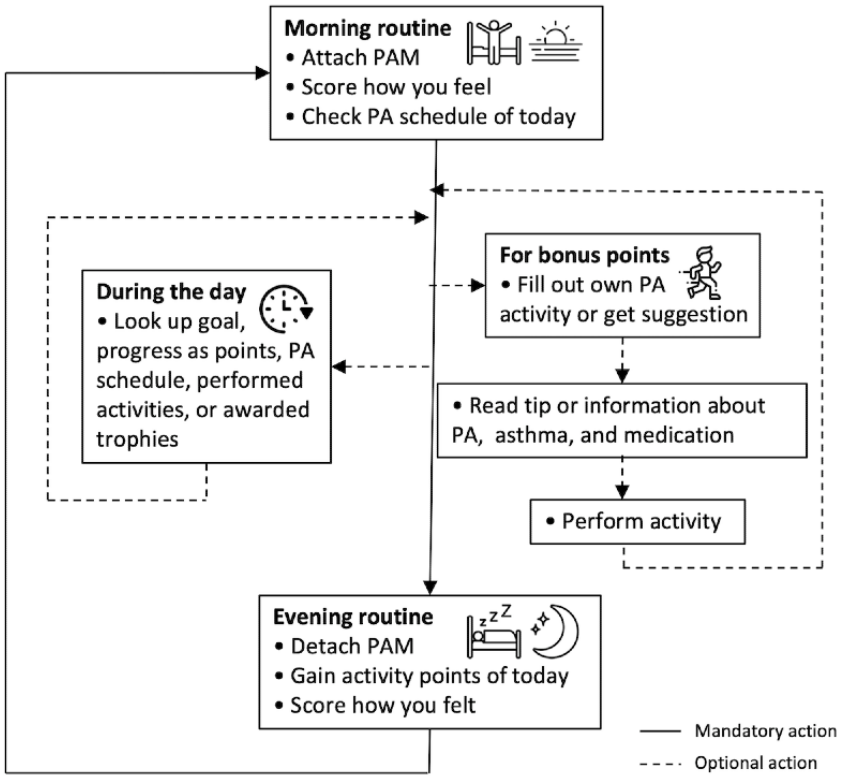
After the baseline week in which children's baseline PA behavior is discovered with the PAM, children also use a smartphone with the *Foxfit* app. Children need not carry around their smartphone all the time. To minimize screen time, the app is not required to track PAs, because the PAM automatically tracks their PA levels. When they are near the Raspberry Pi at home, data are automatically sent to the database on the remote server. Figure 8 shows a timeline of the *Foxfit* intervention in children's everyday life.

In the morning, children attach their PAM to their waistband to track their activity during the day. After that, they open the smartphone app which shows the menu and their personal PA goal. Then, they fill out how they feel that morning to track their asthma symptoms. When they are finished, their PA schedule of the day is shown, and they can start their day.

Although the app is not required to track PA behavior, children can use the app at any moment to register their PAs to become more aware of their own PA behavior. To do so, they open the app and fill out the activity they intend to perform at that moment. After registering, they receive a tip or information on subjects applicable to them. Children can use the app at any moment to look up their personal goal, received points, PA schedule, performed activities, or awarded trophies as well.

Around bedtime, children finish the *Foxfit* day by taking off the PAM and by gaining their daily points. When they open the app, they see the points received during the day. Their spaceship travels from the moon toward the planet based on the received PA points. Similar to the morning routine, children fill out how they felt during the day to track their asthma symptoms.

Figure 8. Timeline of the *Foxfit* intervention in the everyday life of the children. The arrows indicate flows between elements. PA: physical activity; PAM: physical activity monitor.



Step 6: Creating an Evaluation Plan

The developed intervention *Foxfit* will be tested by approximately 15 children (aged 8 to 12 years) with asthma who are inactive. They will all be treated in a hospital in the Netherlands, specialized in pediatric asthma. During the evaluation study, the intervention will be provided by pediatric lung nurse practitioners and functional exercise therapists from those hospitals.

To evaluate the usability and feasibility in children's everyday life and in health care providers' daily practice, both children and health care providers will be interviewed halfway and at the end of their participation about their user experience, adherence, and suggestions for improvement. In addition to the qualitative interviews, they also fill in the System Usability Scale, which measures the usability of the developed intervention [84,85].

To obtain insights into the effect of the intervention, children will fill out several questionnaires at the start and end of the intervention. Those questionnaires regard asthma control, PA participation, enjoyment of PA, and self-perceptions with, respectively the Asthma Control Questionnaire, Activity Questionnaire for Adults and Adolescents, PA enjoyment scale, and the self-perception profile for children (competentie-belevingsschaal voor kinderen) [86-91]. Regarding children's physical fitness and endurance, they perform a 6-minute walk test. In addition, activity data, collected with the PAM, of all participants will be analyzed. We will compare activity levels of the baseline week with activity levels of the last intervention week.

A detailed description of the evaluation together with its corresponding analyses and results will be published separately.

Discussion

Principal Findings

This study illustrated the way IM was used to translate theory and stakeholder experiences into design requirements and functional components and finally develop a complete blended and technology-supported intervention. IM was useful for describing the translation and development in a structured way. The needs and wishes of children with asthma, their parents, and health care providers were combined with scientific evidence regarding behavior change techniques and gamification in a technology-supported intervention. This resulted in the blended intervention *Foxfit* that consists of an activity tracker, a smartphone app for children (aged 8 to 12 years), and a web-based dashboard for health care providers. It focuses on PA in everyday life to improve social participation and contains behavior change elements such as goal setting, rewards, action planning, monitoring, shaping knowledge, a tailored approach, personal coaching and feedback, and a gamified story. Because of the cocreation with future users, the support of the resulting product has been strengthened among children with asthma, their parents, and health care providers. The structured description of the development process ensures that *Foxfit* can be compared easily with other interventions.

Strengths, Limitations, and Future Work

The participation of stakeholders in this project was very strong, because the intervention was not only created for the target group but together with them. Consistent with other studies that

applied IM to develop health interventions for children, we experienced that the process of IM is very time consuming, but the systematic approach improves clarity and analysis of the intervention [41,42]. The theoretical foundation increases the potential of the intervention to realize the desired outcome of supporting PA, because it tends to provoke more significant effects on PA [30,65,92,93]. Moreover, clarity and analysis of the intervention is important, because many gamified interventions for children with chronic disease lack a detailed description of the final intervention [61].

Because of the structured approach, it is relatively easy to apply this approach to other situations. On the one hand, the current gamified story and integrated behavior change techniques can be applied for children with other chronic diseases. Only the content of the knowledge components and the specific questions about asthma symptoms should be adjusted to the new disease. On the other hand, the integrated behavior change techniques can be applied for children in older age groups in case the gamified story is adjusted to their interests and experiences. Because children should be able to handle a smartphone with the *Foxfit* app themselves and understand why they need to wear the activity tracker, we do not think that the intervention is suitable for children younger than the current minimum age of 8 years.

An important factor of *Foxfit* is the support of the health care provider. Because of the blended approach, children receive personal feedback, and the intervention can be tailored. These components are success factors when behavior change is aimed for and requires considerable time from both children and health care providers [67-69]. The intervention takes 7 weeks, which will not be long enough to reach lasting behavior change. However, 7 weeks of monitoring and coaching can lead to positive PA experiences and therefore, improved PA levels. Extended use of the intervention, possibly with decreasing involvement of the health care provider, can help maintain these improved PA levels and form new habits. In the evaluation study we will obtain an indication of the effect of the intervention on improved PA levels.

Time is scarce and expensive for the health care providers. Therefore, we tried to diminish the coaching time by enabling web-based coaching meetings and by minimalizing health care providers' preparation time. Despite these optimizations, it might be difficult to structurally implement the *Foxfit* intervention in the health care system. Because of the advantages of personal feedback from health care providers on children's motivation, we would not recommend to completely omit this. However, the investment time of health care providers might be further diminished through the use of artificial intelligence (AI) for automatically personalizing the intervention. Such a technical system might, for instance, suggest health care providers weekly activity goals based on activity data of the past weeks. This approach takes the attitude of both patients and health care providers toward AI into account. Patients appreciate the human factor of health care providers and might have a negative attitude toward medical AI because of the absence of this humanistic care factor [94]. Moreover, health care providers appreciate AI systems to make the process more efficient or support them by giving suggestions, but ultimately, they want to make their own choices [95]. By giving AI-based suggestions for weekly PA goals, there still is clinical expertise and personal feedback, but the health care providers are supported in their decisions and the process is made more efficient.

Although coaching from the health care provider is an important component of the intervention, it would be interesting to study whether the intervention can be used as standalone app for children as well. Although the intervention was designed for children with asthma who are inactive, there are also several children who already have moderate PA levels but who may still be able to diminish the impact of their disease when their PA levels are improved. A stand-alone app might, for instance, be helpful for them as well as for children who completed the entire blended *Foxfit* intervention previously and now need a repetition. To examine the possibilities of a standalone app, we should study whether some of the health care provider's tasks and dashboard functionalities can be implemented in the smartphone app. Moreover, in some cases, highly educated parents could possibly take over the coaching role of the health care provider when they receive proper instruction.

Regarding the hardware, the PAM that children must wear is extremely easy to use and safe in terms of personal data. No personal data are stored, and the activity data are not shared with a commercial organization. The PAM does not require charging during the intervention and it synchronizes automatically. Thus, the impact on children's everyday life is minimal. However, this system requires Raspberry Pi with Bluetooth and 4G dongle at home. This takes up space, requires an electrical outlet, and comprises of small components, which might be inconvenient when having very young children at home.

Although a large group of children participated in the needs assessment, it was challenging to include children with low PA levels for our pilot test. In the end, more children with sufficient PA levels than children with low PA levels pilot-tested our prototype than children with low PA levels. This might have influenced the feedback of the prototype and therefore the choices made for the final intervention. For example, in the needs assessment, children suggested that competitive elements would help them become physically more active, whereas participants of the stakeholder sessions indicated that competition is not desired, because children's motivation might be reduced when they compete against children having better PA possibilities. A solution might be to compete against each other based on children's personalized goals. In the feasibility and usability tests, user experiences and suggestions for improvement of a larger group of children with asthma who are inactive will be explored.

According to the Fogg Behavior Model, someone will not automatically perform the desired behavior when the opportunity, capability, and motivation are sufficient, but there will be a moment of opportunity in which someone can be persuaded to perform the desired behavior in case he receives a trigger [96,97]. In technology-supported interventions, providing such an external trigger at the moment of opportunity is called a just-in-time notification. Although such notifications

might have reinforced children's motivation, we did not choose to implement them, because one of our requirements was that the intervention should be usable without the children having to carry the smartphone with them all day. The PAM did not technically support such notifications, but incorporating these just-in-time notifications might strengthen the motivational aspect of our intervention in the future.

Conclusions

IM was useful to structure the development of the blended and technology-supported intervention *Foxfit*, and to describe this development process. *Foxfit* facilitates health care providers in promoting PA in children with asthma. The intervention consists of an activity tracker, smartphone app for children, and web-based dashboard for health care providers. Important behavior change elements of the intervention are goal setting, rewards, action planning, monitoring, shaping knowledge, a gamified story, personal coaching and feedback, and a tailored approach. Not only is the intervention based on behavior change techniques, it also meets the needs and wishes of its users because it was cocreated with children with asthma, their parents, and their health care providers.

Acknowledgments

The authors thank all the children, parents, and health care providers who participated in this study. They also thank Rutger van Teutem for his helpful input regarding the design of the intervention and Gainplay Studio for their valuable contribution to the development of the intervention. This study is part of the research project SIMBA, which is funded by the directing body SIA—a part of the Dutch Organization for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek)—under grant HW/Raak PUB 03.047.

Conflicts of Interest

None declared.

References

1. Morris PJ. Physical activity recommendations for children and adolescents with chronic disease. Current Sport Med Report 2008;7(6):353-358. [doi: 10.1249/jsr.0b013e31818f0795]
2. Riner WF, Sellhorst SH. Physical activity and exercise in children with chronic health conditions. J Sport Health Sci 2013 Mar;2(1):12-20. [doi: 10.1016/j.jshs.2012.11.005]
3. Fact Sheet Asthma. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/asthma [accessed 2021-09-14]
4. Fanelli A, Cabral A, Neder J, Martins MA, Carvalho CR. Exercise training on disease control and quality of life in asthmatic children. Med Sci Sports Exerc 2007 Sep;39(9):1474-1480. [doi: 10.1249/mss.0b013e3180d099ad] [Medline: 17805077]
5. Welsh L, Kemp JG, Roberts RG. Effects of physical conditioning on children and adolescents with asthma. Sports Med 2005;35(2):127-141. [doi: 10.2165/00007256-200535020-00003] [Medline: 15707377]
6. Eichenberger PA, Diener SN, Kofmehl R, Spengler CM. Effects of exercise training on airway hyperreactivity in asthma: a systematic review and meta-analysis. Sports Med 2013 Nov 12;43(11):1157-1170. [doi: 10.1007/s40279-013-0077-2] [Medline: 23846823]
7. Global strategy for asthma management and prevention 2021. Global Initiative for Asthma. URL: https://www.ginasthma.org [accessed 2021-09-30]
8. Ferrante G, La Grutta S. The burden of pediatric asthma. Front Pediatr 2018 Jun 22;6:186 [FREE Full text] [doi: 10.3389/fped.2018.00186] [Medline: 29988370]
9. Walsh M, Ryan-Wenger N. Sources of stress in children with asthma. J Sch Health 1992 Dec;62(10):459-463. [doi: 10.1111/j.1746-1561.1992.tb01221.x] [Medline: 1289656]
10. Glazebrook C, McPherson A, Macdonald I, Swift JA, Ramsay C, Newbould R, et al. Asthma as a barrier to children's physical activity: implications for body mass index and mental health. Pediatrics 2006 Dec;118(6):2443-2449. [doi: 10.1542/peds.2006-1846] [Medline: 17142530]
11. Lang D, Butz A, Duggan A, Serwint J. Physical activity in urban school-aged children with asthma. Pediatrics 2004 Apr;113(4):e341-e346. [doi: 10.1542/peds.113.4.e341] [Medline: 15060265]
12. Meyer A, Machnick MA, Behnke W, Braumann KM. [Participation of asthmatic children in gymnastic lessons at school]. Pneumologie 2002 Aug;56(8):486-492. [doi: 10.1055/s-2002-33314] [Medline: 12174333]
13. Yiallouros PK, Economou M, Kolokotroni O, Savva SC, Gavatha M, Ioannou P, et al. Gender differences in objectively assessed physical activity in asthmatic and non-asthmatic children. Pediatr Pulmonol 2015 Apr 28;50(4):317-326. [doi: 10.1002/ppul.23045] [Medline: 24678058]
14. Bindels PJ, Van de Griendt EJ, Grol MH. NHG-standaard astma bij kinderen (derde herziening). Huisarts Wet 2014;57(2):70-80.
15. Beggs S, Foong YC, Le HC, Noor D, Wood-Baker R, Walters JA. Swimming training for asthma in children and adolescents aged 18 years and under. Evid Based Child Health 2013 Sep 18;8(5):1514-1581. [doi: 10.1002/ebch.1935]

16. Wanrooij VH, Willeboordse M, Dompeling E, van de Kant KD. Exercise training in children with asthma: a systematic review. Br J Sports Med 2014 Jul 23;48(13):1024-1031. [doi: 10.1136/bjsports-2012-091347] [Medline: 23525551]
17. Lu K, Forno E. Exercise and lifestyle changes in pediatric asthma. Curr Opin Pulm Med 2020 Jan;26(1):103-111 [FREE Full text] [doi: 10.1097/MCP.0000000000000636] [Medline: 31652153]
18. Zhang W, Wang Q, Liu L, Yang W, Liu H. Effects of physical therapy on lung function in children with asthma: a systematic review and meta-analysis. Pediatr Res 2021 May 03;89(6):1343-1351. [doi: 10.1038/s41390-020-0874-x] [Medline: 32244247]
19. Joschtel B, Gomersall SR, Tweedy S, Petsky H, Chang AB, Trost SG. Effects of exercise training on physical and psychosocial health in children with chronic respiratory disease: a systematic review and meta-analysis. BMJ Open Sport Exerc Med 2018 Oct 01;4(1):e000409 [FREE Full text] [doi: 10.1136/bmjsem-2018-000409] [Medline: 30305925]
20. Clark NM, Valerio MA. The role of behavioural theories in educational interventions for paediatric asthma. Paediatric Respiratory Rev 2003 Dec;4(4):325-333. [doi: 10.1016/s1526-0542(03)00091-5]
21. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. J Cardiovasc Nurs 2013;28(4):320-329 [FREE Full text] [doi: 10.1097/JCN.0b013e318250a3e7] [Medline: 22635061]
22. Müller AM, Alley S, Schoeppe S, Vandelanotte C. Int J Behav Nutr Phys Act 2016 Oct 10;13(1):109 [FREE Full text] [doi: 10.1186/s12966-016-0434-2] [Medline: 27724911]
23. Michie S, Abraham C. Advancing the science of behaviour change: a plea for scientific reporting. Addiction 2008 Sep;103(9):1409-1410. [doi: 10.1111/j.1360-0443.2008.02291.x] [Medline: 18783495]
24. Vandelanotte C, Müller AM, Short CE, Hingle M, Nathan N, Williams SL, et al. Past, present, and future of ehealth and mHealth research to improve physical activity and dietary behaviors. J Nutr Educ Behav 2016 Mar;48(3):219-228. [doi: 10.1016/j.jneb.2015.12.006] [Medline: 26965100]
25. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
26. Lewis GN, Rosie JA. Virtual reality games for movement rehabilitation in neurological conditions: how do we meet the needs and expectations of the users? Disabil Rehabil 2012 Apr 05;34(22):1880-1886. [doi: 10.3109/09638288.2012.670036] [Medline: 22480353]
27. Janssen J, Verschuren O, Renger W, Ermers J, Ketelaar M, van Ee R. Gamification in physical therapy: more than using games. Pediatr Phys Ther 2017 Jan;29(1):95-99. [doi: 10.1097/PEP.0000000000000326] [Medline: 27984481]
28. Patel MS, Small DS, Harrison JD, Fortunato MP, Oon AL, Rareshide CA, et al. Effectiveness of behaviorally designed gamification interventions with social incentives for increasing physical activity among overweight and obese adults across the United States: the STEP UP randomized clinical trial. JAMA Intern Med 2019 Dec 01;179(12):1624-1632 [FREE Full text] [doi: 10.1001/jamainternmed.2019.3505] [Medline: 31498375]
29. Gal R, May AM, van Overmeeren EJ, Simons M, Monninkhof EM. The effect of physical activity interventions comprising wearables and smartphone applications on physical activity: a systematic review and meta-analysis. Sports Med Open 2018 Sep 03;4(1):42 [FREE Full text] [doi: 10.1186/s40798-018-0157-9] [Medline: 30178072]
30. Lau PW, Lau EY, Wong DP, Ransdell L. A systematic review of information and communication technology-based interventions for promoting physical activity behavior change in children and adolescents. J Med Internet Res 2011 Jul 13;13(3):e48 [FREE Full text] [doi: 10.2196/jmir.1533] [Medline: 21749967]
31. Lee W, Reyes-Fernández MC, Posada-Gómez R, Juárez-Martínez U, Martínez-Sibaja A, Alor-Hernández G. Using health games for rehabilitation of patients with infantile cerebral palsy. J Phys Ther Sci 2016 Aug;28(8):2293-2298 [FREE Full text] [doi: 10.1589/jpts.28.2293] [Medline: 27630417]
32. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. J Health Commun 2012;17 Suppl 1:11-21. [doi: 10.1080/10810730.2011.649103] [Medline: 22548594]
33. Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S. IDEAS (integrate, design, assess, and share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. J Med Internet Res 2016 Dec 16;18(12):e317 [FREE Full text] [doi: 10.2196/jmir.5927] [Medline: 27986647]
34. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337(sep29 1):a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
35. Mehra S, Visser B, Dadema T, van den Helder J, Engelbert RH, Weijs PJ, et al. Translating behavior change principles into a blended exercise intervention for older adults: design study. JMIR Res Protoc 2018 May 02;7(5):e117 [FREE Full text] [doi: 10.2196/resprot.9244] [Medline: 29720358]
36. Sporrel K, De Boer RD, Wang S, Nibbeling N, Simons M, Deutekom M, et al. The design and development of a personalized leisure time physical activity application based on behavior change theories, end-user perceptions, and principles from empirical data mining. Front Public Health 2020;8:528472 [FREE Full text] [doi: 10.3389/fpubh.2020.528472] [Medline: 33604321]

37. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
38. Lambert JD, Greaves CJ, Farrand P, Haase AM, Taylor AH. Development of a web-based intervention (eMotion) based on behavioural activation to promote physical activity in people with depression. Mental Health Physical Activity 2017 Oct;13:120-136. [doi: 10.1016/j.mhpa.2017.10.003]
39. Bartholomew LK, Markham CM, Ruiter RAC, Fernández ME, Kok G, Parcel GS. Planning Health Promotion Programs An Intervention Mapping Approach. Hoboken, New Jersey, United States: Wiley; 2016.
40. Oosterom-Calo R, Te Velde SJ, Stut W, Brug J. Development of Motivate4Change using the intervention mapping protocol: an interactive technology physical activity and medication adherence promotion program for hospitalized heart failure patients. JMIR Res Protoc 2015 Jul 20;4(3):e88 [FREE Full text] [doi: 10.2196/resprot.4282] [Medline: 26195072]
41. Lloyd JJ, Logan S, Greaves CJ, Wyatt KM. Evidence, theory and context--using intervention mapping to develop a school-based intervention to prevent obesity in children. Int J Behav Nutr Phys Act 2011 Jul 13;8(1):73 [FREE Full text] [doi: 10.1186/1479-5868-8-73] [Medline: 21752261]
42. Gray-Burrows KA, Day PF, Marshman Z, Aliakbari E, Prady SL, McEachan RR. Using intervention mapping to develop a home-based parental-supervised toothbrushing intervention for young children. Implement Sci 2016 May 06;11(1):61 [FREE Full text] [doi: 10.1186/s13012-016-0416-4] [Medline: 27153832]
43. Pérez-Rodrigo C, Wind M, Hildonen C, Bjelland M, Aranceta J, Klepp K, et al. The pro children intervention: applying the intervention mapping protocol to develop a school-based fruit and vegetable promotion programme. Ann Nutr Metab 2005 Aug 26;49(4):267-277. [doi: 10.1159/000087249] [Medline: 16088090]
44. Brons A, Braam K, Timmerman A, Broekema A, Visser B, van Ewijk B, et al. Promoting factors for physical activity in children with asthma explored through concept mapping. Int J Environ Res Public Health 2019 Nov 13;16(22):4467 [FREE Full text] [doi: 10.3390/ijerph16224467] [Medline: 31766299]
45. van der Ploeg HP, van der Beek AJ, van der Woude LH, van Mechelen W. Physical activity for people with a disability: a conceptual model. Sports Med 2004;34(10):639-649. [doi: 10.2165/00007256-200434100-00002] [Medline: 15335241]
46. Velsor-Friedrich B, Vlasses F, Moberley J, Coover L. Talking with teens about asthma management. J Sch Nurs 2004 Jun 15;20(3):140-148. [doi: 10.1177/10598405040200030401] [Medline: 15147228]
47. Williams B, Powell A, Hoskins G, Neville R. Exploring and explaining low participation in physical activity among children and young people with asthma: a review. BMC Fam Pract 2008 Jun 30;9(1):40 [FREE Full text] [doi: 10.1186/1471-2296-9-40] [Medline: 18590558]
48. Still L, Dolen WK. The perception of asthma severity in children. Curr Allergy Asthma Rep 2016 Jul 22;16(7):50. [doi: 10.1007/s11882-016-0629-2] [Medline: 27333779]
49. Kornblit A, Cain A, Bauman LJ, Brown NM, Reznik M. Parental perspectives of barriers to physical activity in urban schoolchildren with asthma. Acad Pediatr 2018 Apr;18(3):310-316 [FREE Full text] [doi: 10.1016/j.acap.2017.12.011] [Medline: 29309846]
50. Williams B, Hoskins G, Pow J, Neville R, Mukhopadhyay S, Coyle J. Low exercise among children with asthma: a culture of over protection? A qualitative study of experiences and beliefs. Br J Gen Pract 2010 Aug 01;60(577):e319-e326. [doi: 10.3399/bjgp10x515070]
51. Ajzen I. Attitudes, Personality and Behaviour. New York, United States: McGraw-Hill Education; 2005.
52. Stenberg U, Haaland-Øverby M, Koricho AT, Trollvik A, Kristoffersen LR, Dybvig S, et al. How can we support children, adolescents and young adults in managing chronic health challenges? A scoping review on the effects of patient education interventions. Health Expect 2019 Oct;22(5):849-862 [FREE Full text] [doi: 10.1111/hex.12906] [Medline: 31131527]
53. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
54. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011 Presented at: MindTrek '11: Academic MindTrek 2011; Sep 28 - 30, 2011; Tampere Finland. [doi: 10.1145/2181037.2181040]
55. 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]
56. Munson S, Poole E, Perry D, Peyton T. Gamification and health. In: The Gameful World: Approaches, Issues, Applications. Cambridge UK: MIT Press; 2015.
57. Cheek C, Fleming T, Lucassen MF, Bridgman H, Stasiak K, Shepherd M, et al. Integrating health behavior theory and design elements in serious games. JMIR Ment Health 2015 Apr 21;2(2):e11 [FREE Full text] [doi: 10.2196/mental.4133] [Medline: 26543916]
58. King D, Greaves F, Exeter C, Darzi A. 'Gamification': influencing health behaviours with games. J R Soc Med 2013 Mar;106(3):76-78 [FREE Full text] [doi: 10.1177/0141076813480996] [Medline: 23481424]
59. Morschheuser B, Rivera-Pelayo V, Mazarakis A. Interaction and reflection with quantified self and gamification: an experimental study. J Literacy Technol 2014;15(2).

60. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. J Biomed Inform 2017 Jul;71:31-48 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.011] [Medline: 28536062]
61. Bossen D, Broekema A, Visser B, Brons A, Timmerman A, van Etten-Jamaludin F, et al. Effectiveness of serious games to increase physical activity in children with a chronic disease: systematic review with meta-analysis. J Med Internet Res 2020 Apr 01;22(4):e14549 [FREE Full text] [doi: 10.2196/14549] [Medline: 32234697]
62. Judah G, Gardner B, Aunger R. Forming a flossing habit: an exploratory study of the psychological determinants of habit formation. Br J Health Psychol 2013 May;18(2):338-353. [doi: 10.1111/j.2044-8287.2012.02086.x] [Medline: 22989272]
63. Judah G, Gardner B, Kenward MG, DeStavola B, Aunger R. Exploratory study of the impact of perceived reward on habit formation. BMC Psychol 2018 Dec 20;6(1):62 [FREE Full text] [doi: 10.1186/s40359-018-0270-z] [Medline: 30572936]
64. Mekler ED, Brühlmann F, Tuch AN, Opwis K. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. Comput Human Behav 2017 Jun;71:525-534. [doi: 10.1016/j.chb.2015.08.048]
65. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
66. Rhodes RE, Naylor P, McKay HA. Pilot study of a family physical activity planning intervention among parents and their children. J Behav Med 2010 Apr 24;33(2):91-100. [doi: 10.1007/s10865-009-9237-0] [Medline: 19937106]
67. Asbjørnsen RA, Smedsrød ML, Solberg Nes L, Wentzel J, Varsi C, Hjelmesæth J, et al. Persuasive system design principles and behavior change techniques to stimulate motivation and adherence in electronic health interventions to support weight loss maintenance: scoping review. J Med Internet Res 2019 Jun 21;21(6):e14265 [FREE Full text] [doi: 10.2196/14265] [Medline: 31228174]
68. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
69. Orji R, Moffatt K. Persuasive technology for health and wellness: state-of-the-art and emerging trends. Health Informatics J 2018 Mar;24(1):66-91 [FREE Full text] [doi: 10.1177/1460458216650979] [Medline: 27245673]
70. Beckwith VZ, Beckwith J. Motivational interviewing: a communication tool to promote positive behavior change and optimal health outcomes. NASN School Nurse 2020 Apr 26;35(6):344-351. [doi: 10.1177/1942602x20915715]
71. Söderlund LL, Nordqvist C, Angbratt M, Nilsen P. Applying motivational interviewing to counselling overweight and obese children. Health Educ Res 2009 Jun 07;24(3):442-449. [doi: 10.1093/her/cyn039] [Medline: 18996887]
72. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013 Apr 18;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
73. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51(3-4):214-221 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]
74. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JE. Key components in eHealth interventions combining self-tracking and persuasive eCoaching to promote a healthier lifestyle: a scoping review. J Med Internet Res 2017 Aug 01;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
75. Wise M, Gustafson DH, Sorkness CA, Molfenter T, Staresinic A, Meis T, et al. Internet telehealth for pediatric asthma case management: integrating computerized and case manager features for tailoring a web-based asthma education program. Health Promot Pract 2007 Jul;8(3):282-291 [FREE Full text] [doi: 10.1177/1524839906289983] [Medline: 16928987]
76. King G, Lawm M, King S, Rosenbaum P, Kertoy MK, Young NL. A conceptual model of the factors affecting the recreation and leisure participation of children with disabilities. Physical Occup Ther Pediatrics 2009 Jul 29;23(1):63-90. [doi: 10.1080/j006v23n01_05]
77. Meijer SA, Sinnema G, Bijstra JO, Mellenbergh GJ, Wolters WH. Social functioning in children with a chronic illness. J Child Psychol Psychiatry 2003 Oct 13;41(3):309-317. [doi: 10.1111/1469-7610.00615]
78. van Brussel M, van der Net J, Hulzebos E, Helders PJ, Takken T. The Utrecht approach to exercise in chronic childhood conditions: the decade in review. Pediatr Phys Ther 2011;23(1):2-14. [doi: 10.1097/PEP.0b013e318208cb22] [Medline: 21304338]
79. Pam homepage. Pam Coach. URL: https://www.pamcoach.com/index.php?pid=1 [accessed 2021-09-30]
80. Slootmaker SM, Chinapaw MJ, Schuit AJ, Seidell JC, Van Mechelen W. Feasibility and effectiveness of online physical activity advice based on a personal activity monitor: randomized controlled trial. J Med Internet Res 2009 Jul 29;11(3):e27 [FREE Full text] [doi: 10.2196/jmir.1139] [Medline: 19674956]
81. Pol MC, Ter Riet G, van Hartingsveldt M, Kröse B, de Rooij SE, Buurman BM. Effectiveness of sensor monitoring in an occupational therapy rehabilitation program for older individuals after hip fracture, the SO-HIP trial: study protocol of a three-arm stepped wedge cluster randomized trial. BMC Health Serv Res 2017 Jan 03;17(1):3 [FREE Full text] [doi: 10.1186/s12913-016-1934-0] [Medline: 28049480]
82. Slootmaker S, Chin A Paw MJ, Schuit AJ, van Mechelen W, Koppes LL. Concurrent validity of the PAM accelerometer relative to the MTI Actigraph using oxygen consumption as a reference. Scand J Med Sci Sports 2009 Feb;19(1):36-43. [doi: 10.1111/j.1600-0838.2007.00740.x] [Medline: 18266793]

83. Vooijs M, Alpay LL, Snoeck-Stroband JB, Beerthuizen T, Siemonsma PC, Abbink JJ, et al. Validity and usability of low-cost accelerometers for internet-based self-monitoring of physical activity in patients with chronic obstructive pulmonary disease. Interact J Med Res 2014 Oct 27;3(4):e14 [FREE Full text] [doi: 10.2196/ijmr.3056] [Medline: 25347989]
84. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. Boca Raton, Florida, United States: CRC Press; 1996.
85. Brooke J. SUS: a retrospective. J User Exp 2013;8(2):29-40.
86. Juniper E, O'Byrne PM, Guyatt G, Ferrie P, King D. Development and validation of a questionnaire to measure asthma control. Eur Respir J 1999 Oct;14(4):902-907 [FREE Full text] [doi: 10.1034/j.1399-3003.1999.14d29.x] [Medline: 10573240]
87. Juniper EF, Gruffydd-Jones K, Ward S, Svensson K. Asthma control questionnaire in children: validation, measurement properties, interpretation. Eur Respir J 2010 Dec 07;36(6):1410-1416 [FREE Full text] [doi: 10.1183/09031936.00117509] [Medline: 20530041]
88. Chinapaw MJ, Slootmaker SM, Schuit AJ, van Zuidam M, van Mechelen W. Reliability and validity of the activity questionnaire for adults and adolescents (AQuAA). BMC Med Res Methodol 2009 Aug 10;9(1):58 [FREE Full text] [doi: 10.1186/1471-2288-9-58] [Medline: 19664254]
89. Kendzierski D, DeCarlo K. Physical activity enjoyment scale: two validation studies. J Sport Exercise Psychol 1991;13(1):50-64. [doi: 10.1123/jsep.13.1.50]
90. Motl RW, Dishman RK, Saunders R, Dowda M, Felton G, Pate RR. Measuring enjoyment of physical activity in adolescent girls. Am J Prev Med 2001 Aug;21(2):110-117. [doi: 10.1016/s0749-3797(01)00326-9]
91. Veerman J, Straathof M, Treffers P, van den Bergh BRH, ten Brink LT. Handleiding Competentiebelevingsschaal voor Kinderen (CBSK). Lisse: Swets & Zeitlinger; 1997.
92. Michie S, Prestwich A. Are interventions theory-based? Development of a theory coding scheme. Health Psychol 2010 Jan;29(1):1-8. [doi: 10.1037/a0016939] [Medline: 20063930]
93. Painter J, Borba C, Hynes M, Mays D, Glanz K. The use of theory in health behavior research from 2000 to 2005: a systematic review. Ann Behav Med 2008 Jun;35(3):358-362. [doi: 10.1007/s12160-008-9042-y] [Medline: 18633685]
94. Gao S, He L, Chen Y, Li D, Lai K. Public perception of artificial intelligence in medical care: content analysis of social media. J Med Internet Res 2020 Jul 13;22(7):e16649 [FREE Full text] [doi: 10.2196/16649] [Medline: 32673231]
95. Abdullah R, Fakieh B. Health care employees' perceptions of the use of artificial intelligence applications: survey study. J Med Internet Res 2020 May 14;22(5):e17620 [FREE Full text] [doi: 10.2196/17620] [Medline: 32406857]
96. Fogg B. A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology. 2009 Presented at: Persuasive 2009: Persuasive 2009; 4th International Conference on Persuasive Technology; Apr 26 - 29, 2009; Claremont California USA. [doi: 10.1145/1541948.1541999]
97. Fogg BJ. Persuasive technology: using computers to change what we think and do. Ubiquity 2002 Dec 01;2002(December):2. [doi: 10.1145/764008.763957]

Abbreviations

AI: artificial intelligence
CeHReS: Centre for eHealth Research
IDEAS: Integrate, Design, Assess, and Share
IM: intervention mapping
MET: metabolic equivalent of task
MRC: Medical Research Council
PA: physical activity
PAM: physical activity monitor
SSH: Secure Shell

Edited by G Eysenbach; submitted 11.10.21; peer-reviewed by L Happe, M van der Kamp; comments to author 27.11.21; revised version received 18.04.22; accepted 20.04.22; published 25.07.22.

Please cite as:
Brons A, Braam K, Broekema A, Timmerman A, Millenaar K, Engelbert R, Kröse B, Visser B
Translating Promoting Factors and Behavior Change Principles Into a Blended and Technology-Supported Intervention to Stimulate Physical Activity in Children With Asthma (Foxfit): Design Study
JMIR Form Res 2022;6(7):e34121
URL: https://formative.jmir.org/2022/7/e34121
doi:10.2196/34121
PMID:35877162

Short Paper

The Association of Medical Preoperative Evaluation Using Clinical Video Telehealth With Hospital Length of Stay: Descriptive Analysis

Brittany Nicole Burton[1], MD, MHS, MAS; Sara Arastoo[1], MD; Simon Wu[2], MD; Nancy Liu[2], MSN; Michael K Ong[2], MD, PhD; Sondra Vazirani[2], MD, MPH

[1]Department of Anesthesiology and Perioperative Medicine, University of California Los Angeles Health, Los Angeles, CA, United States

[2]Department of Medicine, Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, CA, United States

Corresponding Author:
Brittany Nicole Burton, MD, MHS, MAS
Department of Anesthesiology and Perioperative Medicine
University of California Los Angeles Health
757 Westwood Plaza
Suite 3325
Los Angeles, CA, 90095
United States
Phone: 1 310 825 8325
Email: BBurton@mednet.ucla.edu

Abstract

Background: Preoperative medical evaluation serves to identify risk factors and optimize patients before surgery. Providing a telehealth option in the perioperative setting has played a significant role in reducing barriers to quality perioperative health care.

Objective: We aimed to evaluate how telemedicine preoperative evaluations using Clinical Video Telehealth (CVT) impact hospital length of stay.

Methods: We performed a retrospective chart review between 2016 and 2017 of adult patients who underwent evaluations in our hospitalist-run preoperative medicine clinic. Patients seen in our preoperative CVT program were compared to patients seen in person to evaluate the association of visit type (preoperative CVT versus in-person evaluation) with hospital length of stay, defined as hospital stay from postoperative day 0 to discharge. There were 62 patients included in this retrospective study.

Results: The adjusted incidence rate ratio (IRR) for hospital length of stay was significantly shorter in patients who underwent preoperative CVT compared to an in-person visit (IRR 0.52, 95% CI 0.29-0.92, *P*=.02).

Conclusions: After adjusting for age and comorbidities, we show that preoperative telemedicine in the perioperative setting is associated with a shorter hospital length of stay compared to in-person visits. This suggests that telemedicine can play a viable role in this clinical setting.

(JMIR Form Res 2022;6(7):e38054) doi:10.2196/38054

KEYWORDS

telemedicine; telehealth; eHealth; digital health; hospital; length of stay; veteran's health; video; veteran; preoperative; outpatient; chart review; retrospective; clinical care; effectiveness; efficacy; discharge

Introduction

For the almost 50 million surgeries and procedures performed annually in the United States, preoperative medical evaluation serves to identify and optimize perioperative risk to decrease adverse outcomes and to prevent same-day cancellations of surgery [1]. Traditionally, preoperative evaluations are performed face-to-face in the clinic. Starting in July 2014, given its large catchment area, the Veterans Affairs Greater Los Angeles Healthcare System implemented a telemedicine preoperative medicine clinic using Clinical Video Telehealth (CVT). CVT is a technology that Veterans Affairs (VA) providers have used since the early 2000s. With CVT, clinicians can gather relevant history and conduct a limited physical exam using a camera and digital stethoscope. Since adoption, CVT has found increasing rates of use, especially for patients living

in rural areas, who face significant barriers to completing in-person visits. In the wake of the COVID-19 pandemic, the importance of providing a telehealth option in outpatient care has become even more apparent [2].

Several studies in non-VA settings have demonstrated that preoperative evaluations done via telemedicine are associated with high patient/provider satisfaction, cost savings, and a lower rate of same-day cancellation when compared to in-person evaluations [3-6]. However, the potential limitations of telemedicine preoperative evaluation (eg, not performing a comprehensive physical exam may preclude clinical diagnoses) may lead to subsequent case-cancellation complications. Thus, our project aimed to evaluate how telemedicine preoperative evaluations using CVT impact hospital length of stay.

Methods

Overview

This manuscript follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement guidelines for reporting observational studies. All data used in this study were extracted from electronic medical records.

We performed a retrospective chart review of adult patients who underwent evaluations in our hospitalist-run preoperative medicine clinic. Patients seen in our preoperative CVT program, which started in July 2014, were compared to patients who had in-person visits to evaluate the association of visit type (preoperative CVT versus in-person) with hospital length of stay, defined as hospital stay from postoperative day 0 to discharge. We extracted data from 2016 to 2017. Preoperative CVT involves a thorough history and a full airway exam. Exclusion criteria for the CVT preoperative program were defined at the program's start as American Society of Anesthesiologists (ASA) class 4, or ASA class 3 and uncontrolled blood pressure (>180/100 mm Hg) and/or diabetes (glycated hemoglobin [HbA_{1c}] >9%). The patients needed to meet all criteria to be recommended for an in-person visit and therefore be excluded from CVT. These patients were recommended for in-person evaluation due to comorbidity burden and importance of taking a complete history and conducting a physical exam.

Data Analysis

Statistical analysis was performed using R (version 3.6.1; R Foundation for Statistical Computing). To measure the differences in hospital length of stay among those who received CVT versus face-to-face consultation, chi-square and student *t* tests were used. Multivariable negative binomial regressions were performed, adjusting for age, gender, ASA score, surgery type (major or minor), and Elixhauser comorbidity index. The incidence rate ratio (IRR), 95% CIs, and *P* value were calculated for each estimate.

Ethical Considerations

Our study was reviewed by the Institutional Review Board of the West Los Angeles Veterans Administration Medical Center and was granted an "exempt" status.

Results

There were 62 patients included in this retrospective study. The cancellation rate was 1.74% for CVT versus 3.48% for in-person. Table 1 outlines the distribution of patient characteristics stratified by preoperative visit type. In this unadjusted analysis, there were no significant differences between the cohorts.

Table 2 outlines the negative binomial regression for the association of visit type with hospital length of stay. The age- and Elixhauser score–adjusted incidence rate for hospital length of stay was significantly shorter in patients who underwent preoperative CVT compared to an in-person visit (IRR 0.52, 95% CI 0.29-0.92, *P*=.02).

Table 1. Participant characteristics.

Characteristics	In-person (n=29)	Preoperative Clinical Video Telehealth (n=33)	*P* value[a]
Age (years), mean (SD)	62.83 (11.23)	59.36 (15.43)	.32
Gender, n (%)			.26
Male	28 (96.6)	28 (84.8)	
Female	1 (3.4)	5 (15.2)	
Elixhauser comorbidity score, n (%)			.16
≤1	5 (17.2)	11 (33.3)	
0	6 (20.7)	9 (27.3)	
≥1 and <5	8 (27.6)	10 (30.3)	
≥6 and <10	4 (13.8)	1 (3)	
≥11 and <19	6 (20.7)	2 (6.1)	
Surgical specialty, n (%)			.68
Urology	7 (24.1)	6 (18.2)	
Colorectal	1 (3.4)	2 (6.1)	
Ophthalmology	1 (3.4)	0 (0)	
Plastic surgery	1 (3.4)	5 (15.2)	
General	6 (20.7)	3 (9.1)	
Orthopedics	6 (20.7)	6 (18.2)	
Gynecology	0 (0)	1 (3)	
Ear, nose, and throat	2 (6.9)	4 (12.1)	
Neurosurgery	4 (13.8)	5 (15.2)	
Vascular	1 (3.4)	1 (3)	
ASA class, n (%)[b]			.13
1	1 (3.4)	2 (6.1)	
2	0 (0)	5 (15.2)	
3	27 (93.1)	24 (72.7)	
4	1 (3.4)	2 (6.1)	
Length of stay, mean (SD)	6.55 (9.09)	3.33 (3.97)	.07

[a]Pearson chi-square test for categorical variables. Student *t* test for continuous variables.

[b]ASA: American Society of Anesthesiologists.

Table 2. The association of preoperative visit with hospital length of stay.

	Incidence rate ratio	95% CI	*P* value
Preoperative Clinical Video Telehealth[a]	0.52	0.29-0.92	.02
Elixhauser comorbidity	1.00	0.82-1.22	.27
Age	1.01	0.98-1.03	.98

[a]Reference group for preoperative Clinical Video Telehealth is patients who received medical chart review and did not receive preoperative Clinical Video Telehealth.

Discussion

In summary, we show that preoperative CVT, while holding age and the Elixhauser comorbidity score constant in the model, has an IRR for hospital length of stay that is 0.52 times lower compared to in-person visits. This study found a significant difference in the IRR of postoperative length of stay between patients receiving telehealth versus in-person preoperative evaluations. This suggests that telemedicine can play a viable role in this clinical setting. Telemedicine has the potential to increase care access across all specialties and health care systems. Our findings had several limitations including that the

study was retrospective, was conducted at a single center, and had a low sample size, leading to an increased risk of type II error. Length of stay may be affected by many factors. In the VA patient population, social reasons may affect length of stay more than the typical patient population. There likely is selection bias between those patients who were willing to do CVT versus those who wanted an in-person evaluation.

In our patient population, several patients were more interested in telemedicine compared to in-person visits and we hope to expand to other locations. We plan to apply biomedical informatics to the electronic medical records to extract granular patient data including but not limited to (1) demographic data (age, race, socioeconomic status, and zip code), (2) comorbidities and severity of each comorbidity, (3) postoperative complications, (4) telemedicine-specific data (cancellation rates, missed appointments, and scheduling delays), and (5) patient perceptions and experiences. We hope this research design will help us to identify the benefits and potential disadvantages of telemedicine in the perioperative period. Future studies should be prospective and adequately powered to limit type II error. In addition, future studies should explore how to appropriately triage patients as being "telehealth-appropriate" in the preoperative setting, as well as investigate the effects of preoperative telehealth on other patient-centered outcomes.

Conflicts of Interest

MKO has received royalties from UpToDate Inc. The other authors declare no conflicts of interest.

References

1. Hall MJ, Schwartzman A, Zhang J, Liu X. Ambulatory surgery data from hospitals and ambulatory surgery centers: United States, 2010. Natl Health Stat Report 2017 Mar(102):1-15 [FREE Full text] [Medline: 28256998]
2. Adams SV, Mader MJ, Bollinger MJ, Wong ES, Hudson TJ, Littman AJ. Utilization of interactive clinical video telemedicine by rural and urban veterans in the Veterans Health Administration health care system. J Rural Health 2019 Jun;35(3):308-318. [doi: 10.1111/jrh.12343] [Medline: 30600557]
3. Applegate RL, Gildea B, Patchin R, Rook JL, Wolford B, Nyirady J, et al. Telemedicine pre-anesthesia evaluation: a randomized pilot trial. Telemed J E Health 2013 Mar;19(3):211-216. [doi: 10.1089/tmj.2012.0132] [Medline: 23384334]
4. Kamdar NV, Huverserian A, Jalilian L, Thi W, Duval V, Beck L, et al. Development, implementation, and evaluation of a telemedicine preoperative evaluation initiative at a major academic medical center. Anesth Analg 2020 Dec;131(6):1647-1656 [FREE Full text] [doi: 10.1213/ANE.0000000000005208] [Medline: 32841990]
5. Mullen-Fortino M, Rising KL, Duckworth J, Gwynn V, Sites FD, Hollander JE. Presurgical assessment using telemedicine technology: impact on efficiency, effectiveness, and patient experience of care. Telemed J E Health 2019 Feb;25(2):137-142. [doi: 10.1089/tmj.2017.0133] [Medline: 30048210]
6. Tam A, Leung A, O'Callaghan C, Fagermo N. Role of telehealth in perioperative medicine for regional and rural patients in Queensland. Intern Med J 2017 Aug 06;47(8):933-937. [doi: 10.1111/imj.13484] [Medline: 28485821]

Abbreviations

ASA: American Society of Anesthesiologists
CVT: Clinical Video Telehealth
IRR: incidence rate ratio
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
VA: Veterans Affairs

Edited by A Mavragani; submitted 17.03.22; peer-reviewed by H Karim, B Zhao; comments to author 05.05.22; revised version received 21.06.22; accepted 27.06.22; published 25.07.22.

Please cite as:
Burton BN, Arastoo S, Wu S, Liu N, Ong MK, Vazirani S
The Association of Medical Preoperative Evaluation Using Clinical Video Telehealth With Hospital Length of Stay: Descriptive Analysis
JMIR Form Res 2022;6(7):e38054
URL: https://formative.jmir.org/2022/7/e38054
doi:10.2196/38054
PMID:35877170

Original Paper

Single-Session Interventions Embedded Within Tumblr: Acceptability, Feasibility, and Utility Study

Mallory L Dobias[1], MA; Robert R Morris[2], PhD; Jessica L Schleider[1], PhD

[1]Department of Psychology, Stony Brook University, Stony Brook, NY, United States

[2]Koko, San Francisco, CA, United States

Corresponding Author:
Mallory L Dobias, MA
Department of Psychology
Stony Brook University
Psychology Building B-363
Stony Brook, NY, 11794-2500
United States
Phone: 1 512 818 9040
Email: mallory.dobias@stonybrook.edu

Abstract

Background: Existing mental health treatments are insufficient for addressing mental health needs at scale, particularly for teenagers, who now seek mental health information and support on the web. Single-session interventions (SSIs) may be particularly well suited for dissemination as embedded web-based support options that are easily accessible on popular social platforms.

Objective: We aimed to evaluate the acceptability and effectiveness of three SSIs, each with a duration of 5 to 8 minutes (Project Action Brings Change, Project Stop Adolescent Violence Everywhere, and REFRAME)—embedded as Koko *minicourses* on Tumblr—to improve three key mental health outcomes: hopelessness, self-hate, and the desire to stop self-harm behavior.

Methods: We used quantitative data (ie, star ratings and SSI completion rates) to evaluate acceptability and short-term utility of all 3 SSIs. Paired 2-tailed *t* tests were used to assess changes in hopelessness, self-hate, and the desire to stop future self-harm from before to after the SSI. Where demographic information was available, the analyses were restricted to teenagers (13-19 years). Examples of positive and negative qualitative user feedback (ie, written text responses) were provided for each program.

Results: The SSIs were completed 6179 times between March 2021 and February 2022. All 3 SSIs generated high star ratings (>4 out of 5 stars), with high completion rates (approximately 25%-57%) relative to real-world completion rates among other digital self-help interventions. Paired 2-tailed *t* tests detected significant pre-post reductions in hopelessness for those who completed Project Action Brings Change (P<.001, Cohen d_z=−0.81, 95% CI −0.85 to −0.77) and REFRAME (P<.001, Cohen d_z=−0.88, 95% CI −0.96 to −0.80). Self-hate significantly decreased (P<.001, Cohen d_z=−0.67, 95% CI −0.74 to −0.60), and the desire to stop self-harm significantly increased (P<.001, Cohen d_z=0.40, 95% CI 0.33 to 0.47]) from before to after the completion of Project Stop Adolescent Violence Everywhere. The results remained consistent across sensitivity analyses and after correcting for multiple tests. Examples of positive and negative qualitative user feedback point toward future directions for SSI research.

Conclusions: Very brief SSIs, when embedded within popular social platforms, are one promising and acceptable method for providing free, scalable, and potentially helpful mental health support on the web. Considering the unique barriers to mental health treatment access that many teenagers face, this approach may be especially useful for teenagers without access to other mental health supports.

(JMIR Form Res 2022;6(7):e39004) doi:10.2196/39004

KEYWORDS

web-based intervention; internet intervention; digital intervention; single-session intervention; mental health

Introduction

Limited Accessibility of Existing Mental Health Treatments

Existing mental health treatments have long been inaccessible due to well-established structural (eg, cost, transportation, and time) and individual (eg, stigma and distrust of providers) barriers [1-3]. Furthermore, people experiencing suicidal thoughts or engaging in self-harm often fear the negative consequences of disclosing these experiences (eg, involuntary hospitalization) [4,5], creating additional barriers to treatment among those who may need it most. This is particularly true for teenagers whose caregivers often serve as gatekeepers for their own mental health care [6-8] and who often fear the involvement of a caregiver without their consent [5].

To fill the gap between the *need* for support and *access* to it, many teenagers seek and receive mental health support on the web [9,10]. Specifically, a majority (61%-85%) of teens and young adults have sought information or help for their mental health on the web [10-12]. Teenagers report using web-based mental health resources as a way to obtain free, easily accessible, or anonymous support [12,13], and web-based communities via social platforms (eg, forums and discussion boards) offer a space to share one's experiences with others who have faced similar challenges [14-17]. Youth facing greater psychological distress are also more likely to search for and discuss topics related to their mental health on the web [15]. Thus, these web-based communities in social platforms represent a unique, underexplored avenue for reaching teenagers in need of mental health support.

Offering Mental Health Support via Web-Based Social Platforms

Initial efforts suggest that it is *possible* to provide mental health supports that are integrated into existing web-based social platforms. Koko, a nonprofit, web-based mental health platform, provides on-demand peer support [18,19], crisis triage [20], and self-guided interventions on topics such as body image, self-harm, and stress management. People can access Koko services on various messenger apps (eg, Telegram, Facebook Messenger, and Kik) or via direct message channels of certain social networking platforms (eg, Tumblr); in recent studies, a portion of Koko users were directly referred from a social media integration [20]. However, no previous evaluation of Koko's services has focused entirely on investigating the acceptability and effectiveness of embedding mental health supports within a social networking platform.

Single-Session Interventions as Embedded Web-Based Support

Single-session interventions (SSIs)—brief, targeted interventions designed to be completed within 1 sitting or clinical interaction—provide an especially promising intervention format for reaching young people on the web [21]. Multiple large-scale studies have indicated that SSIs can create meaningful changes for key mental health outcomes. In a recent randomized controlled trial of adolescents (N=2452), a 20- to 30-minute web-based behavioral activation SSI (Project Action Brings Change [ABC]) outperformed a supportive therapy control program at reducing postintervention hopelessness and depressive symptoms 3 months later [22]. Another trial (N=565) found Project Stop Adolescent Violence Everywhere (SAVE), a web-based 30-minute SSI targeting self-injurious thoughts and behaviors in teenagers, improved postintervention self-hatred, and desires to stop future self-harm, relative to an active control [23]. Taken together, substantive evidence suggests the efficacy of web-based SSIs in improving mental health outcomes among young people.

Despite the great *potential* of SSIs to increase access to mental health supports, little large-scale research has evaluated their acceptability and effectiveness when disseminated in web-based settings that young people already routinely access. One study evaluated the acceptability and preliminary utility of a free, open-access platform offering 3 web-based SSIs for youth [24]. After viewing paid advertisements on a social media platform (ie, Instagram), nearly 700 youths clicked on the advertisement and viewed the web-based platform within 6 months. The youths rated all 3 SSIs as acceptable, and analyses indicated significant pre- to postintervention reductions in hopelessness and self-hate among 187 SSI completers. However, to complete an SSI program, the youth had to be (1) aware of the open-access platform and (2) motivated to visit the platform on a separate, unfamiliar website and finish a 30-minute activity. Reducing friction (ie, points of difficulty) between individuals and their technology during key help-seeking thresholds may be particularly important for encouraging greater access to mental health supports on the web [25]. Rather than relying on individuals' ability and motivation to seek external mental health resources in moments of heightened distress (eg, standalone web-based resources, crisis hotlines, and text lines), SSIs could provide accessible, in-the-moment anonymous supports for young people who are *already embedded* in popular social platforms.

Considerations for Intervention Design

SSIs designed for these web-based, embedded contexts should be streamlined to improve engagement while retaining their therapeutic value. Once disseminated in the real world, digital interventions often encounter issues with user engagement (ie, low uptake and low completion) [26-28]. Substantially reducing the time and effort required to complete web-based SSIs (eg, condensing content and reducing the number of writing prompts) may improve user engagement, particularly for SSIs delivered as in-the-moment supports at times of elevated distress. Determining which elements of SSIs to condense versus preserve presents a particular design challenge for the creation of SSIs embedded on the web.

Several intervention design principles, drawn from basic research in social psychology, education, and marketing [29], have been theorized to support the use of web-based SSIs for mental health: (1) using neuroscientific evidence to normalize youths' experiences and boost message credibility; (2) centering youths' expertise in their own lived experiences; (3) asking youths to share what they have learned with others, using their own words; and (4) providing testimonials from others experiencing similar and relevant challenges [21]. These core

principles have been featured in much of the existing research evaluating web-based SSIs [22-24]. However, limited research to date has evaluated whether *very* brief, streamlined SSIs (ie, SSIs much shorter than 30 minutes) that retain these core principles may still provide some benefit as quick, free mental health supports embedded on the web.

This Study

As a majority of existing mental health treatments remain inaccessible, many young people seek mental health support via social platforms on the web. SSIs are uniquely well positioned for integration within web-based social platforms, providing free, brief, and anonymous mental health support options at scale. However, little research has formally evaluated the acceptability and effectiveness of very brief SSIs designed for this context. This study adapted 3 web-based SSIs (5-8 minutes) that were offered as Koko minicourses on Tumblr, a microblogging and social networking website with 135 million active users monthly [30]. The aims of this study were 2-fold: (1) to evaluate the *acceptability* of SSIs in this context via user data and acceptability ratings and (2) to assess *short-term effectiveness* of all 3 SSIs in improving key mental health outcomes (ie, hopelessness, self-hate, desire to discontinue self-harm) from before to after the SSI. Given the unique barriers to mental health treatment access faced by teenagers, the analyses were restricted to teenagers (ages 13-19 years) where possible. The results may inform continued efforts to increase access to free, anonymous, and timely mental health support options.

Methods

Ethical Considerations

For this study, we collected anonymous data exclusively from individuals on Tumblr—all of whom were introduced to the service by either (1) clicking on a featured advertisement from Tumblr (eg, "take control by taking this mood-boosting minicourse") or (2) direct referral from the platform. As all data were part of a completely anonymous program evaluation, this study was deemed as nonhuman subjects research in consultation with the institutional review board at Stony Brook University. In addition, Koko's privacy policy and terms of service acknowledge that anonymized data may be shared for research purposes.

Recruitment and Procedure

The direct referral pathway for each of the 3 SSIs was similar. Users who searched for mental health topics on Tumblr were shown an in-app overlay with links to various resources, such as The National Suicide Prevention Lifeline. A set of over 1300 keywords and their derivations were used to detect terms such as "self-harm" or "depression" as well as slang and obfuscations, such as "sewer-slide" and "'s3lf h@rm." In addition to links to crisis lines, users were also sent a direct message from Koko through the Tumblr direct message channel. Specifically, they were sent the following automated message from a chatbot called "Kokobot":

> *Hi! I'm Kokobot [wave emoji]. I'm working with Tumblr to connect people who are interested in mental health topics. Type "hi" to get started...*

Next, users were onboarded to the service and asked to describe a recent negative situation that they have been facing, along with any associated negative thoughts. From there, a set of text-based classifiers for mental health [20,31] categorized the user's posts and directed them to one of several resources, including peer support, crisis lines, and SSIs. Users could also access any of these services from a main menu at any point while using the platform. Users whose text descriptions were flagged by a crisis model were asked to specify whether their current struggles were related to suicidal thoughts, abuse, eating disorders, or self-harm. If a user disclosed that they were struggling with self-harm, they were shown crisis lines from around the world as well as a link to Project SAVE.

The 3 SSIs were initially introduced as Koko minicourses at 3 separate times: March 2021 (Project ABC), June 2021 (REFRAME), and July 2021 (Project SAVE). Across all 3 SSIs, the data for this study were collected through February 2022. Preintervention data were collected immediately before beginning each SSI (ie, each Koko minicourse), and postintervention data were collected immediately following the completion of a program.

SSI Programs

Project ABC Single-Session Intervention

Project ABC minicourse was a briefer 5- to 8-minute version of the original 20- to 30-minute Project ABC SSI evaluated in earlier randomized trial research [22]. As with the full-length program, the abbreviated Project ABC SSI used principles of behavioral activation, encouraging individuals to "take action" and engage in pleasurable behaviors that align with personal values to boost mood and build self-efficacy. Similar to the original SSI, the brief minicourse involved the following components: (1) psychoeducation, describing how taking values-based *actions* can boost *mood* over time; (2) values assessment, where individuals identify a top value for them (eg, academics, friendships, hobbies, family, or staying active); (3) "action plan" exercise, where individuals develop a personalized plan for engaging in meaningful activities of their choice; (4) roadblock exercise designed to identify real-life obstacles and how to address them; and (5) writing prompts, where individuals can share what they have learned with other Koko users. This shortened version of the Project ABC SSI contained condensed intervention content, fewer writing prompts, and greater emphasis on multiple-choice, interactive options, relative to the original 30-minute program.

Project SAVE Single-Session Intervention

Before disseminating as a Koko minicourse, an abbreviated (8-minute) version of the Project SAVE SSI was adapted from an original 30-minute program [32]. Project SAVE was designed to reduce the use of self-harm behaviors to cope with high emotional distress, especially in the context of self-hatred or desires to punish oneself [33]. Similar to the original SSI, the 8-minute version of Project SAVE included evidence-based techniques that are common in cognitive behavior therapies (eg,

psychoeducation and secondary distress tolerance skills). Specifically, Project SAVE included four key components: (1) information about how changing *actions* (eg, reducing self-harm behavior, being kinder to yourself) can change *feelings* and *thoughts* for the better; (2) statistics and testimonials from other teenagers with lived experiences of coping with self-hatred or self-harm; (3) alternative coping strategies to use in place of self-harm; and (4) offering advice to others using Koko based on what they learned. Compared with the original, the shortened Project SAVE SSI contained condensed material, fewer writing exercises, and more multiple-choice, interactive options.

The REFRAME Single-Session Intervention

The REFRAME SSI (5 minutes) teaches cognitive reappraisal, an emotion regulatory strategy that involves modifying one's interpretation of stressful situations [34]. A recent multicountry study with 21,644 individuals found that a brief web-based training on cognitive reappraisal reduced negative emotions and increased positive emotions about COVID-19–related stressors [35]. The REFRAME SSI on Koko included similar components, but it did not specifically target COVID-19–related stressors and it did not distinguish between multiple reappraisal strategies (eg, reconstrual vs repurposing). It included (1) a brief introduction to cognitive reappraisal; (2) a simplified description of its neural correlates; (3) practice examples with vignettes taken from the Koko peer support platform; and (4) a set of brief prompts to help users engage positive reappraisals in their own lives (eg, "This could get better because..."; "This isn't 100% my fault because..."). Users were also taught how to practice this skill on the Koko peer support platform and that helping others reappraise may confer positive psychological outcomes [36,37].

Measures

Demographics

Demographic information (age, gender, and race and ethnicity) was collected as part of the preintervention measures for individuals completing Project SAVE.

Hopelessness

The Beck Hopelessness Scale-4 is a brief and reliable measure used to assess hopelessness in young people [38,39]. At pre- and postintervention time points for the Project ABC and REFRAME SSIs, responders indicated their agreement "right now, in the present moment" with each of 4 statements (eg, "I feel that my future is hopeless and things will not improve") on a 4-point Likert scale (1="absolutely disagree"; 4="absolutely agree"). Hopelessness scores for each person (4-item average) ranged from 1 to 4, with higher values indicating greater hopelessness. Internal consistency was Cronbach α=.84 and Cronbach α=.89 for pre- and post-Project ABC time points, and α=.81 and Cronbach α=.87 for pre- and post-REFRAME SSI time points, respectively.

Self-hate

The Self-Hate Scale is a reliable 7-item measure used to assess self-hatred in young people [24,40]. At pre- and postintervention time points for the Project SAVE SSI, respondents indicated how true each statement (eg, "I hate myself") is for them, "right now, in the present moment," on a 7-point Likert scale (1="not at all true for me"; 4="somewhat true for me"; 7="true for me"). The self-hate scores for each person (7-item average) ranged from 1 to 7, with higher values indicating greater self-hate. Internal consistency was Cronbach α=.90 and Cronbach α=.94 for pre- and post-Project SAVE SSI time points, respectively.

Desire to Discontinue Self-harm

Desire to discontinue self-harm behavior was indexed using a single item adapted from the Self-Injurious Thoughts and Behaviors Interview–Revised [23,41]. Individuals were asked, "How much do you want to stop purposefully hurting yourself without wanting to die?" and told to rate their answer on a 5-point Likert scale (1="I don't want to stop at all"; 5="I definitely want to stop"). A sixth option (eg, "I have stopped engaging in these behaviors") was available for individuals not currently engaging in self-harm behaviors. Data for these individuals were not included in the analyses evaluating pre- to post-SSI changes in this outcome.

SSI Feedback

After completing all 3 SSIs, users were prompted to provide a quantitative "star rating" of the program, from 1 to 5 stars, where higher star values indicate higher ratings or more positive feedback. In addition, after the intervention, all individuals were asked if they would like to provide qualitative feedback via an optional writing prompt (eg, "Do you have any feedback for us?").

Statistical Analysis

Power

Previous web-based SSI research suggests small to large within-group effect sizes for key outcomes (including hopelessness and self-hatred) using more naturalistic study designs (ie, not randomized controlled trials) [24]. Thus, we estimated our statistical power to detect a small within-group effect (Cohen d=0.20) using a paired 2-tailed t test in the smallest of the 3 samples (ie, REFRAME, n=768 pairs).

Data Exclusion

We assessed the total number of views, starts, and completions for each SSI, as well as item-level drop-off data within each SSI, to describe broad usage patterns without excluding any data. Demographic and outcome data were only recorded and made available once individuals advanced through to the end of a program and clicked "submit." All pre-post analyses, star ratings, and qualitative data were therefore conducted and reported within program "completers" (ie, individuals who completed an SSI). In addition, where demographic data were available (Project SAVE SSI data), we restricted our analysis to teenagers (ie, excluding individuals who reported ages outside of 13-19 years). As Project SAVE was designed for teenagers engaging in self-harm, only individuals who self-disclosed recently engaging in self-harm (via the Koko onboarding pathways described earlier in this section) were directed to complete this program.

Usage Patterns and SSI Feedback

For each SSI, we evaluated usage patterns and feedback, including the number of people who viewed (ie, opened the first page of the program), started (ie, advanced past the first page), completed (ie, advanced through the entire program and beyond the questions), and provided "star ratings" (ie, quantitative ranking of 1-5 stars, with higher stars reflecting a higher rating) for each SSI. In addition, we calculated the program completion rates (percentage completed out of those who started) and average star ratings for each SSI. To illustrate the types of qualitative feedback each SSI received, we extracted specific examples of positive and negative feedback (see the Results section).

Aggregate-level data were available for views and user dropout, for every page within each SSI, among the entire sample (ie, without focusing solely on program completers). We reported these results by plotting the percentage dropout from the number of views for each page within each SSI.

Evaluating Pre-Post Changes

We evaluated pre- to postintervention changes for 4 outcomes: hopelessness (via two separate tests, 1 for Project ABC and 1 for REFRAME) as well as self-hate and desire to discontinue self-harm (Project SAVE only). After checking appropriate assumptions (ie, verifying that pre-post difference scores for each outcome were approximately normally distributed) [42], we performed 2-tailed, paired *t* tests across all outcomes for all programs. We calculated Cohen d_z for paired *t* tests using *t* values, sample sizes, and the Measure of the Effect package in R (version 4.0.0) [43] to evaluate the magnitude of within-group changes for each outcome. For outcomes where difference scores were not normally distributed, we completed an additional sensitivity analysis using the nonparametric Wilcoxon signed-rank test to evaluate within-group changes from before to after the intervention [44]. We applied a false discovery rate correction to all 4 *P* values to reduce the likelihood of false positives [45]. Thus, we interpreted the results for each outcome as significant if the false discovery rate–corrected *P* values were <.05. All analysis code and deidentified data for this study have been made publicly available [46].

Results

Sample Characteristics

As Koko provides anonymous mental health support, individuals are not required to provide potentially identifiable demographic information (eg, age, gender, and race and ethnicity) to complete minicourses. Specific demographic information for this study (age, gender, and race and ethnicity) was only available for the Project SAVE SSI, as this minicourse was introduced to Koko *after* Project ABC and REFRAME SSIs had been introduced. For the Project SAVE minicourse, individuals were given the option (but were not required) to report demographic information about themselves.

The Project SAVE analyses excluded individuals who were not teenagers between the ages of 13 and 19 years, resulting in a final sample of 1194 individuals (72.28% of the total Project SAVE data). Among this group, the average age of the individuals who completed Project SAVE was 15.71 (SD 1.83) years. The top three most commonly endorsed gender identities were female (419/1194, 35.09%), nonbinary (189/1194, 15.83%), and not sure (98/1194, 8.21%). The top three most commonly endorsed racial or ethnic identities were White (607/1194, 50.84%); Asian (172/1194, 14.41%); and Hispanic or Latinx (112/1194, 9.38%). Table 1 presents complete details on participant demographics.

Table 1. Gender identity and race and ethnicity for Project Stop Adolescent Violence Everywhere (N=1194).

Demographics	Values, n (%)
Gender	
Agender	35 (2.93)
Androgynous	13 (1.09)
Female	419 (35.09)
Female to male transgender	74 (6.20)
Gender expansive	15 (1.26)
Gender identity not listed	18 (1.51)
Gender information missing	219 (18.34)
Intersex	3 (0.25)
Male	27 (2.26)
Male to female transgender	2 (0.17)
Nonbinary	189 (15.83)
Not sure	98 (8.21)
Prefer not to say	26 (2.18)
Transfeminine gender	2 (0.17)
Trans man	10 (0.84)
Transmasculine gender	44 (3.69)
Transgender	15 (1.26)
Two-spirited	2 (0.17)
Race and ethnicity[a]	
Asian	172 (14.41)
Black or African American	66 (5.53)
Hispanic or Latinx	112 (9.38)
Native American or Alaska Native	23 (1.93)
Native Hawaiian or other Pacific Islander	12 (1.01)
White	607 (50.84)
Prefer not to answer	125 (10.47)

[a]Individuals could select multiple racial and ethnic identities.

Statistical Analysis

Power

Power analyses indicated that we had 99% power to detect a small effect size (Cohen *d*=0.20) using a 2-tailed paired *t* test in the smallest of the 3 current samples (REFRAME, n=768 pairs).

Usage Patterns and SSI Feedback

Project ABC

Project ABC was viewed 17,620 times, started 14,434 times, and completed 3679 times—with a 25.49% (3679/14,434) completion rate among starters across the 12-month study period. Among the Project ABC completers, 3412 (92.74%) provided a star rating with an average star rating of 4.27 (SD 0.94; median 5) stars. In total, 1217 (33.08%) provided qualitative feedback on Project ABC (see Table 2 for examples of user feedback). The dropout percent of the number of views for each page for all 3 SSIs is plotted in Figure 1. Notably, high levels of dropout (relative to page views, 7%-13%) consistently occurred on pages where Project ABC users were asked to respond to a writing prompt.

Finally, Project ABC received far more views, starts, and completions than either of the other two SSI programs (4065 and 2174 views for Project SAVE and REFRAME, respectively), as Tumblr advertised the Project ABC SSI as a featured minicourse between December 2021 and February 2022, resulting in higher traffic to this SSI.

Table 2. Examples of positive and critical feedback for all single-session interventions (SSIs).

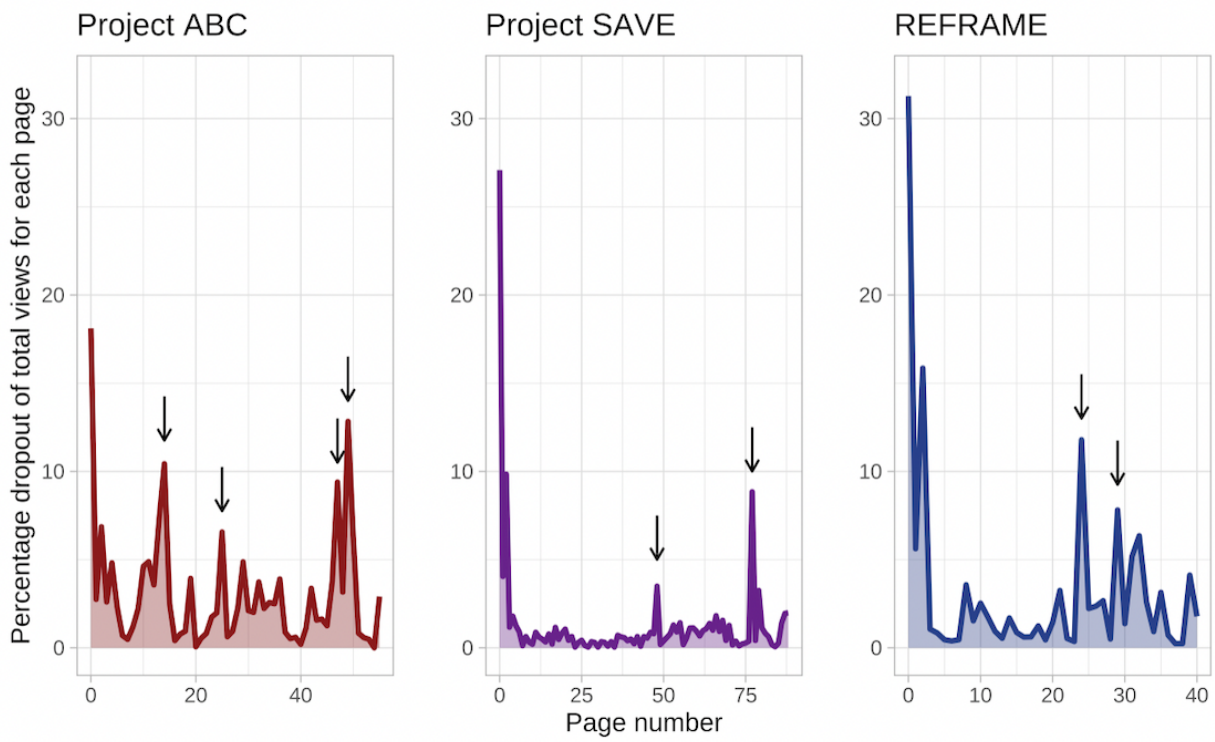
SSI	Positive feedback	Critical feedback
ABC[a]	"Thanks y'all! been dealing w some serious mental health issues and having places to remind me of my agency and joy is really helpful."	"Despite the great intentions and work I think there is situations that are very complicated and having this intermediate bot, very few info of the person you are helping it's too simplistic."
	"this was really really helpful and i'm seriously going to try my goal/plan. I also feel awake and motivated enough to study. I'd love to see more in the future."	"This helped me see get through my rain cloud but now I kinda feel stressed."
	"Hey, this was surprisingly well done. I went in expecting it would be terrible. But you're making mental health a really approachable topic for people. Thanks for working on reaching out to others. I'd love to see you continue with these mini-courses."	"This is a good idea, but only works for the things one has control over. If you are terminally ill, just lost a loved one or in another uncontrollable challenge, none of these can help."
REFRAME[b]	"This is such a great way to de-stress, I mean re-frame your stress, it's definitely a bit helpful."	"Less simplification"
	"I decided to pick up my phone and do this while I was procrastinating. It's so crazy how this seemingly small task changed my perspective."	"i think that one of the problems i see is that this requires people to be more specific and there isn't a sense of connection."
	"I love this so much!! It actually made me feel better, which I didn't think it would! Thank you <3."	"Please include physical stress reducers, as well. It is difficult to focus on reducing stress when I can't properly string thoughts together."
SAVE[c]	"This is amazing. My thoughts of self-harm faded a bit, and prompted me to do the things alternative to when I have self-hate thoughts."	"It would've been more helpful if reasons for self-harm outside of self-hatred were explored. I'm currently dealing with external circumstances that are overwhelming me and my response is to harm myself. It feels like the only way to release my emotions."
	"This is the most convincing thing I've ever heard as to why not to self harm. Thank you so much this is so helpful."	"I have dealed with this so long that there wasn't really anything new to me, so it really didn't help."
	"Thank you. This course has definitely calmed me down when I was having a breakdown and thinking of hurting myself."	"Talk about slowly building up to recovery instead of just jumping right in. Talk about how to deal with intense emotions and more."

[a]Project Action Brings Change (ABC) SSI.

[b]REFRAME SSI.

[c]Project Stop Adolescent Violence Everywhere (SAVE) SSI.

Figure 1. Percentage dropout on each page of all 3 single-session interventions (SSIs), out of the number of individuals who viewed that page. Arrows reflect points where writing prompts were introduced, in each of the 3 SSIs. Spikes in dropout tended to occur after initially opening each SSI, as well as on pages requesting written responses. ABC: Action Brings Change; REFRAME; SAVE: Stop Adolescent Violence Everywhere.



Project SAVE

In 12 months, Project SAVE was viewed 4065 times, started 2961 times, and completed 1652 times; 55.79% (1652/2961) of those who started Project SAVE completed it. After excluding individuals with ages outside our desired range (13-19 years), 1194 observations remained for analysis. Among those who completed the minicourse, 954 (79.90%) provided a star rating for Project SAVE, with an average rating of 4.22 (SD 0.97; median 5) stars. A total of 209 (17.50%) participants provided qualitative feedback on Project SAVE (Table 2). Similar to the Project ABC SSI, Project SAVE observed higher dropout rates (relative to page views, 3%-9%) on pages where users were asked to enter a written response (Figure 1).

REFRAME

REFRAME was viewed 2174 times, started 1498 times, and completed 848 times within 12 months (848/1498, 56.60% completion rate among the starters). Among REFRAME completers, 732 (86.32%) provided a star rating for the REFRAME SSI, with an average rating of 4.31 (SD 0.93; median 5) stars; 246 (29.01%) provided qualitative feedback on the REFRAME minicourse (Table 2). Consistent with dropout patterns for the other 2 SSIs, REFRAME observed higher dropout rates (relative to page views, 8%-12%) on pages requesting writing from users.

Evaluating Pre-Post Changes

Among individuals who completed Project ABC, hopelessness significantly decreased from before to after the SSI ($t_{3,565}$=−48.48, P<.001). Similarly, individuals who completed the REFRAME SSI reported significant reductions in hopelessness from before to after the intervention (t_{767}=−24.41, P<.001). Project SAVE completers reported significant pre- to postintervention reductions in self-hate ($t_{1,007}$=−21.30, P<.001) and an increase in desire to discontinue self-harm (t_{839}=11.48, P<.001). All the means, SDs, and within-group effect sizes are listed in Table 3.

Table 3. Means, SDs, and effect sizes for all single-session intervention (SSI) outcomes.

Outcome and SSI	Before the SSI, mean (SD)	After the SSI, mean (SD)	Cohen d_z (95% CI)
Hopelessness			
ABC[a]	2.60 (0.78)	2.16 (0.80)	−0.81 (−0.85 to −0.77)
REFRAME[b]	2.86 (0.74)	2.31 (0.78)	−0.88 (−0.96 to −0.80)
Self-hate			
SAVE[c]	5.69 (1.27)	5.07 (1.64)	−0.67 (−0.74 to −0.60)
Desire to stop harm[d]			
SAVE	2.63 (1.20)	2.97 (1.32)	0.40 (0.33 to 0.47)

[a]Project Action Brings Change (ABC) SSI.

[b]REFRAME SSI.

[c]Project Stop Adolescent Violence Everywhere (SAVE) SSI.

[d]Desire to discontinue self-harm behavior.

In addition, Wilcoxon signed-rank tests were performed as sensitivity analyses for all outcomes due to the relatively nonnormal distributions of difference scores. For all outcomes, results were consistent with *t* test analyses, indicating significant pre-post reductions in hopelessness ($P<.001$; Project ABC and REFRAME) and self-hate ($P<.001$; Project SAVE), as well as a significant increase in desire to discontinue self-harm ($P<.001$; Project SAVE) from before to after the intervention. Distributions for all outcomes at pre- and post-SSI time points are shown in Figures 2 and 3.

Figure 2. Hopelessness ratings before and after the Project Action Brings Change (ABC) single-session intervention (SSI; left) and before and after the REFRAME SSI (right). Higher scores reflect higher levels of hopelessness.

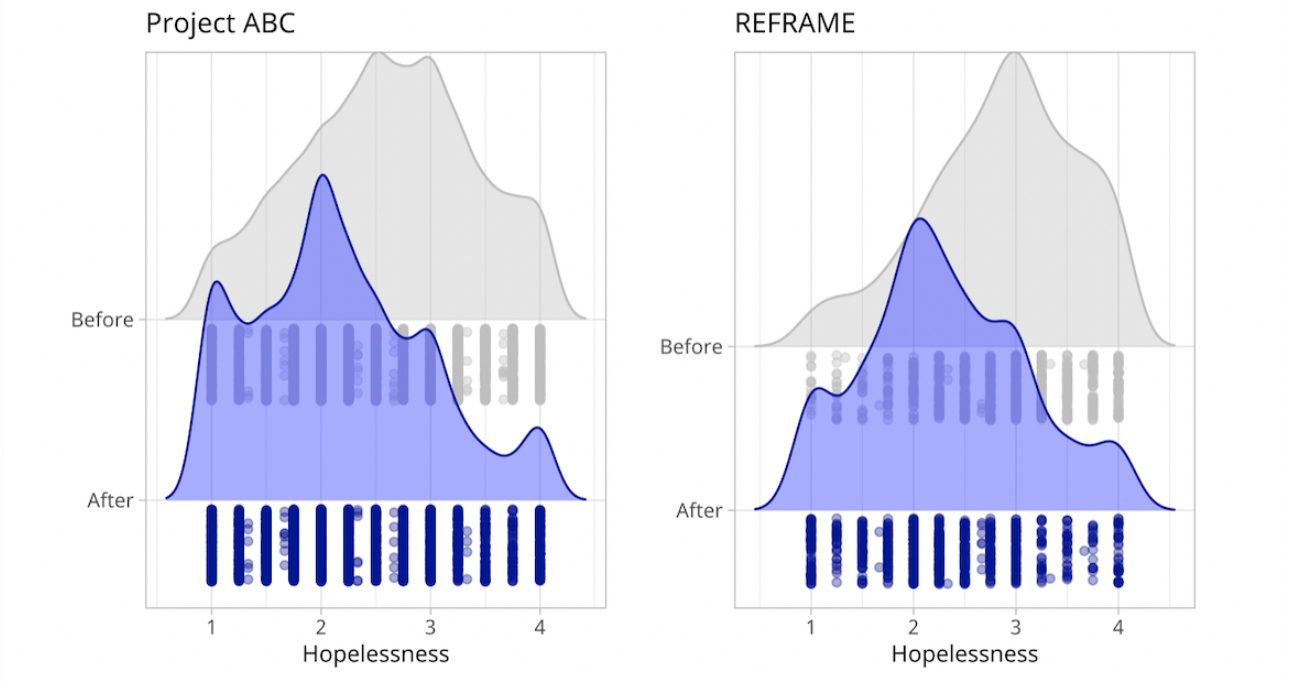
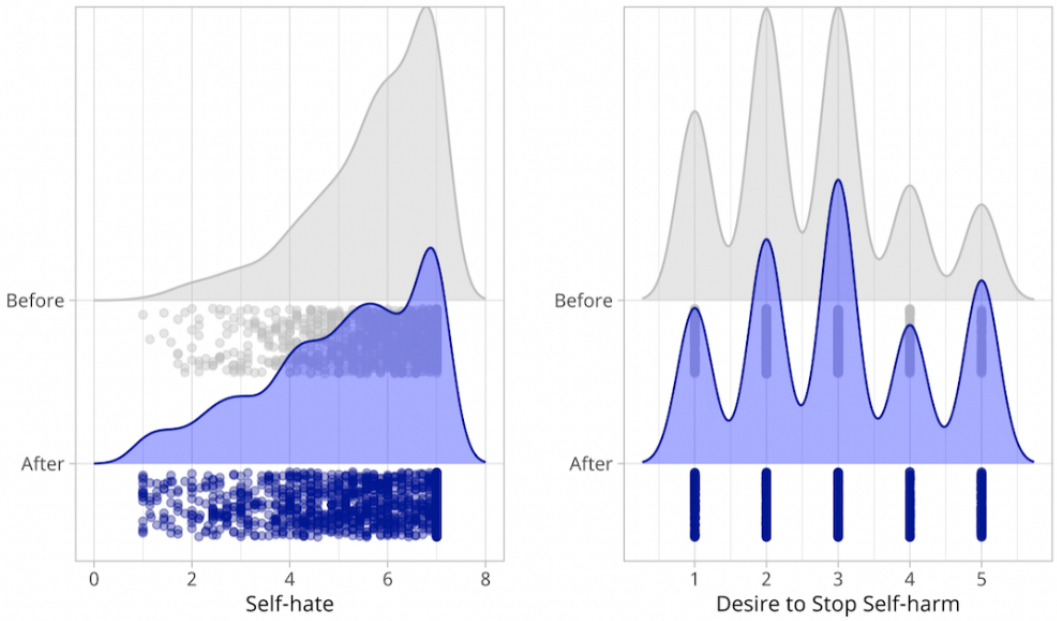


Figure 3. Self-hate ratings before and after the Project Stop Adolescent Violence Everywhere single-session intervention (left), where higher scores reflect higher levels of self-hatred. (Right) Desire to stop future self-harm behavior, where higher scores reflect higher desire to stop future self-harm.



Discussion

Principal Findings

Within 12 months, 3 very brief (5-8 minutes) web-based SSIs were viewed >18,800 times and completed >6100 times. Offered as Koko minicourses and embedded within a popular social platform, all 3 SSIs received high-quality ratings (average ratings >4 out of 5 stars). Of the participants who started an SSI, between 25% and 57% of them completed one. Among those who completed Project ABC and REFRAME minicourses, individuals reported a decrease in hopelessness from before to after the SSI. For those who completed Project SAVE, individuals reported decreased self-hatred and increased desire to stop future self-harm behavior from before to after the intervention. This study represents a real-world evaluation of the acceptability and short-term utility of web-based SSIs as very brief, anonymous, "in-the-moment" mental health supports that can be integrated within major social platforms such as Tumblr.

Comparison With Prior Work

Consistent with existing research on web-based SSIs [24], all 3 interventions were (1) rated as acceptable by users and (2) associated with improvements across primary outcomes. Furthermore, despite reducing intervention length by *nearly 75%*, effect sizes for these 5- to 8-minute SSIs were remarkably similar to effect sizes for SSIs designed to last 30 minutes (Cohen d_z =0.71 vs 0.81 (hopelessness) and Cohen d_z =0.61 vs 0.67 (self-hate) for full and abbreviated SSIs, respectively) [24]. These results support a growing body of evidence that suggests even extremely brief digital interventions may improve clinically relevant outcomes [20,47].

One possible reason for the relative similarity of postintervention effect sizes observed in this study (5-8–minute SSIs) versus within-group effects in earlier randomized trials (30-minute SSIs) [22,23] might be the shared design features and principles across the "brief" and "very brief" versions of these interventions. Given that these core principles have been featured in all versions of these SSIs to date, it is possible that these common principles are more important than the exact length of the SSI. Future research should formally evaluate this possibility in a head-to-head trial by comparing SSIs of various lengths.

Our results may also indicate the *real-world utility* of these very brief web-based SSIs. Naturalistic study designs similar to that used in this study are crucial for evaluating acceptability and utility among web-based SSIs, as randomized trials overestimate user engagement levels for unguided, digital mental health interventions—often producing usage rates that are *4 times* greater than rates observed in real-world settings [27]. Once these interventions are disseminated outside of randomized trials, a small minority of individuals (0.5%-28.6%) complete all content [28]. Therefore, relative to real-world completion rates of other digital self-help interventions for mental health, completion rates for the 3 SSIs tested in this study (25%-57%) were extraordinarily high.

Low real-world completion rates for self-help digital interventions, especially relative to guided digital supports [26,48,49], are unsurprising; unguided supports require substantial emotional and mental bandwidth (eg, energy, sustained attention, and motivation) to find and engage with an unguided mental health support tool at the very moments when these capacities may be at their lowest (eg, when experiencing elevated distress). Our study sought to minimize the bandwidth necessary to engage with 3 single-session, unguided supports by (1) dramatically reducing their length and (2) bringing these interventions to web-based spaces where people already are. Future research should explicitly assess whether these strategies improve completion rates across other unguided digital mental health supports.

Embedding SSIs within a popular social platform (Tumblr) also likely increased the *visibility* and uptake of these SSIs. For example, after 6 months of recruitment using paid advertising and leveraging substantial media coverage, 1 study found that 700 youths viewed an open-access, web-based mental health platform featuring 3 SSIs [24]. The 3 SSIs in this study received a combined ≥18,800 views within 12 months—nearly 6240 of which were views of interventions that were never featured in a Tumblr-based advertisement (Project SAVE and REFRAME). These numbers suggest considerable interest in accessible, anonymous, "in-the-moment" support options, as well as the potential to *sustainably* offer these supports at high scale and relatively low cost.

Notably, existing research identifies safety as a primary ethical concern for researchers and stakeholders interested in using digital mental health tools among youth [50]; many digital resources lack evidence or clinical validation and others may not provide sufficient protection for sensitive user data. By using SSIs that have previously been evaluated in rigorous randomized trial research [22,23] and by collecting anonymous pre- and post-SSI user data, in this study, we aimed to mitigate these common concerns. Thus, calls for clearer guidelines for safety and quality in mental health apps [51] should also be accompanied by clearer guidelines, both for academic- and industry-based program developers and researchers, on how to responsibly and ethically disseminate mental health resources via partnerships with social networking platforms.

Strengths and Limitations

This study has several strengths. Although randomized trial research overestimates user engagement for digital interventions [27], it constitutes much of the existing research on web-based SSIs to date [22,23,52]. Therefore, the naturalistic design of this study provides a valuable, more accurate estimate of the potential for web-based SSIs to offer mental health supports at scale. Koko's partnership with Tumblr also made it possible to offer anonymous, "in-the-moment" support for people seeking mental health–related content within the platform, rather than requiring individuals to search for external resources. Finally, we conducted a fine-grained analysis of usage data (ie, dropout) going item by item for each SSI. It was especially valuable to recognize high levels of user dropout relative to page views (3%-13%) on pages requesting written responses. Further studies should seek to understand patterns in user dropout and how they may inform digital intervention design. In addition, future evaluations may also want to expand upon this study by conducting smaller-scale usability testing [53] and research explicitly designed for the in-depth analysis of rich qualitative user feedback [54].

In addition to the aforementioned strengths, this study has some limitations. First, demographic information was not available for all individuals included in this study. Demographic information that *was* available from 1 of the 3 SSIs indicated substantial missing data (several users reported skipping demographic items to ensure they could not be identified). Black, Hispanic, and Asian youths consistently access mental health treatment at lower rates than their non-Hispanic White peers [55,56], yet up to 90% of youths in mental health clinical trials are White [57,58]. Digital interventions provide one possible avenue toward reducing disparities in access to mental health resources among racial and ethnic minority populations [59,60]; a vast majority of adolescents in the United States have internet access via either smartphones or desktops or laptops (88% and 95%, respectively) [61]. While there are disparities in broadband access by socioeconomic status, geographic region, educational attainment, and race and ethnicity, some evidence suggests that these gaps in access may be decreasing [62]. Should SSIs hope to provide truly *accessible* and *equitable* mental health supports, they must be accompanied with consistent evaluations of who has access to them, with researchers actively working to prevent further (or exacerbated) inequity via a new treatment modality [60].

In addition, as is often the case in web-based mental health support research [24,63,64], cisgender boys were underrepresented in this sample. Given unique barriers to seeking mental health treatments faced by cisgender boys and young men (eg, internalized masculine gender norms, such as "toughness"; elevated mental health stigma) [65,66], future mixed methods research may want to evaluate whether boys view SSIs as a less stigmatizing or more approachable mental health support option.

Finally, given that this study represents an unmasked evaluation of within-group intervention effects in a completers-only sample, our ability to draw causal inferences was limited. For example, those who completed an SSI may have provided more positive star ratings for the programs than those who exited before finishing the program. However, 2 of the 3 SSIs featured in this study have demonstrated positive effects on identical outcomes in large-scale, triple-masked, and randomized research [22,23]. SSIs have shown considerable promise in multiple studies and study designs.

Conclusions

Very brief SSIs (5-8 minutes) can be embedded within web-based social platforms as anonymous, in-the-moment mental health supports capable of reaching many individuals within months. Among those who complete these SSIs, individuals generally rate them as acceptable, and pre-post evaluations suggest that they may be helpful in reducing hopelessness and self-hate as well as in increasing the desire to stop self-harm. These pre-post findings, combined with results of earlier randomized trial research, suggest that SSIs delivered in this context may be a sustainable approach for providing the much-needed mental health resources, particularly for teenagers who may not have access to other mental health supports. Considering the substantial unmet need for mental health care among teenagers in the United States [67], SSIs may provide a valuable and complementary source of mental health support among a broader ecosystem of mental health treatment options (eg, school- and community-based programs, mental health screening, and gatekeeper training), all of which are necessary for reducing mental illness at scale.

Acknowledgments

JLS received funding from the National Institute of Health Office of the Director (DP5OD028123), National Institute of Mental Health (R43MH128075), National Science Foundation (2141710), Health Research and Services Association (U3NHP45406-01-00), Upswing Fund for Adolescent Mental Health, Society for Clinical Child and Adolescent Psychology, and the Klingenstein Third Generation Foundation.

MLD received funding from the Graduate Council Fellowship at Stony Brook University and has previously received research funding from the Psi Chi International Honors Society and the University of Denver Faculty Research Fund.

Authors' Contributions

JLS, RM, and MLD conceptualized the project, contributed to the study design, and developed the original versions of the single-session interventions. RM adapted the original intervention materials for integration into the Koko. MLD performed data analyses and wrote the initial draft of the manuscript. All the authors contributed to the review and editing of the final manuscript.

Conflicts of Interest

MLD receives book royalties from New Harbinger and funding from the Graduate Council Fellowship at Stony Brook University. She has previously received research funding from Psi Chi International Honors Society and University of Denver Faculty Research Fund. RM is the cofounder and chief executive officer of Koko, a nonprofit organization that receives financial aid from industry partners, private donors, and Hopelab Ventures. He also serves as a scientific advisor for Homecoming, a digital platform to help support psychedelic-assisted psychotherapy. JLS serves on the Scientific Advisory Board for Walden Wise and the Clinical Advisory Board for Koko; is the cofounder and codirector of Single Session Support Solutions, Inc.

References

1. Brown A, Rice SM, Rickwood DJ, Parker AG. Systematic review of barriers and facilitators to accessing and engaging with mental health care among at-risk young people. Asia Pac Psychiatry 2016 Mar;8(1):3-22. [doi: 10.1111/appy.12199] [Medline: 26238088]
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. Chen L, Crum RM, Martins SS, Kaufmann CN, Strain EC, Mojtabai R. Service use and barriers to mental health care among adults with major depression and comorbid substance dependence. Psychiatr Serv 2013 Sep 01;64(9):863-870 [FREE Full text] [doi: 10.1176/appi.ps.201200289] [Medline: 23728427]
4. Hom MA, Stanley IH, Joiner TE. Evaluating factors and interventions that influence help-seeking and mental health service utilization among suicidal individuals: a review of the literature. Clin Psychol Rev 2015 Aug;40:28-39. [doi: 10.1016/j.cpr.2015.05.006] [Medline: 26048165]
5. Fox KR, Bettis AH, Burke TA, Hart EA, Wang SB. Exploring adolescent experiences with disclosing self-injurious thoughts and behaviors across settings. Res Child Adolesc Psychopathol 2022 May;50(5):669-681. [doi: 10.1007/s10802-021-00878-x] [Medline: 34705197]
6. Boyd RC, Butler L, Benton TD. Understanding adolescents' experiences with depression and behavioral health treatment. J Behav Health Serv Res 2018 Jan;45(1):105-111. [doi: 10.1007/s11414-017-9558-7] [Medline: 28488156]
7. Breland DJ, McCarty CA, Zhou C, McCauley E, Rockhill C, Katon W, et al. Determinants of mental health service use among depressed adolescents. Gen Hosp Psychiatry 2014;36(3):296-301 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.12.003] [Medline: 24417955]
8. Rickwood DJ, Mazzer KR, Telford NR. Social influences on seeking help from mental health services, in-person and online, during adolescence and young adulthood. BMC Psychiatry 2015 Mar 07;15:40 [FREE Full text] [doi: 10.1186/s12888-015-0429-6] [Medline: 25886609]
9. Pretorius C, Chambers D, Coyle D. Young people's online help-seeking and mental health difficulties: systematic narrative review. J Med Internet Res 2019 Nov 19;21(11):e13873 [FREE Full text] [doi: 10.2196/13873] [Medline: 31742562]
10. Coping with COVID-19: how young people use digital media to manage their mental health. Common Sense Media. 2021 Mar 15. URL: https://www.commonsensemedia.org/research/coping-with-covid19-how-young-people-use-digital-media-to-manage-their-mental-health [accessed 2022-06-29]
11. Wetterlin FM, Mar MY, Neilson EK, Werker GR, Krausz M. eMental health experiences and expectations: a survey of youths' Web-based resource preferences in Canada. J Med Internet Res 2014 Dec 17;16(12):e293 [FREE Full text] [doi: 10.2196/jmir.3526] [Medline: 25519847]
12. Pretorius C, Chambers D, Cowan B, Coyle D. Young people seeking help online for mental health: cross-sectional survey study. JMIR Ment Health 2019 Aug 26;6(8):e13524 [FREE Full text] [doi: 10.2196/13524] [Medline: 31452519]

13. Fish JN, McInroy LB, Paceley MS, Williams ND, Henderson S, Levine DS, et al. "I'm kinda stuck at home with unsupportive parents right now": LGBTQ youths' experiences with COVID-19 and the importance of online support. J Adolesc Health 2020 Sep;67(3):450-452 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.06.002] [Medline: 32591304]
14. Horgan A, Sweeney J. Young students' use of the Internet for mental health information and support. J Psychiatr Ment Health Nurs 2010 Mar;17(2):117-123. [doi: 10.1111/j.1365-2850.2009.01497.x] [Medline: 20465756]
15. Burns JM, Birrell E, Bismark M, Pirkis J, Davenport TA, Hickie IB, et al. The role of technology in Australian youth mental health reform. Aust Health Rev 2016 Nov;40(5):584-590. [doi: 10.1071/AH15115] [Medline: 26934382]
16. Williams AJ, Nielsen E, Coulson NS. "They aren't all like that": perceptions of clinical services, as told by self-harm online communities. J Health Psychol 2020;25(13-14):2164-2177 [FREE Full text] [doi: 10.1177/1359105318788403] [Medline: 30024273]
17. Jones R, Sharkey S, Ford T, Emmens T, Hewis E, Smithson J, et al. Online discussion forums for young people who self-harm: user views. Psychiatrist 2018 Jan 02;35(10):364-368. [doi: 10.1192/pb.bp.110.033449]
18. Morris RR, Schueller SM, Picard RW. Efficacy of a Web-based, crowdsourced peer-to-peer cognitive reappraisal platform for depression: randomized controlled trial. J Med Internet Res 2015 Mar 30;17(3):e72 [FREE Full text] [doi: 10.2196/jmir.4167] [Medline: 25835472]
19. Doré BP, Morris RR. Linguistic synchrony predicts the immediate and lasting impact of text-based emotional support. Psychol Sci 2018 Oct;29(10):1716-1723. [doi: 10.1177/0956797618779971] [Medline: 30020863]
20. Jaroszewski AC, Morris RR, Nock MK. Randomized controlled trial of an online machine learning-driven risk assessment and intervention platform for increasing the use of crisis services. J Consult Clin Psychol 2019 Apr;87(4):370-379. [doi: 10.1037/ccp0000389] [Medline: 30883164]
21. Schleider JL, Dobias ML, Sung JY, Mullarkey MC. Future directions in single-session youth mental health interventions. J Clin Child Adolesc Psychol 2020;49(2):264-278 [FREE Full text] [doi: 10.1080/15374416.2019.1683852] [Medline: 31799863]
22. Schleider JL, Mullarkey MC, Fox KR, Dobias ML, Shroff A, Hart EA, et al. A randomized trial of online single-session interventions for adolescent depression during COVID-19. Nat Hum Behav 2022 Feb;6(2):258-268 [FREE Full text] [doi: 10.1038/s41562-021-01235-0] [Medline: 34887544]
23. Dobias ML, Schleider JL, Jans L, Fox KR. An online, single-session intervention for adolescent self-injurious thoughts and behaviors: results from a randomized trial. Behav Res Ther 2021 Dec;147:103983. [doi: 10.1016/j.brat.2021.103983] [Medline: 34688102]
24. Schleider JL, Dobias M, Sung J, Mumper E, Mullarkey MC. Acceptability and utility of an open-access, online single-session intervention platform for adolescent mental health. JMIR Ment Health 2020 Jun 30;7(6):e20513 [FREE Full text] [doi: 10.2196/20513] [Medline: 32602846]
25. Ash J, Anderson B, Gordon R, Langley P. Digital interface design and power: friction, threshold, transition. Environ Plan D 2018 Apr 13;36(6):1136-1153. [doi: 10.1177/0263775818767426]
26. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
27. Baumel A, Edan S, Kane JM. Is there a trial bias impacting user engagement with unguided e-mental health interventions? A systematic comparison of published reports and real-world usage of the same programs. Transl Behav Med 2019 Nov 25;9(6):1020-1033. [doi: 10.1093/tbm/ibz147] [Medline: 31689344]
28. 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]
29. Lewin GW. Constructs in field theory. In: Lewin K, editor. Resolving Social Conflicts and Field Theory in Social Science. Washington, D.C., United States: American Psychological Association; 1944:191-199.
30. Press information. Tumblr. URL: https://www.tumblr.com/press [accessed 2022-06-29]
31. Kshirsagar R, Morris R, Bowman S. Detecting and explaining crisis. In: Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality. 2017 Presented at: CLPsych '17; August 3, 2017; Vancouver, BC, Canada. [doi: 10.18653/v1/w17-3108]
32. Dobias ML, Schleider JL, Fox KR. Project SAVE: stop adolescent violence everywhere. Open Science Framework. 2020. URL: https://osf.io/vguf4/ [accessed 2022-06-29]
33. Klonsky ED. The functions of deliberate self-injury: a review of the evidence. Clin Psychol Rev 2007 Mar;27(2):226-239. [doi: 10.1016/j.cpr.2006.08.002] [Medline: 17014942]
34. McRae K, Gross JJ. Emotion regulation. Emotion 2020 Feb;20(1):1-9. [doi: 10.1037/emo0000703] [Medline: 31961170]
35. Wang K, Goldenberg A, Dorison CA, Miller JK, Uusberg A, Lerner JS, et al. A multi-country test of brief reappraisal interventions on emotions during the COVID-19 pandemic. Nat Hum Behav 2021 Aug;5(8):1089-1110 [FREE Full text] [doi: 10.1038/s41562-021-01173-x] [Medline: 34341554]

36. Doré BP, Morris RR, Burr DA, Picard RW, Ochsner KN. Helping others regulate emotion predicts increased regulation of one's own emotions and decreased symptoms of depression. Pers Soc Psychol Bull 2017 May;43(5):729-739. [doi: 10.1177/0146167217695558] [Medline: 28903637]
37. Arbel R, Khouri M, Sagi J. Reappraising negative emotions reduces distress during the COVID-19 outbreak. PsyArXiv. Preprint posted online on September 28, 2020. [doi: 10.31234/osf.io/y25gx]
38. Rhoades H, Rusow JA, Bond D, Lanteigne A, Fulginiti A, Goldbach JT. Homelessness, mental health and suicidality among LGBTQ youth accessing crisis services. Child Psychiatry Hum Dev 2018 Aug 10;49(4):643-651. [doi: 10.1007/s10578-018-0780-1] [Medline: 29322361]
39. Perczel Forintos D, Rózsa S, Pilling J, Kopp M. Proposal for a short version of the Beck Hopelessness Scale based on a national representative survey in Hungary. Community Ment Health J 2013 Dec;49(6):822-830. [doi: 10.1007/s10597-013-9619-1] [Medline: 23756722]
40. Turnell AI, Fassnacht DB, Batterham PJ, Calear AL, Kyrios M. The Self-Hate Scale: development and validation of a brief measure and its relationship to suicidal ideation. J Affect Disord 2019 Feb 15;245:779-787. [doi: 10.1016/j.jad.2018.11.047] [Medline: 30448763]
41. Fox KR, Harris JA, Wang SB, Millner AJ, Deming CA, Nock MK. Self-injurious thoughts and behaviors interview-revised: development, reliability, and validity. Psychol Assess 2020 Jul;32(7):677-689. [doi: 10.1037/pas0000819] [Medline: 32324021]
42. Field A, Miles J, Field Z. Discovering Statistics Using R. Thousand Oaks, CA, USA: Sage Publications; 2012:390-391.
43. MOTE: Magnitude of the Effect - An Effect Size and CI calculator. GitHub. URL: http://github.com/doomlab/MOTE [accessed 2022-06-29]
44. Bauer DF. Constructing confidence sets using rank statistics. J Am Stat Assoc 1972 Sep;67(339):687-690. [doi: 10.1080/01621459.1972.10481279]
45. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 2018 Dec 05;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
46. Dobias M, Morris R, Schleider J. Analysis Code and De-identified Data. Open Science Framework. Preprint posted online on February 22, 2021. [FREE Full text] [doi: 10.31219/osf.io/k3prh]
47. Baumel A, Fleming T, Schueller SM. Digital micro interventions for behavioral and mental health gains: core components and conceptualization of digital micro intervention care. J Med Internet Res 2020 Oct 29;22(10):e20631 [FREE Full text] [doi: 10.2196/20631] [Medline: 33118946]
48. Moshe I, Terhorst Y, Philippi P, Domhardt M, Cuijpers P, Cristea I, et al. Digital interventions for the treatment of depression: a meta-analytic review. Psychol Bull 2021 Aug;147(8):749-786. [doi: 10.1037/bul0000334] [Medline: 34898233]
49. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions — a systematic review. Internet Interven 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
50. Wies B, Landers C, Ienca M. Digital mental health for young people: a scoping review of ethical promises and challenges. Front Digit Health 2021 Sep 6;3:697072 [FREE Full text] [doi: 10.3389/fdgth.2021.697072] [Medline: 34713173]
51. Torous J, Roberts LW. Needed innovation in digital health and smartphone applications for mental health: transparency and trust. JAMA Psychiatry 2017 May 01;74(5):437-438. [doi: 10.1001/jamapsychiatry.2017.0262] [Medline: 28384700]
52. Sung JY, Mumper E, Schleider JL. Empowering anxious parents to manage child avoidance behaviors: randomized control trial of a single-session intervention for parental accommodation. JMIR Ment Health 2021 Jul 06;8(7):e29538 [FREE Full text] [doi: 10.2196/29538] [Medline: 34255718]
53. Issa T, Isaias P. Usability and human computer interaction (HCI). In: Issa T, Isaias P, editors. Sustainable Design: HCI, Usability and Environmental Concerns. London, UK: Springer; 2015:19-36.
54. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
55. Cummings JR, Druss BG. Racial/ethnic differences in mental health service use among adolescents with major depression. J Am Acad Child Adolesc Psychiatry 2011 Feb;50(2):160-170 [FREE Full text] [doi: 10.1016/j.jaac.2010.11.004] [Medline: 21241953]
56. Elster A, Jarosik J, VanGeest J, Fleming M. Racial and ethnic disparities in health care for adolescents: a systematic review of the literature. Arch Pediatr Adolesc Med 2003 Sep;157(9):867-874. [doi: 10.1001/archpedi.157.9.867] [Medline: 12963591]
57. Kemp J, Barker D, Benito K, Herren J, Freeman J. Moderators of psychosocial treatment for pediatric obsessive-compulsive disorder: summary and recommendations for future directions. J Clin Child Adolesc Psychol 2021;50(4):478-485. [doi: 10.1080/15374416.2020.1790378] [Medline: 32706265]
58. Weisz JR, Kuppens S, Ng MY, Eckshtain D, Ugueto AM, Vaughn-Coaxum R, et al. What five decades of research tells us about the effects of youth psychological therapy: a multilevel meta-analysis and implications for science and practice. Am Psychol 2017;72(2):79-117. [doi: 10.1037/a0040360] [Medline: 28221063]
59. Ramos G, Chavira DA. Use of technology to provide mental health care for racial and ethnic minorities: evidence, promise, and challenges. Cogn Behav Pract 2022 Feb;29(1):15-40. [doi: 10.1016/j.cbpra.2019.10.004]

60. Friis-Healy EA, Nagy GA, Kollins SH. It is time to REACT: opportunities for digital mental health apps to reduce mental health disparities in racially and ethnically minoritized groups. JMIR Ment Health 2021 Jan 26;8(1):e25456 [FREE Full text] [doi: 10.2196/25456] [Medline: 33406050]
61. Anderson M, Jiang J. Teens, Social Media and Technology 2018. Pew Research Center. 2018. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed 2022-06-29]
62. Bauerly BC, McCord RF, Hulkower R, Pepin D. Broadband access as a public health issue: the role of law in expanding broadband access and connecting underserved communities for better health outcomes. J Law Med Ethics 2019 Jun;47(2_suppl):39-42 [FREE Full text] [doi: 10.1177/1073110519857314] [Medline: 31298126]
63. Kramer J, Conijn B, Oijevaar P, Riper H. Effectiveness of a Web-based solution-focused brief chat treatment for depressed adolescents and young adults: randomized controlled trial. J Med Internet Res 2014 May 29;16(5):e141 [FREE Full text] [doi: 10.2196/jmir.3261] [Medline: 24874006]
64. van der Zanden R, Kramer J, Gerrits R, Cuijpers P. Effectiveness of an online group course for depression in adolescents and young adults: a randomized trial. J Med Internet Res 2012 Jun 07;14(3):e86 [FREE Full text] [doi: 10.2196/jmir.2033] [Medline: 22677437]
65. Chandra A, Minkovitz CS. Stigma starts early: gender differences in teen willingness to use mental health services. J Adolesc Health 2006 Jun;38(6):754.e1-754.e8. [doi: 10.1016/j.jadohealth.2005.08.011] [Medline: 16730608]
66. Sileo KM, Kershaw TS. Dimensions of masculine norms, depression, and mental health service utilization: results from a prospective cohort study among emerging adult men in the United States. Am J Mens Health 2020;14(1):1557988320906980 [FREE Full text] [doi: 10.1177/1557988320906980] [Medline: 32079448]
67. Kazdin AE. Annual Research Review: expanding mental health services through novel models of intervention delivery. J Child Psychol Psychiatry 2019 Apr;60(4):455-472. [doi: 10.1111/jcpp.12937] [Medline: 29900543]

Abbreviations

ABC: Action Brings Change
SAVE: Stop Adolescent Violence Everywhere
SSI: single-session intervention

Edited by A Mavragani; submitted 25.04.22; peer-reviewed by D Coyle, L Balcombe; comments to author 16.05.22; revised version received 05.06.22; accepted 08.06.22; published 25.07.22.

Please cite as:
Dobias ML, Morris RR, Schleider JL
Single-Session Interventions Embedded Within Tumblr: Acceptability, Feasibility, and Utility Study
JMIR Form Res 2022;6(7):e39004
URL: https://formative.jmir.org/2022/7/e39004
doi:10.2196/39004
PMID:35877163

Original Paper

Development and Initial Testing of a Personalized, Adaptive, and Socially Focused Web Tool to Support Physical Activity Among Women in Midlife: Multidisciplinary and User-Centered Design Approach

Danielle Arigo[1,2], PhD; Andrea F Lobo[3], PhD; M Cole Ainsworth[1], PhD; Kiri Baga[1], BA; Kristen Pasko[1], MA

[1]Department of Psychology, Rowan University, Glassboro, NJ, United States

[2]Department of Family Medicine, Rowan School of Osteopathic Medicine, Stratford, NJ, United States

[3]Department of Computer Science, Rowan University, Glassboro, NJ, United States

Corresponding Author:
Danielle Arigo, PhD
Department of Psychology
Rowan University
201 Mullica Hill Road
Robinson 116G
Glassboro, NJ, 08028
United States
Phone: 1 8562564500 ext 53775
Email: arigo@rowan.edu

Abstract

Background: Women in midlife are vulnerable to developing cardiovascular disease, particularly those who have conditions such as hypertension. Physical activity (PA) can reduce risk, but efforts to promote PA in this population have been only modestly effective. More attention to social influences on PA behavior may be useful, particularly social support and social comparison processes. Activating these processes with digital tools can provide easy access that is flexible to the needs of women in midlife.

Objective: This paper describes the user-centered design processes of developing and conducting initial evaluation of a personalized and adaptive web application, tailored to the social needs of women in midlife. The goal was to gather feedback from the population of interest, before and during the design process.

Methods: This study was conducted in 4 stages. The first and second authors (DA and AFL) developed technical specifications, informed by their experience with the population of interest. We collected feedback on potential content for the web application with women in midlife using both interviews (5/10, 50%; mean age 47.4, SD 6.66 years; mean BMI 35.3, SD 9.55 kg/m^2) and surveys (5/10, 50%; mean age 51, SD 6.60 years; mean BMI 32.7, SD 8.39 kg/m^2). We used their feedback to inform support messages and peer profiles (ie, sources of social comparison information). Nine members of the behavioral science team and 3 testers unfamiliar with the web application completed internal testing. We conducted naturalistic functionality testing with a different group of women in midlife (n=5; mean age 50, SD 6.26 years; mean BMI 30.1, SD 5.83 kg/m^2), who used the web application as intended for 4 days and provided feedback at the end of this period.

Results: Iterative storyboard development resulted in programming specifications for a prototype of the web application. We used content feedback to select and refine the support messages and peer profiles to be added. The following 2 rounds of internal testing identified bugs and other problems regarding the web application's functioning and full data collection procedure. Problems were addressed or logged for future consideration. Naturalistic functionality testing revealed minimal further problems; findings showed preliminary acceptability of the web application and suggested that women may select different social content across days.

Conclusions: A multidisciplinary and user-centered design approach led to a personalized and adaptive web application, tailored to the social needs of women in midlife. Findings from testing with this population demonstrated the feasibility and acceptability of the new application and supported further development toward its use in daily life. We describe several potential uses of the

web application and next steps for its development. We also discuss the lessons learned and offer recommendations for future collaborations between behavioral and computer scientists to develop similar tools.

(JMIR Form Res 2022;6(7):e36280) doi:10.2196/36280

KEYWORDS

user-centered design; digital health; eHealth; women's health; midlife; physical activity; social support; social comparison; mobile phone

Introduction

Background

Cardiovascular disease (CVD) remains the leading cause of mortality worldwide, and its public health burden continues to grow [1-3]. Risk for CVD increases with age, with the increased prevalence of risk conditions such as type 2 diabetes and hypertension [4]. Women in midlife (aged 40-60 years [5]) with these conditions are particularly vulnerable; their risk is already heightened by the processes of aging [6], associated weight gain [7] and menopause transition [4,8]. Engaging in regular physical activity (PA) substantially reduces their risk for CVD [9] and has many other physical and mental health benefits for women [10]. However, the gender disparity in PA engagement that favors men across the life span widens during midlife [11], and less than half of the women in the United States in this age range meet US national PA guidelines [12]. Consequently, promoting PA among women in midlife with elevated CVD risk has received considerable empirical and clinical attention as a key pathway for improving cardiovascular health in this population [4,13].

Behavior change techniques such as goal setting, intention formation, and self-monitoring show positive effects on PA adoption and maintenance, at least in the short term (ie, 6 months to 1 year [14,15]). However, these techniques are not widely effective for increasing PA among women in midlife [16]. In addition to widespread challenges such as low motivation [17], many women in midlife face barriers to prioritizing their own health in daily life. Women in this group often serve a variety of caretaking roles in addition to other personal and professional responsibilities [18], and many perceive that taking time for their own PA is selfish [19]. Thus, there is a clear need for novel, low-burden PA resources that are tailored specifically for this population. This paper describes a user-centered design (UCD) approach to address the PA needs and preferences of women in midlife, with a tailored digital tool.

Social Influences on PA Among Women in Midlife

Great attention to *social influences on PA* in interventions may be particularly useful for women in midlife, as women in this age range often cite lack of social support for PA as a key barrier to engagement [19]. Social support for behavior change is a multifaceted construct that involves various types of behavior or communication from others. For example, informational support involves providing resources or suggestions (eg, reasons or ways to change a behavior), whereas emotional support involves demonstrating compassion (eg, encouragement or validation). Individuals differ in the type or types of support they prefer and find helpful [20], and their preferences and responses may fluctuate across days or weeks (eg, based on mood or perceived barriers [21]). A simple way to provide support in digital interventions for health behavior change is via text-based messages, which can be assigned by the system or self-selected by users to meet their immediate needs. Common categories of support in such interventions include information (ie, advice and tips for ways to change health behaviors), encouragement (ie, motivational messages to reinforce and promote health behavior efforts or success), and accountability (ie, tips for increasing commitment to health behavior goals or plans [22]).

Similarly, social comparisons, or self-evaluations relative to others [23], show influence on PA among women in midlife. This group expresses desire for and positive subjective response to PA role models, particularly those who are similar to them (eg, regarding demographic background, caregiving responsibilities, and workload [24,25]). A range of PA-based comparisons can have positive effects. Comparing one's own PA behavior with that of someone who engages in more PA (eg, steps or minutes of aerobic exercise per day), or *upward comparison*, can provide inspiration, motivation, and guidance to achieve a similar outcome [26]. Comparing one's own PA behavior with that of someone with the same level of PA engagement, or *lateral comparison*, can provide assurance of *keeping up* to a relevant standard [27]. Comparing one's own PA behavior with that of someone who engages in less PA, or *downward comparison*, can prompt satisfaction with one's own PA behavior and motivation to continue PA efforts to stay ahead of others [28]. Importantly, both upward and downward comparisons can have the opposite effects: upward comparisons can lead to dejection and disappointment (particularly when there is a very large or consistent discrepancy between one's own and the target's PA [29,30]), and downward comparisons can confirm that one's own PA situation is bad (or likely to become so [31]), and lead to decreased PA. Lateral comparisons are least likely to prompt negative responses; however, the range of their effects is currently unclear. The consequences of all 3 types of comparisons for PA also differ between people and fluctuate within the same person over time [32-36].

Some existing studies have attempted to activate social support and comparison processes during group interventions to promote PA among women in midlife [37,38]. Others have matched intervention participants with one another to create PA partnerships that foster support and comparison, with matching based on demographic characteristics [39,40]. Findings from these studies suggest that women in midlife will engage in discussions that offer mutual support for PA between participants and will respond to social comparison opportunities (eg, by indicating that they are motivated to keep up with others

in the program or use them as role models). Importantly, however, a subset of these studies show that women differ in their responses to opportunities for support or comparison and that these opportunities can result in decreased PA motivation or behavior [39,40]. A reason for this heterogeneity in response is that existing approaches do little to ensure that opportunities for support or comparison are those that match individual preferences. Furthermore, users' preferences may not be most effective for fostering PA adoption or maintenance [41]. In addition to facilitating on-demand PA resources, digital tools that allow self-selection of support and comparison prompts can provide useful information about alignment between preferences and needs and inform improvements to intervention approaches (for women in midlife and more broadly).

Digital PA Support: Personalization, Adaptation, and UCD

Historically, evidence-based interventions to promote health behaviors such as PA have been fixed, such that all participants receive the same intervention content over time [42]. Digital tools have the potential to account for individual differences in participants' responses to intervention content. *Personalization*, or the use of specific content for its alignment with individual characteristics of the user, is intended to increase users' attention by signaling that information is meant specifically for them. Personalized content is expected to be more relevant than generic content, and thus, increase the power of communication [43]. Common personalization strategies in digital PA interventions include individualized feedback, such as customizing messages to users based on their baseline PA level, and user targeting, which adapts content based on user demographics (eg, gender and BMI [44]).

However, even personalized interventions can be insufficient, as they do not account for changes or natural fluctuations in users' needs and preferences. This is a key issue for digital behavior change tools; personalization may only address who a user was in the past, rather than who they currently are, or may not meet a user's needs in the moment [45]. *Adaptation* is a more advanced form of personalization, whereby application content is adjusted based on recent or immediate context [44]. This ensures that digital tools communicate information to users as it is currently relevant, rather than relying on data collected when the user adopts the tool initially. For example, a digital tool may account for a user's recent daily steps when delivering feedback in an intervention for PA, so that the feedback is relevant to that day's PA behavior. Digital platforms may be most effective when they are both personalized and adaptive [46,47]. As noted, there are individual differences and time-related and context-related variations in individuals' preferences for and behavioral responses to social support and comparison, both generally [36,48,49] and specifically among women in midlife who have elevated CVD risk [32,50].

Effectively personalizing and adapting digital platforms require a detailed understanding of the target population. *UCD*, which encompasses several principles and methods aimed at deeply understanding intended end users [51], is widely used in industry settings (eg, product or service development) and is steadily gaining attention in academic research as the demand for engaging digital health platforms and services grows. UCD is iterative—new insights are used to drive and refine development throughout the process (eg, user testing informs design modifications). UCD methods have been used to develop digital health tools for many health behaviors and outcomes, including weight control and binge eating (via digital diaries and personas [52]), medication adherence (via interviews, needs assessments, and focus groups [53]), and chronic obstructive pulmonary disease (via background analysis, prototyping, and usability testing [54]). With respect to digital PA tools, UCD methods have been used to personalize apps for walking among breast cancer survivors (via needs assessment, prototyping, and usability testing [44]), sitting time reduction among office workers (via usability testing and user interviews [55]), and leisure time PA among adults living in close proximity to parks (via needs assessment and user evaluation [56]). Few existing studies, if any, have used UCD principles to develop resources for women in midlife, and no existing study has developed a digital PA tool specifically for this population.

Aims of This Study

There is a clear need for more effective PA promotion resources for women in midlife, and digital tools are appealing because of their flexibility and ease of access. To address this need, our team of behavioral and computer scientists developed a personalized and adaptive web application that is tailored to the needs of this population. This tool is accessed via a web browser and allows users to self-select PA-related social support and social comparison opportunities, which are intended to bolster PA self-efficacy and motivation. Other behavior change techniques activated by the web application include PA self-monitoring, goal setting, and intention formation or planning [26,57]. Social support is provided via short, text-based messages, and comparisons are prompted via exposure to the profile of a peer. As described in the following sections, personalization occurs with respect to the user's age and racial and ethnic identification, and the application adapts to the user's level of PA behavior at each use. Of note, the web application is designed for brief, repeated use—specifically, 10 minutes per day for multiple consecutive days—on days when users wear a PA monitor (eg, pedometer or Fitbit). This allows for the assessment of variability in selections and responses, including subjective perception of selected content and objective PA behavior on a given day.

The overarching goal of this paper was to describe the UCD process of developing and conducting the initial evaluation of the web application prototype, including the results of feasibility, acceptability, and functionality testing with real end users from the target population of women in midlife with elevated CVD risk. We also note some of the lessons learned and present recommendations for future collaborations between behavioral and computer scientists to develop similar digital tools. This study was conducted in four stages: (1) initial application development, (2) feedback on the potential content of the web application (ie, content of peer profiles and messages) from the population of interest, (3) internal functionality testing, and (4) functionality testing with a different set of participants from the population of interest. Table 1 provides a summary of the goals, users, and methods for each stage.

Table 1. Summary of the 4 stages of web application development and testing.

Stage	Goal	Users	Methods
Initial development	Final storyboard and specifications to guide prototype programming	N/A[a]	Discussions between investigators, iterative storyboard generation, and feedback
Content feedback	Selection of profile images, formats, and messages in response to end users' input	A total of 10 women in midlife, with ≥1 CVD[b] risk conditions	In total, 5 qualitative interviews and 5 survey responses
Internal testing			
Round 1	Identification of bugs or other problems to address with respect to web application functioning	A total of 9 behavioral science trainees (clinical psychology and health behavior) who were familiar with the web application	Totally, 9 sets of positive and negative testing
Round 2	Identification of bugs or other problems to address with respect to full data collection procedure	A total of 3 behavioral science trainees (clinical psychology) who were not familiar with the web application	In total, 3 in vivo tests, each over 3 days (morning web application use, pedometer wear, and end-of-day survey)
Naturalistic functionality testing	Identification of bugs or other problems to address with respect to full data collection procedure and obtaining end-user feedback	A total of 5 women in midlife, with ≥1 CVD risk conditions	In total, 5 in vivo tests, each over 4 days (morning web application use, pedometer wear, and end-of-day survey)

[a]N/A: not applicable.

[b]CVD: cardiovascular disease.

Methods

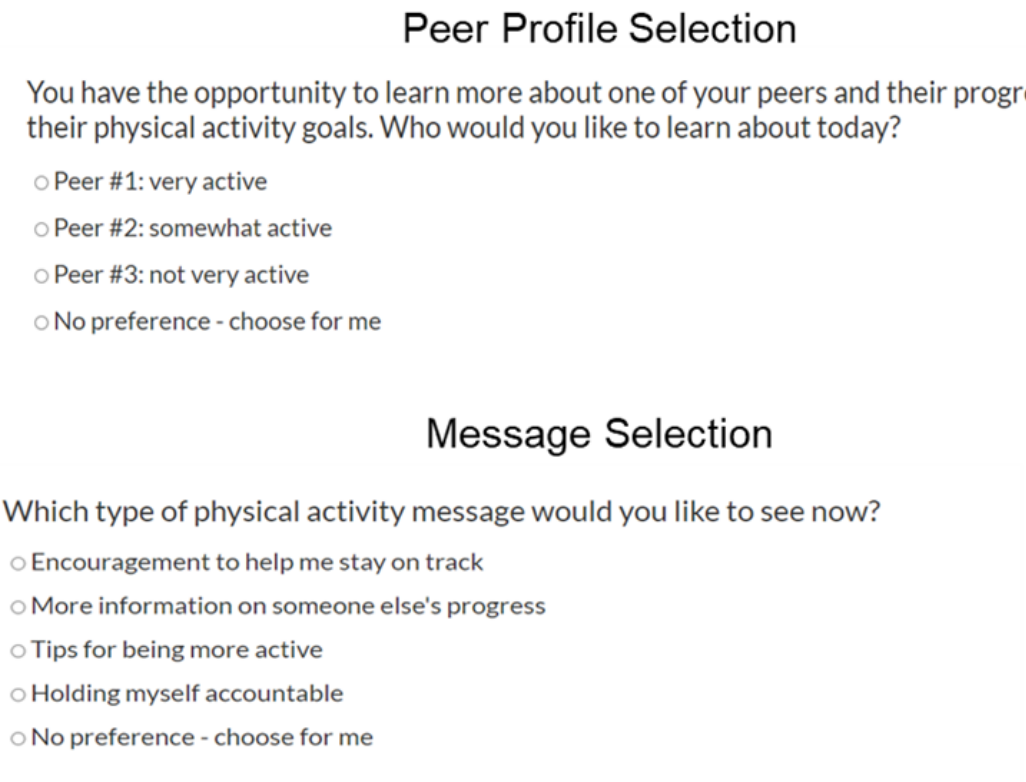
Description of Web Application Content and Use Flow

Each user was provided a study identifier and unique link to the web application. Their researcher-created profile included their age and racial or ethnic identification, for personalization (see next paragraph); the web application does not collect any identifying information during use. The initial page of the application reminds users about the content they will encounter and instructions for optimal use of the application (eg, allocate approximately 10 minutes to engage with it via a web browser each morning). Then, the users are guided to report their step and active minute goals for the day from their PA monitors and reflect on their satisfaction with their PA behavior from the previous day.

Next, social comparisons are prompted by exposure to a peer profile, which shows an image of a fictitious peer and brief text with this peer's characteristics (ie, another woman who appears to be between 40 and 60 years of age, who is interested in increasing her PA; refer to Figure 1 for an example). Each profile text lists the peer's average number of steps and active minutes per week, their favorite way to be physically active, their biggest challenge to being active, and their social context (ie, work situation, family, and caregiving responsibilities). Users can select one of four options: a highly active peer (upward target), a moderately active peer (lateral target), a not-so-active peer (downward target), or *no preference*. In the latter case, the system randomly selects an upward, lateral, or downward target with equal probability.

For the first peer shown at each daily use, personalization is achieved by showing images of peers whose age and racial or ethnic background align with those of the user; users who do not provide a racial or ethnic identification are indicated as race or ethnicity unspecified and view a peer image whose identification is randomly selected at each use. Each image is tagged with 2 relevant age ranges and 2 relevant racial or ethnic identifiers. Within these specifications, the application randomly selects an image and peer text from a database for each profile, ensuring that a user is not shown the same image or text on 2 consecutive profiles. Adaptation occurs by anchoring the peers' average daily steps and active minutes to the user's reported PA behavior from the previous day, and peers' PA metrics are updated every day in response to changes in the user's PA behavior. Specifically, highly active peers' steps and active minutes are 130% of that of the user from the previous day, moderately active peers' PA is 95% to 105% of that of the user from the previous day, and not-very-active peers' PA is 68% of that of the user from the previous day. If PA behavior data are unavailable for the previous day, these adaptations are made using prespecified ranges. Users review the provided peer profile and then respond to a series of questions about their perceptions of the peer described.

Figure 1. Web application screens that facilitate selection for peer profiles (social comparison targets) and messages (social support sources).



Following peer profile reflections, social support is provided via exposure to a brief support message. Users are able to select one of four options for message category: encouragement, tips (ie, suggestions for ways to be more active and information about the benefits of activity for women), accountability (ie, suggestions for holding yourself accountable to your PA goals), or *no preference*. In the latter case, the system selects the user's least-recently-used message category. Then, the system randomly selects encouragement, tips, or accountability message from a database (Figure 1). The system ensures that a user is not shown the same text on 2 consecutive messages. Users review the provided message and respond to a series of questions about their perceptions of the message, including how helpful they found it to be.

After completing these items, users are asked whether they would like to see a third message (yes or no). Those who select *no* are directed to the last set of questions for the day, which prompts PA intention formation via text entry. Those who select *yes* are offered a new set of options: encouragement, tips, accountability, another peer, or no preference. Users who attempt to select the same message category twice in the same use episode are prompted to *keep things interesting—select another type of message*, and the system does not allow them to move forward until they select a different category. Users who select to view a second peer for the day are not able to choose the type; they are randomized to a profile using the rules described previously for a *no preference* selection. After viewing the third message, the user responds to a series of questions about their perceptions and then is prompted to form PA intentions or plans via text entry, as described previously. Textbox 1 lists all the behavior change techniques activated by web application use.

Textbox 1. Physical activity (PA) behavior change techniques activated by the new web application.

Self-monitoring

- Use of pedometer, which is worn on days of web application use
- Reflection on the previous day's PA progress

Goal setting

- Identification of step and active minute goals for the day

Social support

- Selection of support message type to support PA motivation and behavior
- Reflection on message content and application

Social comparison

- Selection of comparison type (peer profile) to support PA motivation and behavior, personalized to match user's age and racial or ethnic identification and based on user's recent PA behavior (adapted to steps and active minutes)
- Reflection on response

Planning or intention formation

- Open-ended response to the following prompt: "Please describe your plan for reaching your activity goals today. How will you get your steps or active minutes in today?"

Recruitment and Participants

Initial Application Development

PhD-level researchers in clinical health psychology (DA) and computer science (AFL) met regularly throughout 2019 and 2020 to develop the purpose, functions, and technical specifications of the web application (described in detail in the *Procedures* section). Both investigators were women; 50% (1/2) of them identified as White and 50% (1/2) identified as Latina.

Content Feedback From the Population of Interest

Women from the target population were recruited for 90-minute feedback interviews conducted via Zoom (Zoom Technologies, Inc). Eligibility criteria required that women be aged between 40 and 60 years (inclusive), with ≥1 risk conditions for CVD (ie, hypertension or prehypertension, prediabetes or type 2 diabetes, hyperlipidemia or hypercholesterolemia, metabolic syndrome, current smoker, or quit smoking in the past 3 months); that they were not currently pregnant; and that they had access to Zoom for the interview. Recruitment was conducted via electronic advertisements posted to the supporting institution's announcement service and social media sites such as Twitter. Interviews were conducted in January 2021. Of the 6 women who were initially contacted about their interest, 5 (83%) enrolled and completed the interview. The average participant's age was 47.4 (SD 6.66) years and BMI was 35.3 (SD 9.55) kg/m^2. As all but one of these participants (4/5, 80%) identified as White, 5 additional participants were recruited to provide feedback via electronic survey, to ensure representation of perspectives of women from a range of backgrounds. For this group, the average participant's age was 51 (SD 6.60) years and their BMI was 32.7 (SD 8.39) kg/m^2. Of the 5 participants who completed the survey, 2 (40%) identified as Black, 2 (40%) identified as Latina, and 1 (20%) identified as East Asian.

Internal Functionality Testing

Investigators familiarized their team members (ie, not users from the target population) with the application's purpose and previous stages of development. Then, the entire team conducted internal testing to confirm that the application functioned as intended. A total of 9 testers were included in the first round, of which 3 (33%) were doctoral students, 4 (44%) were undergraduate students in psychology, 1 (11%) was PhD-level program manager, and 1 (11%) was PhD-level investigator in clinical health psychology. All of them have backgrounds in digital health research. Of these 9 testers, 5 (56%) were women and 7 (78%) identified as White. In the second round, testers were 3 doctoral students in clinical psychology, without familiarity with the web application. All of them (3/3, 100%) were women aged 25 to 30 years; 67% (2/3) of them identified as White and 33% (1/3) identified as Black.

Feasibility, Acceptability, and Functionality Testing With the Population of Interest

Women were recruited using electronic advertisements posted to the supporting institution's announcement service and social media sites (eg, Twitter), to engage in 4-day naturalistic functionality testing. In addition to the eligibility criteria described for content feedback, eligibility required that women did not have an active injury or illness that impeded their PA. A total of 11 women expressed interest in participating, and 5 (45%) women enrolled and completed these tests in June 2021. The average participants' mean age was 50 (SD 6.26) years and BMI was 30.1 (SD 5.83) kg/m^2. Of the 5 participants, 2 (40%) identified as Black, 1 (20%) identified as Latina, and 2 (40%) identified as White. None of these women (0/5, 0%) participated in the earlier phases of this study and thus, were naïve to the content of the web application. These users tested the application each morning for 4 consecutive days. They also wore either their personal PA monitors or a study-provided pedometer,

completed end-of-day surveys on each day, and engaged in a brief exit interview at the end of data collection.

Measures

Content Ratings (Initial Feedback)

Qualitative responses were noted for participants' impressions of peers' ages, racial or ethnic backgrounds, and professional circumstances for each image shown and for preferences regarding peer profile content. Participants who engaged in the initial interviews were asked to rate each PA message on a scale of 1 (*not at all*) to 10 (*extremely*), regarding (1) how much they liked each message and (2) how helpful they thought it would be for supporting their PA behavior. Participants were asked to briefly articulate their reasons for these ratings.

PA Monitors and End-of-Day Surveys (Second Phase of Internal Testing as Well as Feasibility and Naturalistic Functionality Testing)

Women who participated in the second phase of internal testing (3 trainees; 3 days each) and in feasibility/naturalistic functionality testing (5 women in midlife; 4 days each) wore a PA monitoring device during each test day. These participants used either their own personal device (eg, Fitbit or Apple Watch; worn on the wrist) or a study-provided Accusplit AX2720MV pedometer (worn on the waist or hip). These devices captured steps and active minutes (ie, high-intensity activity) throughout the day, and participants reported their totals from these devices at the start and end of each day. Participants reported their starting PA at the start of web application use each morning and daily totals in end-of-day surveys (sent via SMS text message or email).

Web Application Use and Technical Problems (Second Phase of Internal Testing as Well as Feasibility and Naturalistic Functionality Testing)

We determined the feasibility of web application use as intended by the percentage of times the application was accessed successfully and returned no errors (or other noteworthy problems), relative to the number of expected uses. Additional relevant data were related to devices and browsers used to access the web application, nature and replicability of reported errors, and application navigation information (eg, peer and message category selections).

Exit Interview (Feasibility/Acceptability/Naturalistic Functionality Testing Only)

Women in midlife who completed feasibility/naturalistic functionality testing engaged in a brief exit interview after the end of their tests, conducted via Zoom (15-30 minutes). Participants were asked about their overall experience with the web application, each component of the application, and whether they would be interested in future use. Given the complexity and lack of consensus regarding measuring acceptability in digital health research [58], responses in these domains were used to informally summarize participants' views of acceptability (as described in the following sections).

Procedures

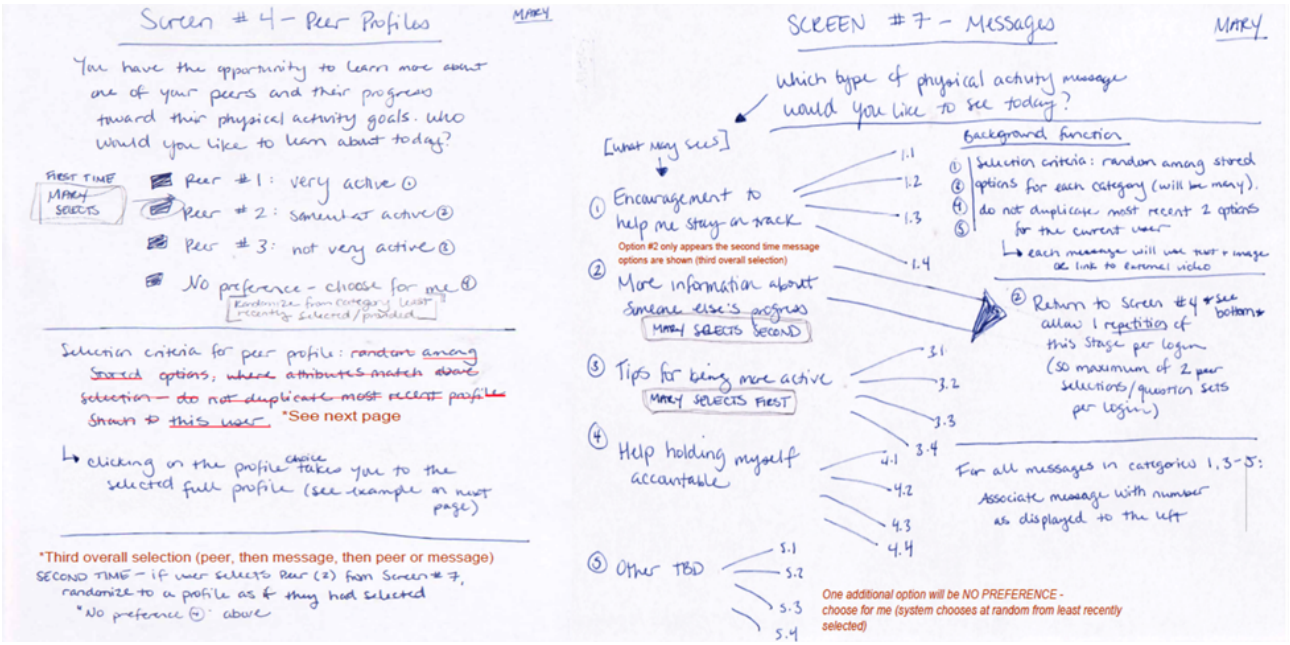
Initial Development

Collaboration began in 2018, with support from internal and external funding provided to the first author (DA). Investigators met regularly in 2019 and 2020 to discuss the goals of a new tool, how it may be tailored to the needs of women in midlife with elevated CVD risk, and the process of identifying technical specifications. Although we considered a stand-alone mobile app as the outcome, we selected a web application owing to its versatility and ease of access across devices and platforms. During this time, the first author built on her previous studies with the population of interest [37,39,59] by conducting additional observational and qualitative studies [32,33,60,61] and collecting preliminary data to assess their needs and preferences. Insights from previous studies, discussions between investigators, and new participant responses informed the refinement of the goals of the new tool.

In the spring and summer of 2020, investigators developed storyboards to specify the requirements of the web application. We developed a total of 7 storyboard drafts using an interactive process, which was led by the first author (DA) and modified with questions and input from the computer science expert (second author [AFL]). Figure 2 shows screenshots of the final storyboard, illustrating some of the important steps of the user experience with the web application. This resource included 37 sheets describing the intended progress through 15 pages of the application, including descriptions of progress for 2 fictional users and skeleton of a database that would house participants' selections and responses as they used the application. The storyboarding process led to the creation of a penultimate set of technical specifications, which the computer scientist (AFL) used to develop the code for the initial draft of the application.

Development of the application occurred during the summer and fall of 2020. We chose the Django framework for its integration of front-end and back-end development and its secure superuser view that allows team members to edit the database entries for peer texts and support messages. The application was deployed on an Amazon Web Services ec2 instance using the SQLite database. During this time, the first author (DA) and trainees in behavioral science created a preliminary set of peer profiles and support messages. This included images, profile text, and message text. Images of women who appeared to be aged between 40 and 60 years were selected from open-access web-based databases [62] to represent a range of racial or ethnic and professional backgrounds. Profile text was generated based on the investigators' knowledge of and experience with prompting PA-based social comparisons, creating 5 versions of the same 2 profiles that offered different levels of detail. Finally, 18 support messages (6 each for encouragement, tips, and accountability) were generated based on those used in previous studies [63-65], with modifications to address the needs of women in midlife and attention to the possibility of limited access to fitness facilities during the COVID-19 pandemic (eg, exercises that are easy to do at home, without equipment).

Figure 2. Pages from the final version of the web application storyboard, describing content and function for selection of peer profiles (social comparison targets) and messages (social support sources). Mary refers to the selections of a fictitious user, to indicate the expected flow.

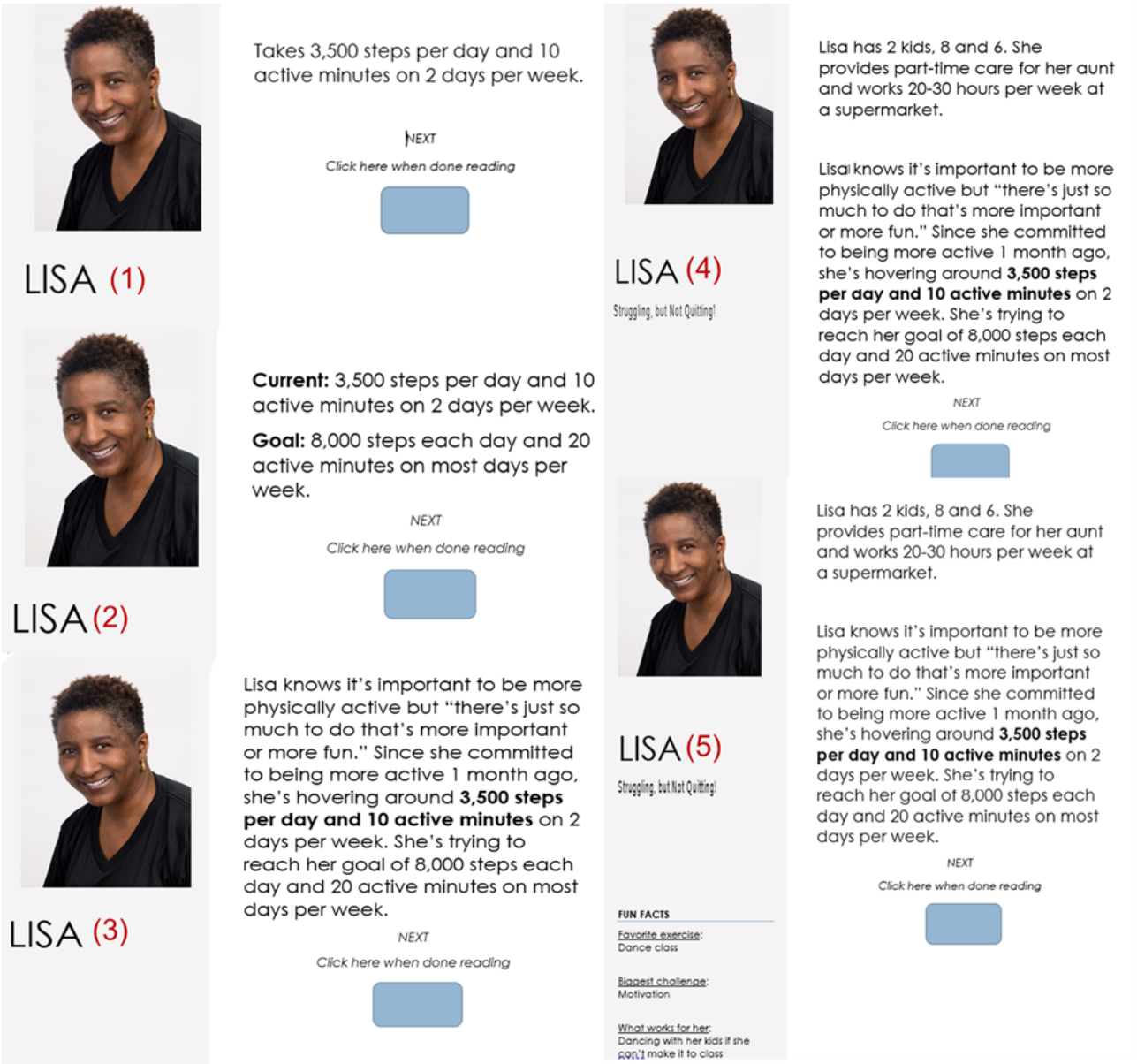


Content Feedback With the Population of Interest

Procedures were approved by the supporting institutional review boards (IRBs), and the participants provided written informed consent via Adobe electronic signature (5/10, 50%). Participants met with the first author (DA) for individual, 90-minute interviews via Zoom. These interviews focused on gathering women's perceptions, preferences, and suggestions for improvements to sample peer images, profile text, and messages. After initial overview of the session and introduction to the purpose of the web application, participants were asked to view a series of peer images and indicate their perceptions of the woman's age, racial or ethnic identification, and profession. Then, they viewed 5 versions of each of the 2 peer profiles and were asked to verbally evaluate each version (ie, strong vs weak preference for each, relative to the previous version, and their reasons). These profiles showed minimal information initially and increased in the number of facts presented in each version (Figure 3). Next, participants were asked to read each of the 18 messages and rate each one on the scales described previously. Finally, they were asked to rate their interest in using the web application to support their PA efforts in the future on a scale of 1 (not at all interested) to 5 (extremely interested). Women who provided feedback via electronic survey (5/10, 50%) provided informed consent by clicking a button to verify that they consented to proceeding with the survey; they were given a brief explanation of the web application and asked to respond to a subset of the stimuli described previously.

The interviewer (first author [DA]) took detailed notes during the Zoom feedback sessions to capture numeric ratings, narrative explanations for these ratings, and other suggestions from participants. Then, the interviewer and another member of the behavioral science team reviewed and synthesized these notes and responses from the participants who completed the survey. They identified common themes and recommendations for changes, which were implemented as described in the following sections. A third member of the behavioral science team was consulted for input on messages that received a wide range of responses, and disagreements about changes were resolved by consensus.

Figure 3. Versions of possible peer profiles rated by end users during content feedback stage.



Internal Functionality Testing

In the first phase, research team members were randomly assigned to user accounts representing women in the target age range, from varying racial or ethnic backgrounds, and asked to conduct a series of tests of the web application. First, they completed a round of positive testing (using valid data entry) to verify that the application worked as expected. Then, they completed a round of negative testing (using invalid or improper data entry) to check for unexpected conditions. Each user conducted between 5 and 20 tests over the course of 4 months, with application updates occurring between each round of tests (February 2021 to May 2021). A subset of team members (5/9, 56%) also exported data files for recent use episodes to check for errors in recording the use variables in the database.

Users in the second phase of testing received instructions via email for testing the application between 8 AM and 9 AM each day for 3 consecutive days, wearing their personal PA monitors, and completing end-of-day surveys on each day. Of the 3 users, 1 (33%) user tested as a woman with racial or ethnic identification the same as her own (Black), and 2 (67%) tested as women not matched to their racial or ethnic identification (the White trainee users tested as a Latina woman and as a woman whose identification was unspecified, respectively). End-of-day surveys were sent via SMS text messages at 9 PM each night and were completed via mobile device by 9:45 PM. Before users accessed the web application on the following day, members of the research team entered their step and active minute totals to the application, to enable adaptation of the PA levels shown in peer profiles. The investigators were available during the application test window each morning (8 AM to 9 AM) to receive and address reports of errors and access difficulties.

Feasibility, Acceptability, and Functionality Testing With the Population of Interest

All procedures were approved by the relevant IRBs, and participants provided written informed consent via Adobe electronic signature (5/5, 100%). Each participant met with the

first author (DA) via Zoom for introduction to the web application and study procedures, during which the first author used screen share to show participants an episode of web application use in real time. Participants who requested to use a study pedometer (3/5, 60%) received these devices and printed instructions in the mail before this meeting (with prepaid postage for the return mailing) and were invited to ask questions about intended use for data collection. Next, participants received an overview of the end-of-day survey content and procedures and engaged in a live test of receiving surveys via SMS text messages. The interviewer collected information about participants' wake and sleep times and discussed their preferred time for using the application and receiving end-of-day surveys, according to their wake and sleep times. Specifically, participants were asked to use the web application within 2 hours of waking up and to complete the end-of-day survey within 1 hour before going to sleep. The investigators were available during the application test windows each morning to receive and address any reports of errors or access difficulties. Finally, before participants accessed the web application on the following day, the first author entered the step and active minute totals to the application, to enable adaptation of the PA levels shown in peer profiles. Following the procedures outlined previously, the interviewer (first author) took detailed notes during the Zoom exit interviews to capture participants' views of acceptability. The interviewer and another member of the behavioral science team reviewed and synthesized these notes to generate a summary of the acceptability feedback.

Ethics Approval

The IRB at Rowan University approved this study (PRO-2020-163).

Results

Development

Refinement of the initial idea for the web application occurred over a period of 1 year. This period involved regular discussions among investigators, storyboarding, identifying technical specifications, and designing a database. Investigators created 7 drafts of the web application storyboard using a process of iterative feedback. At each stage, discussions generated questions to be answered and functions to be refined, which would allow for individual user personalization and day-to-day adaptation. The initial version of the storyboard used 8 pages to describe the 9 distinct screens needed to achieve a full-use episode. The final version used 37 pages to specify back-end decision points, 15 distinct screens needed to achieve a full-use episode, and full examples of 2 different users' experiences of a full-use episode. Separate files illustrated the database (where users' selections, entries, time spent in using the application, and end-of-day data are collected and stored), administration pages (for creating users, entering end-of-day data, and downloading database files), and a penultimate set of technical specifications. The software implementation, deployment, testing, and support tasks required approximately 3 months of effort from the computer science expert (second author [AFL]). Finally, we deployed and tested the web application as described previously.

Content Feedback

Gathering feedback on potential web application content occurred during the later stages of programming, over a 4-week period. Participants who engaged in 90-minute interviews via Zoom (5/10, 50%) reported strong overall interest in using the web tool under development to support their PA efforts (mean score 4.5 out of 5). Regarding peer images, there was high agreement regarding perceived age (87%), racial or ethnic background (88%), and types of careers (85%). A total of 2 images were perceived as depicting women aged between 40 and 46 years, 3 images as women aged between 45 and 52 years, and 2 images as women aged between 52 and 60 years. A total of 3 images were perceived as depicting Black women (or multiracial with Black as an identification), 4 as Latina women (or multiethnic with Latina as an identification), 2 as White women (or multiracial with White as an identification), and 1 as Middle Eastern or South Asian (Indian or multiracial with one of these as an identification). A total of 3 images were described as portraying women with administratively focused jobs (eg, office assistants and business managers), 2 as teachers, and 3 as other (eg, coaches and retired). All images were deemed acceptable for the stated purpose, and participants indicated that some variety across images (eg, in types of clothing) may make the peer profiles relatable to a wide range of women in midlife.

A total of 2 images were perceived as *intimidating* by ≥1 participant. The first image was described as projecting a "no nonsense" attitude (however, still moderately welcoming), whereas the second appeared to "have it all" (ie, fit, affluent, and carefree). These happened to be the 2 images used for the next set of prompts related to the amount and types of information that participants perceived as useful for inclusion in peer profiles. For this set of prompts, the participants viewed 5 different versions of each woman's profile, each increasing the amount of information presented. For both profiles, participants indicated that the middle version provided the optimal amount of information (ie, version 3 in Figure 3). Although 60% (3/5) of the participants specifically expressed interest in the addition of *fun facts* in each profile (as these made the profiled women seem more *relatable*), they indicated that the short versions were the best for providing useful information efficiently.

Participants' overall recommendations for PA messages were to keep them short; however, there was considerable heterogeneity in the subjective response to each message. Consider the following message: "We're often told how important physical activity is to our health. But why is it important to you? Remind yourself of YOUR reasons to be more active." This message received a score of 0 from one participant (regarding how much she liked the message and found it helpful); she indicated that the message implied that PA "must not be important if I'm not doing it," which was not motivating. However, the same message received a score of 10 from another participant (who said it was "thought-provoking and affirming"), and the average ratings for liking and helpfulness were 6.2.

Examples of the highest-rated and lowest-rated messages are listed in Table 2. The lowest-rated messages were removed or

substantially revised; all other messages received minor adjustments to wording or formatting (eg, whether parts of the text were underlined or bulleted), based on participant feedback.

Table 2. Sample messages and ratings and comments by end users, obtained during the content feedback stage.

Messages	Ratings (1-10 scale)		Comments
	Liking, mean (SD); range	Helpfulness, mean (SD); range	
Do you schedule meetings and appointments on your calendar, and usually show up to them? Could you do the same for physical activity? Make it easy – add time to be active on your calendar, and make it a point to show up like you would with any other obligation!	4.1 (1.67); 3.5-7	4.3 (2.11); 3.5-7	• "I agree that this could work but it causes me stress - I have too many appointments already." • "This is a good suggestion, but reword it so that it's not asking so many questions."
Do you have responsibilities such as childcare, or taking care of others you care about? Lack of time or energy due to these responsibilities can make physical activity difficult, but it doesn't mean that you can't be more physically active. Try going for walks, light stretching, or standing while watching TV. You can do all this with your loved ones, or try it during short windows when you have a moment to yourself.	5.4 (2.97); 2-10	5.3 (1.99); 2-7	• "This one is really long and it's not personal - seems too general." • "Too long but it's a good reminder."
You don't have to work on your physical activity goals alone! Support from family and friends can help you to achieve your goals. Let someone close to you know about your goals, so they can provide you with support and help hold you to them!	8 (2.12); 5-10	7.8 (1.92); 5-10	• "This is good advice. It helps hold you accountable to tell someone else." • "If I were going to do this I would have done it already."
When setting physical activity goals each week, try to make them clear and measurable. Instead of a goal of "walking more," try: walking 3 times this week for 10 minutes each time. This will help you stay on task and make it easier to know when you've met your goal. Then you get a sense of accomplishment when you know you've met your goal!	8.3 (1.20); 7-10	8.1 (1.34); 7-10	• "I like this a lot. It's clear and actionable." • "I agree! Great suggestion."

Internal Testing

Round 1

Of the 25 positive tests conducted, 2 (8%) resulted in initial errors, and they were resolved. A further 16% (4/25) of the tests revealed bugs, including failing to offer a second message and repeating the same content twice in 1 use episode. Both of these problems were resolved after the conclusion of round 1 testing. Negative testing (involving 20 use episodes) identified two application navigation pitfalls: using the back button allowed for many selections in the same use episode and entering spaces or other non–language-based text was accepted as valid responses to the open-ended questions. Across both types of tests, several team members also indicated that certain details of peer profiles did not match their intended ages. For example, a subset of photos of women in their late fifties was paired with profiles describing them as having very young children.

At the end of round 1, the investigators consulted and agreed that the application should insist that the user proceed strictly in the order described previously, irrespective of the user's attempts to use their browser's back button. The second author (AFL) made system modifications such that the user is required to complete a day's survey on a single device and browser, and the user's attempt to use their browser's back button redirects them to the next screen. These updates were completed before the functionality testing with end users. The application specifications did not include procedures for preventing nonlanguage text responses to the open-ended questions. In contrast, the peer profiles were intended to be tagged with age indicators (similar to peer images), but this step was skipped in the build process. As the latter two issues were deemed substantial revisions, they were noted and postponed until a future round of substantive updates.

Round 2

Across 3 testers and 3 days of use episodes (total of 9 episodes), 11% (1/9) of the test uses of the web application resulted in an error. The error for the assigned client could not be resolved and the tester was assigned a new client ID. After this substitution, all use episodes were successful, and all end-of-day surveys were completed as intended. Of the 3 testers, 1 (33%) tester noted that she received the same message twice in the same use episode, despite choosing different message categories, and this error was logged for correction. All testers (3/3, 100%) expressed their perception that the web application would be useful for supporting women's PA efforts and indicated that the web application's personalization and adaptation features were effective for tailoring their experience of peer profiles. Testers also offered additional feedback on the wording of peer profiles, messages, and end-of-day surveys to clarify the specific experiences of interest (eg, a missing text box for entering details if the *other* response to a multiple-choice item was selected).

Feasibility, Acceptability, and Naturalistic Functionality Testing

Testing with a different group of women from the target population was conducted over 4 weeks. All the participants (5/5, 100%) were able to access the web application on their personal devices. This included the use of desktop and laptop computers, tablets, and smartphones, across Chrome and Safari browsers. A total of 60% (3/5) of the participants did not encounter any errors or problems with access on any of their days of participation. Of the 5 participants, 1 (20%) participant received an error after starting use (ie, getting to the starting page without entering any data), closing her browser, and attempting to restart later. As the application allows only 1 use per client per day, this difficulty demonstrated correct functionality. The participant was able to access the application on the same day after her first attempt was deleted by the research staff. The final participant received error messages for attempts to access the web application on her smartphone but encountered no problems when she switched to her laptop computer the following day. The initial problem could not be duplicated by the research staff but was noted in case of repeated difficulties. Thus, of the 20 expected use episodes (5 participants; 4 days each), 18 (90%) were completed and resulted in use data without incident and the remaining 2 (10%) were missing, owing to access difficulties described previously. In total, 85% (17/20) of the episodes were informed by previous end-of-day step and active minute totals.

Participants' peer profile and message selections are summarized in Table 3. Across participants and use episodes, initial peer type selections were most frequently lateral or downward targets (6/18, 33% each). Upward targets and *no preference—choose for me* were selected during 17% (3/18) of episodes each, by a different participant each time; *no preference* was selected on different days of use each time, whereas 67% (2/3) of the selections of upward targets were on the last days of use. All of the participants (5/5, 100%) selected at least two different peer options. Of the peers actually displayed, lateral or downward targets were seen in 39% (7/18) of episodes each, and upward targets were seen in 22% (4/18) of episodes. As noted, these directions were relative to participant's PA behavior on the previous day. Initial message type selections were predominantly those providing accountability (7/18, 39%), followed by encouragement (6/18, 33%) and tips (3/18, 17%). *No preference—choose for me* was selected only twice, by a different participant and on a different day of observation each time. A second peer and an encouragement message were chosen twice each; a tip message and *no preference* were selected once.

Table 3. Summary of peer profile and message selection types during end-user functionality testing stage (5 users; 4 days each).

Categories	Selections of episodes (n=18), n (%)
Upward peer	3 (17)
Lateral peer	6 (33)
Downward peer	6 (33)
No preference for peer	3 (17)
Encouragement message	6 (33)
Tips message	3 (17)
Accountability message	7 (39)
No preference for message	2 (11)

During the exit interviews, all participants (5/5, 100%) indicated that their overall experience with the web application was positive; 80% (4/5) of them indicated that they found the application very helpful in setting a positive tone for their PA for the day and reminding them to be more active. For example, a participant stated that she noticed herself walking her dog more often as a way to increase her PA. The final participant reported that she found the web application acceptable. As she was previously very active and used several of the application's suggested approaches at that time, the content was "not new" to her. However, she reported an expectation that it would be useful for someone just starting an effort to be more active. All participants (5/5, 100%) indicated that they would be interested in using an expanded version of the web application to support their ongoing PA efforts, with 60% (3/5) of them specifying that use would be most beneficial in conjunction with PA coaching (ie, additional, formal guidance to build behavioral and psychological skills).

Discussion

Principal Findings

Women in midlife—particularly those who have health conditions that exacerbate their increasing risk for CVD—would benefit from PA-specific resources that address their social preferences and needs. This paper describes a multidisciplinary, iterative, UCD approach for developing a personalized, adaptive web application for this purpose. Our approach involved four stages: initial design and building, gathering content feedback, internal testing, and naturalistic functionality testing with end users. Changes were made to the application's content, function, and navigation in response to feedback at each stage, which resulted in a prototype that showed evidence of feasibility and acceptability among the target population.

Our initial findings also suggest the possibility of heterogeneity in the PA preferences and experiences of women in midlife. Some social content was consistently rated as favorable, whereas other content generated wide ranges of favorability, and women

showed noteworthy variability in their support message and peer (comparison target) selections in their daily lives. Some of this variability may be due to curiosity and desire to explore the available options [66,67]. The content that women in midlife want or believe will be most helpful also likely varies from day to day (eg, with mood and other contextual factors). An important next step is to conduct studies with long test periods and large samples to determine the extent of stability versus ongoing variability in women's selections and responses. Incorporating UCD principles and methods throughout the development process provided insight into the needs and preferences of our target population. By taking an iterative approach, we identified issues early, which allowed for refinement before further testing. This process was made easy by regular, planned discussions between the behavioral science and computer science teams and ensured a shared vision and scope for the project.

Uses and Future Directions for a New PA Web Tool

There are multiple potential uses for our new web application, and our future studies will investigate the utility of the tool for each purpose. First, as suggested by members of the target population, this tool may be used as an adjunct to and extension of more traditional PA coaching. Individual or group-based coaching, which often occurs via live interaction (in person or through phone calls or video calls), can provide a range of psychological and behavioral skills for adopting or maintaining regular PA routines (eg, intention formation and planning [68]). The web application, available daily, can supplement these interactions by providing content that reinforces skills and guides actions on demand in women's daily lives. Similarly, the web application can serve as a low-intensity introduction to some basic PA skills (eg, goal setting) and help women to determine the types of support or behaviors that may work best for their goals. Given its low intensity, the web application is unlikely to be widely effective as a stand-alone intervention for PA adoption. However, there may be a subset of women in midlife who benefit from the minimal, noninvasive, personalized, and adaptive support that it offers, either for initial PA adoption or for reinitiation of previous PA routines (eg, as women recover from PA changes owing to the COVID-19 pandemic [69,70]).

We also see potential for more innovative uses of the web application, beyond traditional PA support. For example, as noted, what users want and what works for prompting behavior change in daily life do not always align [41]. When combined with objective PA monitoring, data from the new tool can facilitate the assessment of alignment between a user's social PA content preference (selection), their subjective response to this content (liking or perceived helpfulness), and their daily PA behavior. This can be informative in 2 ways. First, from a basic science perspective, such assessment on a large scale will indicate the prevalence and distinct types of misalignments. This can inform our understanding of phenomena such as cognitive biases that interfere with PA, including strong preferences for and liking of content that is self-enhancing (eg, downward comparison targets [71]), despite the behavioral benefit of exposure to less-enhancing contact, such as upward comparison targets [72]. Second, from an intervention perspective, information about any observed misalignment of preferences and behavioral responses can be offered as feedback (eg, from PA coaches). This can enable women to gain insight into patterns that facilitate or inhibit their PA behavior change.

A final potential use of the new web application is for assessing the alignment of preferences and responses between users, which also has implications for PA interventions. Specifically, the procedures described in the naturalistic functionality testing phase can be used to generate a profile or type for each user. Content selections and both subjective and behavioral responses can determine the *optimal* social support and comparison exposures for promoting a user's engagement with the tool and PA behavior [45]. At the same time, responses to survey items and PA behavior data can determine the type or types of support and comparison targets that the user would offer to other women in midlife. Then, this profile of optimal social input and likely output can be matched with a complementary profile to create PA partner dyads that are likely to meet each other's social needs. This process can maximize the power of the social environment of PA interventions to promote PA among women in midlife.

The present set of studies represents only the formative steps in a long process to fully develop and test a new digital health tool. Consistent with many of the existing UCD studies, we used both quantitative and qualitative methods with modest sample sizes [73]. Each of these potential uses of the web application will require the noted substantive revisions to the tool (which are currently underway) and large-scale testing with large and more diverse samples of women in midlife.

Lessons Learned and Recommendations for Future Digital Health Collaborations

The present series of studies and the creation of a new digital health tool resulted from close collaboration between behavioral and computer scientists. This type of collaboration is becoming more common and desirable to ensure that the development of a tool to support health behavior change is informed by a range of relevant expertise [74]. Such partnerships can be extremely fruitful but are not without challenges. For example, the fields of behavioral and computer science prioritize different methods and often share little common language. In this study, although we developed the storyboard and specifications for the web application over several months and a series of iterations, potentially important navigation rules were left unspecified (ie, use of the web browser's back button, entry of spaces, or nonlanguage text as open-ended responses). These and similar oversights may be owing to behavioral scientists' implicit assumptions and lack of familiarity with the problems that such errors can cause, and thus, failing to specify all desired rules ahead of time. To avoid the need for major changes after the initial tool is built—particularly if funding for such changes is limited or uncertain—there is additional pressure on computer science team members to foresee all types of possible pitfalls. In an academic setting, where funding for the development of digital tools may be scarce, early discussions about how and when substantive, unforeseen changes will be handled are critical.

Behavioral scientists also tend to focus, from the beginning, on a process of human subjects research that involves regulatory oversight, informed consent discussions, certain types of documentation (eg, test user feedback), and multiple attempts to secure additional (limited) funding. This process is intentionally methodical and can be quite slow as a result [74]. Computer scientists may not be familiar with the IRB processes, training required to participate in studies with human participants, and delays that these may cause. Computer scientists may prefer agile and iterative software development methods with multiple rounds of user feedback, but may not realize, at the beginning, the time and effort required to obtain IRB approvals and recruit participants. For this collaboration, although we were aware of the potential regulatory challenges from the beginning, we encountered difficulties regarding questions about permissions for server housing and IRB questions about the collection of protected health information via the new web application. The web application was designed to protect users' identities and does not collect protected health information; however, the process of effectively communicating this to the IRB and securing permission to house the web application on a commercial server was lengthy and involved. This was one of the several delays in the large process of web application development and testing that contributed to the extension of the project time line. The onset of the COVID-19 pandemic was the cause of another delay, as part of the investigators' research time was redirected to other professional and personal responsibilities [75].

Strengths, Limitations, and Conclusions

Despite these delays, this study had several strengths, including a base of foundational work with the population of interest [32,37,60,61,69], adherence to widely recommended UCD principles [51,73], emphasis on recruiting women from racial or ethnic minority backgrounds to maximize diversity in small samples, and collaboration between flexible researchers who are committed to the project. This process also benefited from multiple phases of internal testing, including a range of testers who were familiar with and those who were naïve to the web application specifications and phases of both positive and negative testing. There were also noteworthy limitations to our study. For example, all data collection processes and feedback sessions with women in midlife were conducted by the first author (DA), who is the principal investigator of the project. This was intentional to ensure consistency in interview style and optimal synthesis of findings across study stages, particularly for the small samples recruited and the resources available. However, knowing that they were speaking to the project director may have heightened the participants' social desirability or self-presentation behaviors. The first author's direct involvement can also enable confirmation bias regarding the conclusions drawn (eg, from interview feedback). Interview scripts were designed to probe dislikes or nonpreferred content to limit this concern, and additional team members were consulted to provide an additional perspective, but it remains a possibility. This may be particularly relevant for the assessment of acceptability, which was performed informally in this study. In future studies, it would be ideal to record interviews and have multiple raters use a standard set of guidelines [76].

However, it is important to note that these procedures may not be possible at formative study stages owing to limited financial support to protect team members' time for such activities. Relative to the resources available for later stages of the process, such as efficacy testing for an existing digital health tool, research funding for the formative stages of tool development is often much more modest. Consequently, creative thinking, patience, and persistence are critical to the early success of such a venture. Together, this series of studies provides a useful example of how to approach these formative stages to gain necessary insights from the population of interest and to identify pitfalls and the next steps for future studies.

Acknowledgments

These studies were supported by the National Institutes of Health (NHLBI K23136657; principal investigator: DA). The authors would like to thank Megan M Brown for her assistance with data collection and Laura Travers, Ahmad Hussein, Leah Schumacher, Daija Jackson, Samuel Hart, Heather Mulvenna, Bernard Kwiatek, and Kristen Edwards for their assistance with internal testing.

Conflicts of Interest

None declared.

References

1. Roth GA, Mensah GA, Fuster V. The global burden of cardiovascular diseases and risks: a compass for global action. J Am Coll Cardiol 2020 Dec 22;76(25):2980-2981 [FREE Full text] [doi: 10.1016/j.jacc.2020.11.021] [Medline: 33309174]
2. Cardiovascular diseases (CVDs). World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) [accessed 2022-01-08]
3. GBD 2015 MortalityCauses of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct 08;388(10053):1459-1544 [FREE Full text] [doi: 10.1016/S0140-6736(16)31012-1] [Medline: 27733281]
4. Rodgers JL, Jones J, Bolleddu SI, Vanthenapalli S, Rodgers LE, Shah K, et al. Cardiovascular risks associated with gender and aging. J Cardiovasc Dev Dis 2019 Apr 27;6(2):19 [FREE Full text] [doi: 10.3390/jcdd6020019] [Medline: 31035613]

5. Brim O, Ryff C, Kessler R. How Healthy Are We?: A National Study of Well-Being at Midlife. Chicago, IL: University of Chicago Press; 2003.
6. North BJ, Sinclair DA. The intersection between aging and cardiovascular disease. Circ Res 2012 Apr 13;110(8):1097-1108 [FREE Full text] [doi: 10.1161/CIRCRESAHA.111.246876] [Medline: 22499900]
7. Lovejoy JC. Weight gain in women at midlife: the influence of menopause. Obesity Manag 2009 Feb;5(2):52-56. [doi: 10.1089/obe.2009.0203]
8. Karvinen S, Jergenson MJ, Hyvärinen M, Aukee P, Tammelin T, Sipilä S, et al. Menopausal status and physical activity are independently associated with cardiovascular risk factors of healthy middle-aged women: cross-sectional and longitudinal evidence. Front Endocrinol (Lausanne) 2019;10:589 [FREE Full text] [doi: 10.3389/fendo.2019.00589] [Medline: 31543865]
9. Lear SA, Hu W, Rangarajan S, Gasevic D, Leong D, Iqbal R, et al. The effect of physical activity on mortality and cardiovascular disease in 130 000 people from 17 high-income, middle-income, and low-income countries: the PURE study. Lancet 2017 Dec 16;390(10113):2643-2654. [doi: 10.1016/S0140-6736(17)31634-3] [Medline: 28943267]
10. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity--a systematic review of longitudinal studies. BMC Public Health 2013 Sep 08;13:813 [FREE Full text] [doi: 10.1186/1471-2458-13-813] [Medline: 24010994]
11. Finkel D, Andel R, Pedersen NL. Gender differences in longitudinal trajectories of change in physical, social, and cognitive/sedentary leisure activities. J Gerontol B Psychol Sci Soc Sci 2018 Oct 10;73(8):1491-1500 [FREE Full text] [doi: 10.1093/geronb/gbw116] [Medline: 27624718]
12. 2018 Physical activity guidelines. US Department of Health and Human Services. URL: https://health.gov/sites/default/files/2019-09/Physical_Activity_Guidelines_2nd_edition.pdf [accessed 2022-01-08]
13. Arigo D, Romano K, Pasko K, Travers L, Ainsworth MC, Jackson DA, et al. A scoping review of behavior change techniques used to promote physical activity among women in midlife. PsyArXiv 2021 Preprint posted online August 31, 2021 [FREE Full text] [doi: 10.31234/osf.io/d4y3p]
14. Bélanger-Gravel A, Godin G, Amireault S. A meta-analytic review of the effect of implementation intentions on physical activity. Health Psychol Rev 2013 Mar;7(1):23-54. [doi: 10.1080/17437199.2011.560095]
15. Munson S, Consolvo S. Exploring goal-setting, rewards, self-monitoring, and sharing to motivate physical activity. In: Proceedings of the 6th International Conference on Pervasive Computing Technologies for Healthcare. 2012 Presented at: 6th International Conference on Pervasive Computing Technologies for Healthcare; May 21-24, 2012; San Diego, United States. [doi: 10.4108/icst.pervasivehealth.2012.248691]
16. Murray JM, Brennan SF, French DP, Patterson CC, Kee F, Hunter RF. Effectiveness of physical activity interventions in achieving behaviour change maintenance in young and middle aged adults: a systematic review and meta-analysis. Soc Sci Med 2017 Nov;192:125-133. [doi: 10.1016/j.socscimed.2017.09.021] [Medline: 28965003]
17. Herazo-Beltrán Y, Pinillos Y, Vidarte J, Crissien E, Suarez D, García R. Predictors of perceived barriers to physical activity in the general adult population: a cross-sectional study. Braz J Phys Ther 2017;21(1):44-50 [FREE Full text] [doi: 10.1016/j.bjpt.2016.04.003] [Medline: 28442074]
18. Infurna FJ, Gerstorf D, Lachman ME. Midlife in the 2020s: opportunities and challenges. Am Psychol 2020 May;75(4):470-485 [FREE Full text] [doi: 10.1037/amp0000591] [Medline: 32378943]
19. Hendry P, Solmon M, Choate LH, Autrey P, Landry JB. Midlife women's negotiations of barriers to and facilitators of physical activity: implications for counselors. Adultspan J 2010;9(1):50-64 [FREE Full text] [doi: 10.1002/j.2161-0029.2010.tb00071.x]
20. Taylor S. Social support: a review. In: The Oxford Handbook of Health Psychology. Oxfordshire, England, UK: Oxford University Press; 2011.
21. Janssen LH, Elzinga BM, Verkuil B, Hillegers MH, Keijsers L. The link between parental support and adolescent negative mood in daily life: between-person heterogeneity in within-person processes. J Youth Adolesc 2021 Feb;50(2):271-285 [FREE Full text] [doi: 10.1007/s10964-020-01323-w] [Medline: 32997209]
22. Hwang KO, Ottenbacher AJ, Green AP, Cannon-Diehl MR, Richardson O, Bernstam EV, et al. Social support in an internet weight loss community. Int J Med Inform 2010 Jan;79(1):5-13 [FREE Full text] [doi: 10.1016/j.ijmedinf.2009.10.003] [Medline: 19945338]
23. Festinger L. A theory of social comparison processes. Human Relation 2016 Apr 22;7(2):117-140. [doi: 10.1177/001872675400700202]
24. Cannioto RA. Physical activity barriers, behaviors, and beliefs of overweight and obese working women: a preliminary analysis. Women Sport Physical Activity J 2010 Apr;19(1):70-85. [doi: 10.1123/wspaj.19.1.70]
25. Rowland SA, Berg KE, Kupzyk KA, Pullen CH, Cohen MZ, Schulz PS, et al. Feasibility and effect of a peer modeling workplace physical activity intervention for women. Workplace Health Saf 2018 Sep;66(9):428-436. [doi: 10.1177/2165079917753690] [Medline: 29929437]
26. Bandura A. Health promotion from the perspective of social cognitive theory. Psychol Health 1998 Jul;13(4):623-649. [doi: 10.1080/08870449808407422]
27. Helgeson VS, Mickelson KD. Motives for social comparison. Pers Soc Psychol Bull 2016 Jul 02;21(11):1200-1209. [doi: 10.1177/01461672952111008]

28. Rancourt D, Leahey TM, LaRose JG, Crowther JH. Effects of weight-focused social comparisons on diet and activity outcomes in overweight and obese young women. Obesity (Silver Spring) 2015 Jan;23(1):85-89 [FREE Full text] [doi: 10.1002/oby.20953] [Medline: 25407550]
29. Buunk BP. Comparison direction and comparison dimension among disabled individuals: toward a refined conceptualization of social comparison under stress. Pers Soc Psychol Bull 2016 Jul 02;21(4):316-330. [doi: 10.1177/0146167295214002]
30. Muller D, Fayant MP. On being exposed to superior others: consequences of self-threatening upward social comparisons. Social Personality Psychol Compass 2010;4(8):621-634. [doi: 10.1111/j.1751-9004.2010.00279.x]
31. Buunk B, Ybema J. Social comparisons and occupational stress: the identification-contrast model. In: Health, Coping, and Well-being Perspectives From Social Comparison Theory. New York: Psychology Press; 1997.
32. Arigo D, Mogle JA, Smyth JM. Relations between social comparisons and physical activity among women in midlife with elevated risk for cardiovascular disease: an ecological momentary assessment study. J Behav Med 2021 Oct;44(5):579-590 [FREE Full text] [doi: 10.1007/s10865-021-00229-7] [Medline: 33982214]
33. Arigo D, Mogle JA, Brown MM, Pasko K, Travers L, Sweeder L, et al. Methods to assess social comparison processes within persons in daily life: a scoping review. Front Psychol 2019;10:2909 [FREE Full text] [doi: 10.3389/fpsyg.2019.02909] [Medline: 32038352]
34. Arigo D, Pasko K, Mogle JA. Daily relations between social perceptions and physical activity among college women. Psychol Sport Exerc 2020 Mar;47:101528 [FREE Full text] [doi: 10.1016/j.psychsport.2019.04.018] [Medline: 32831642]
35. Arigo D, Smyth JM, Suls JM. Perceptions of similarity and response to selected comparison targets in type 2 diabetes. Psychol Health 2015;30(10):1206-1220. [doi: 10.1080/08870446.2015.1040018] [Medline: 25871344]
36. Van der Zee K, Buunk B, Sanderman R, Botke G, van den Bergh F. Social comparison and coping with cancer treatment. Personal Individual Differences 2000 Jan;28(1):17-34. [doi: 10.1016/s0191-8869(99)00045-8]
37. Butryn ML, Arigo D, Raggio GA, Colasanti M, Forman EM. Enhancing physical activity promotion in midlife women with technology-based self-monitoring and social connectivity: a pilot study. J Health Psychol 2016 Aug;21(8):1548-1555. [doi: 10.1177/1359105314558895] [Medline: 25488937]
38. Peterson JA, Yates BC, Atwood JR, Hertzog M. Effects of a physical activity intervention for women. West J Nurs Res 2005 Feb 01;27(1):93-110. [doi: 10.1177/0193945904270912] [Medline: 15659587]
39. Arigo D. Promoting physical activity among women using wearable technology and online social connectivity: a feasibility study. Health Psychol Behav Med 2015 Dec 31;3(1):391-409. [doi: 10.1080/21642850.2015.1118350]
40. Arigo D, Schumacher LM, Pinkasavage E, Butryn ML. Addressing barriers to physical activity among women: a feasibility study using social networking-enabled technology. Digit Health 2015;1:2055207615583564 [FREE Full text] [doi: 10.1177/2055207615583564] [Medline: 29942539]
41. Smith KL, Kerr DA, Fenner AA, Straker LM. Adolescents just do not know what they want: a qualitative study to describe obese adolescents' experiences of text messaging to support behavior change maintenance post intervention. J Med Internet Res 2014 Apr 08;16(4):e103 [FREE Full text] [doi: 10.2196/jmir.3113] [Medline: 24713407]
42. Collins LM, Murphy SA, Bierman KL. A conceptual framework for adaptive preventive interventions. Prev Sci 2004 Sep;5(3):185-196 [FREE Full text] [doi: 10.1023/b:prev.0000037641.26017.00] [Medline: 15470938]
43. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res 2008 Jun;23(3):454-466 [FREE Full text] [doi: 10.1093/her/cyn004] [Medline: 18349033]
44. Monteiro-Guerra FM, Rivera-Romero O, Fernandez-Luque L, Caulfield B. Personalization in real-time physical activity coaching using mobile applications: a scoping review. IEEE J Biomed Health Inform 2020 Jun;24(6):1738-1751. [doi: 10.1109/JBHI.2019.2947243] [Medline: 31751254]
45. Zhu J, Dallal DH, Gray RC, Villareale J, Ontañón S, Forman EM, et al. Personalization paradox in behavior change apps: lessons from a social comparison-based personalized app for physical activity. Proc ACM Human Comput Interact 2021 Apr 13;5(CSCW1):1-21. [doi: 10.1145/3449190]
46. Everett E, Kane B, Yoo A, Dobs A, Mathioudakis N. A novel approach for fully automated, personalized health coaching for adults with prediabetes: pilot clinical trial. J Med Internet Res 2018 Feb 27;20(2):e72 [FREE Full text] [doi: 10.2196/jmir.9723] [Medline: 29487046]
47. Zuckerman O, Gal-Oz A. Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity. Pers Ubiquit Comput 2014 Jul 5;18(7):1705-1719. [doi: 10.1007/s00779-014-0783-2]
48. Wilson JM, Smith K, Strough J, Delaney R. Knowing you are there makes the difference: perceived social support, preferences for using support, and health. J Women Aging 2021;33(4):396-410. [doi: 10.1080/08952841.2020.1860633] [Medline: 33347380]
49. Gibbons FX, Buunk BP. Individual differences in social comparison: development of a scale of social comparison orientation. J Pers Soc Psychol 1999 Jan;76(1):129-142. [doi: 10.1037//0022-3514.76.1.129] [Medline: 9972558]
50. Arigo D, Brown M, Shank F, Young C. Ecological momentary assessment of relations between social interactions and physical activity outcomes among women in midlife with CVD risk conditions. Annals of Behavioral Medicine 2022 (forthcoming).

51. Abras C, Maloney-Krichmar D, Preece J. User-centered design. In: Encyclopedia of Human-Computer Interaction. Thousand Oaks, California: SAGE Publications; 2004.
52. Graham AK, Neubert SW, Chang A, Liu J, Fu E, Green EA, et al. Applying user-centered design methods to understand users' day-to-day experiences can inform a mobile intervention for binge eating and weight management. Front Digit Health 2021;3:651749 [FREE Full text] [doi: 10.3389/fdgth.2021.651749] [Medline: 34713124]
53. Alberts NM, Badawy SM, Hodges J, Estepp JH, Nwosu C, Khan H, et al. Development of the incharge health mobile app to improve adherence to hydroxyurea in patients with sickle cell disease: user-centered design approach. JMIR Mhealth Uhealth 2020 May 08;8(5):e14884 [FREE Full text] [doi: 10.2196/14884] [Medline: 32383683]
54. Korpershoek YJ, Hermsen S, Schoonhoven L, Schuurmans MJ, Trappenburg JC. User-centered design of a mobile health intervention to enhance exacerbation-related self-management in patients with chronic obstructive pulmonary disease (Copilot): mixed methods study. J Med Internet Res 2020 Jun 15;22(6):e15449 [FREE Full text] [doi: 10.2196/15449] [Medline: 32538793]
55. van Dantzig S, Geleijnse G, van Halteren AT. Toward a persuasive mobile application to reduce sedentary behavior. Pers Ubiquit Comput 2012 Jul 12;17(6):1237-1246. [doi: 10.1007/s00779-012-0588-0]
56. Sporrel K, De Boer RD, Wang S, Nibbeling N, Simons M, Deutekom M, et al. The design and development of a personalized leisure time physical activity application based on behavior change theories, end-user perceptions, and principles from empirical data mining. Front Public Health 2020;8:528472 [FREE Full text] [doi: 10.3389/fpubh.2020.528472] [Medline: 33604321]
57. 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]
58. Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. Transl Behav Med 2021 Jul 29;11(7):1473-1480 [FREE Full text] [doi: 10.1093/tbm/ibab048] [Medline: 33963864]
59. Grossman JA, Arigo D, Bachman JL. Meaningful weight loss in obese postmenopausal women: a pilot study of high-intensity interval training and wearable technology. Menopause 2018 Apr;25(4):465-470. [doi: 10.1097/GME.0000000000001013] [Medline: 29088015]
60. Arigo D, Brown MM, Pasko K, Ainsworth MC, Travers L, Gupta A, et al. Rationale and design of the women's health and daily experiences project: protocol for an ecological momentary assessment study to identify real-time predictors of midlife women's physical activity. JMIR Res Protoc 2020 Oct 15;9(10):e19044 [FREE Full text] [doi: 10.2196/19044] [Medline: 33055065]
61. Arigo D, Mogle JA, Brown MM, Gupta A. A multi-study approach to refining ecological momentary assessment measures for use among midlife women with elevated risk for cardiovascular disease. Mhealth 2021;7:53 [FREE Full text] [doi: 10.21037/mhealth-20-143] [Medline: 34805384]
62. Pxfuel homepage. Pxfuel. URL: https://www.pxfuel.com/ [accessed 2022-07-11]
63. Legler S, Celano CM, Beale EE, Hoeppner BB, Huffman JC. Use of text messages to increase positive affect and promote physical activity in patients with heart disease: the Promoting Activity in Cardiac Patients via Text Messages (PACT) pilot study. Curr Psychol 2020;39:648-655 [FREE Full text] [doi: 10.1007/s12144-018-9785-y] [Medline: 32982125]
64. Thakkar J, Redfern J, Thiagalingam A, Chow CK. Patterns, predictors and effects of texting intervention on physical activity in CHD - insights from the TEXT ME randomized clinical trial. Eur J Prev Cardiol 2016 Nov;23(17):1894-1902. [doi: 10.1177/2047487316664190] [Medline: 27512051]
65. Spark LC, Fjeldsoe BS, Eakin EG, Reeves MM. Efficacy of a text message-delivered extended contact intervention on maintenance of weight loss, physical activity, and dietary behavior change. JMIR Mhealth Uhealth 2015 Sep 15;3(3):e88 [FREE Full text] [doi: 10.2196/mhealth.4114] [Medline: 26373696]
66. Wittmann BC, Bunzeck N, Dolan RJ, Düzel E. Anticipation of novelty recruits reward system and hippocampus while promoting recollection. Neuroimage 2007 Oct 15;38(1):194-202 [FREE Full text] [doi: 10.1016/j.neuroimage.2007.06.038] [Medline: 17764976]
67. Wittmann BC, Daw ND, Seymour B, Dolan RJ. Striatal activity underlies novelty-based choice in humans. Neuron 2008 Jun 26;58(6):967-973 [FREE Full text] [doi: 10.1016/j.neuron.2008.04.027] [Medline: 18579085]
68. Howlett N, Trivedi D, Troop NA, Chater AM. Are physical activity interventions for healthy inactive adults effective in promoting behavior change and maintenance, and which behavior change techniques are effective? A systematic review and meta-analysis. Transl Behav Med 2019 Jan 01;9(1):147-157 [FREE Full text] [doi: 10.1093/tbm/iby010] [Medline: 29506209]
69. Brown MM, Arigo D. Changes in life circumstances and mental health symptoms during the COVID-19 pandemic among midlife women with elevated risk for cardiovascular disease. J Women Aging 2021 Aug 25:1-12. [doi: 10.1080/08952841.2021.1967654] [Medline: 34432597]
70. Stockwell S, Trott M, Tully M, Shin J, Barnett Y, Butler L, et al. Changes in physical activity and sedentary behaviours from before to during the COVID-19 pandemic lockdown: a systematic review. BMJ Open Sport Exerc Med 2021;7(1):e000960 [FREE Full text] [doi: 10.1136/bmjsem-2020-000960] [Medline: 34192010]

71. Wills TA. Downward comparison principles in social psychology. Psychol Bull 1981;90(2):245-271. [doi: 10.1037/0033-2909.90.2.245]
72. Meier A, Schäfer S. Positive side of social comparison on social network sites: how envy can drive inspiration on instagram. Cyberpsychol Behav Soc Netw 2018 Jul;21(7):411-417. [doi: 10.1089/cyber.2017.0708] [Medline: 29995526]
73. Göttgens I, Oertelt-Prigione S. The application of human-centered design approaches in health research and innovation: a narrative review of current practices. JMIR Mhealth Uhealth 2021 Dec 06;9(12):e28102 [FREE Full text] [doi: 10.2196/28102] [Medline: 34874893]
74. Arigo D, Jake-Schoffman DE, Wolin K, Beckjord E, Hekler EB, Pagoto SL. The history and future of digital health in the field of behavioral medicine. J Behav Med 2019 Feb;42(1):67-83 [FREE Full text] [doi: 10.1007/s10865-018-9966-z] [Medline: 30825090]
75. Krukowski RA, Jagsi R, Cardel MI. Academic productivity differences by gender and child age in science, technology, engineering, mathematics, and medicine faculty during the COVID-19 pandemic. J Womens Health (Larchmt) 2021 Mar;30(3):341-347 [FREE Full text] [doi: 10.1089/jwh.2020.8710] [Medline: 33216682]
76. 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]

Abbreviations

CVD: cardiovascular disease
IRB: institutional review board
PA: physical activity
UCD: user-centered design

Edited by A Mavragani; submitted 08.01.22; peer-reviewed by J Hawkins, H Lee, E Dervic, A Graham; comments to author 16.02.22; revised version received 18.03.22; accepted 10.04.22; published 26.07.22.

Please cite as:
Arigo D, Lobo AF, Ainsworth MC, Baga K, Pasko K
Development and Initial Testing of a Personalized, Adaptive, and Socially Focused Web Tool to Support Physical Activity Among Women in Midlife: Multidisciplinary and User-Centered Design Approach
JMIR Form Res 2022;6(7):e36280
URL: https://formative.jmir.org/2022/7/e36280
doi:10.2196/36280
PMID:35881431

Original Paper

Detection of Potential Arbovirus Infections and Pregnancy Complications in Pregnant Women in Jamaica Using a Smartphone App (ZIKApp): Pilot Evaluation Study

Elisa Ruiz-Burga[1], PhD; Patricia Bruijning-Verhagen[2], PhD; Paulette Palmer[3], MPH; Annalisa Sandcroft[3], BScN, RN; Georgina Fernandes[1], MSc; Marieke de Hoog[4], PhD; Lenroy Bryan[5], BSc; Russell Pierre[3], MD; Heather Bailey[6], PhD; Carlo Giaquinto[7], MD; Claire Thorne[1], PhD; Celia D C Christie[3], DMPeds; ZIKAction Consortium[8]

[1]Population, Policy & Practice Research and Teaching Department, Great Ormond Street Institute of Child Health, University College London, London, United Kingdom

[2]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands

[3]Department of Child and Adolescent Health, University of the West Indies, Kingston, Jamaica

[4]University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands

[5]Department of Obstetrics and Gynaecology, University of the West Indies, Kingston, Jamaica

[6]Institute for Global Health, University College London, London, United Kingdom

[7]Dipartimento di Salute della Donna e del Bambino, Università degli Studi di Padova, Padova, Italy

[8]Fondazione Penta Onlus, Padova, Italy

Corresponding Author:
Elisa Ruiz-Burga, PhD
Population, Policy & Practice Research and Teaching Department
Great Ormond Street Institute of Child Health
University College London
30 Guilford St,
London, WC1N 1EH
United Kingdom
Phone: 44 02079052396
Email: e.burga@ucl.ac.uk

Abstract

Background: There is growing evidence of the benefits of mobile health technology, which include symptom tracking apps for research, surveillance, and prevention. No study has yet addressed arbovirus symptom tracking in pregnancy.

Objective: This study aimed to evaluate the use of a smartphone app (*ZIKApp*) to self-report arbovirus symptoms and pregnancy complications and to assess compliance with daily symptom diaries during pregnancy in a cohort of women in an arbovirus-endemic, subtropical, middle-income country (Jamaica).

Methods: Pregnant women aged ≥16 years, having a smartphone, and planning on giving birth at the recruiting center were enrolled between February 2020 and July 2020. ZIKApp comprised a daily symptom diary based on algorithms to identify potential episodes of arbovirus infection and pregnancy complications. Sociodemographic, epidemiological, and obstetric information was collected at enrollment, with additional review of medical records, and users' perception was collected through an exit survey. Descriptive analyses and logistic regression analysis of possible factors associated with diary adherence were performed.

Results: Of the 173 women enrolled, 157 (90.8%) used ZIKApp for a median duration of 155 (IQR 127-173) days until pregnancy end, 6 (3.5%) used the app for <7 days, and 10 (5.8%) exited the study early. For each successive 30-day period from enrollment up to 150 days after enrollment, of these 157 women, 121 (77.1%) to 129 (82.2%) completed their daily symptom diary; 50 (31.8%) to 56 (35.7%) did so on the same day. Overall, 31.8% (50/157) of the women had *good adherence* to diary reporting (ie, they completed the task on the same day or 2 to 3 days later for ≥80% of the days enrolled). There were 3-fold higher odds of good adherence for participants aged >34 years versus those aged 25 to 29 years (adjusted odds ratio 3.14, 95% CI 1.10-8.98) and 2-fold higher odds for women with tertiary versus secondary education (adjusted odds ratio 2.26, 95% CI 1.06-4.83). Of the 161 women who ever made a diary entry, 5454 individual symptom reports were made (median 17 per woman; IQR 4-42; range

0-278); 9 (5.6%) women reported symptom combinations triggering a *potential arbovirus episode* (none had an adverse pregnancy outcome) and 55 (34.2%) reported painful uterine contractions or vaginal bleeding, mainly in the month before delivery. Overall, 51.8% (71/137) of the women rated the app as an excellent experience and were less likely to be poor diary adherers (P=.04) and 99.3% (138/139) reported that the app was easy to understand and use.

Conclusions: This pilot found a high adherence to ZIKApp. It demonstrated the feasibility and usability of the app in an arbovirus-endemic region, supporting its future development to contribute to surveillance and diagnosis of arbovirus infections in pregnancy and to optimize maternal care.

(JMIR Form Res 2022;6(7):e34423) doi:10.2196/34423

KEYWORDS

mHealth; digital health; arbovirus; pregnancy; adherence; compliance; low- and middle-income countries; LMIC; maternal health; pregnancy complications; prenatal care; pregnancy outcomes; mobile phone

Introduction

Background

Zika virus (ZIKV), chikungunya virus (CHIKV), and dengue virus (DENV) are arboviruses transmitted primarily by the *Aedes aegypti* mosquito that have caused multiple epidemics in recent years, notably the explosive epidemic of CHIKV in Latin America and the Caribbean that started in 2013, followed by the re-emergence of ZIKV a year later [1,2]. These arboviruses are of major public health concern and have been associated with significant morbidity alongside substantial economic impacts [3-5]. Although ZIKV outbreaks in particular have been highlighted with respect to maternal and infant health, owing to the causal link between ZIKV and microcephaly and other fetal and infant abnormalities [6], CHIKV and DENV are also both vertically transmitted, and all 3 viruses have been associated with adverse pregnancy outcomes (eg, preterm delivery, miscarriage, and stillbirth) and infant sequelae in the context of congenital infection [7-10].

Jamaica has experienced several DENV outbreaks of increasing intensity, severity, and magnitude in recent years (2010, 2012, and 2018-2019), with increased attributable morbidity and mortality in the very young, including reported cases of neonatal microcephaly [11-14]. Jamaica also experienced an explosive CHIKV epidemic in 2014 [3,15-18]: >80% of the general and antenatal populations were affected, with significant perinatal (maternal and newborn) illnesses, increased attributable neonatal morbidity and mortality [3,16,19], and >2500 deaths during the epidemic year nationally [18].

The ZIKV epidemic in Jamaica occurred in 2016-2017 [20-23]. National hospital-based surveillance revealed increased annual rates of severe microcephaly from 23.6 per 100,000 live births in 2010 (during the period of DENV outbreaks) to 41.7 per 100,000 live births in 2017 in association with the ZIKV epidemic [14], alongside surveillance reports of the congenital syndrome associated with ZIKV and related fetal brain disruption sequences temporally and spatially related to reported ZIKV cases [14]. The congenital syndrome was also being recognized simultaneously in 0.8% to 2.2% of the newborns in 3 urban public maternity hospitals [23], with a 15.6% ZIKV immunoglobulin G antibody seroprevalence in pregnancy reported [19]. Other complications included Guillain-Barré syndrome and varied neurological presentations in children and adults [21,22].

The COVID-19 pandemic has seen a growing number of symptom tracker mobile phone apps that have helped to develop an understanding of an emerging infection and its associated clinical manifestations [24-27]. Before the emergence of SARS-CoV-2, there was already an increasing body of research demonstrating the benefits of mobile health (mHealth) technology for remote monitoring of symptoms, public health surveillance, education, and prevention [28], including for arboviruses [29-32]. With respect to pregnancy, mHealth apps have been used for multiple interventions, including to optimize gestational weight gain, to increase intake of vegetables and fruit, for smoking cessation, to identify specific symptoms of pre-eclampsia, for drug safety monitoring, and to support health care delivery for prevention of asthma and infections [33-38]. However, there is limited literature on the levels of adherence to mHealth apps for daily reporting of symptoms during pregnancy. This is a significant gap in evidence because the requirement for daily reporting could be an important factor potentially limiting long-term adherence.

Objectives

This pilot study aimed to evaluate the use of ZIKApp by pregnant women to report symptoms indicative of an arbovirus infection and pregnancy complications, including evaluating adherence with daily symptom diary reporting up to the end of pregnancy. The study setting was Jamaica, an example of an arbovirus-endemic, subtropical, middle-income country with high levels of smartphone penetration.

Methods

Recruitment

This pilot study was conducted at the University Hospital of the West Indies (UHWI), a university teaching hospital that performed approximately 1500 deliveries in 2016, as part of research conducted by the ZIKAction consortium, which conducts maternal and child health–focused research on ZIKV and other arboviruses in Latin America and the Caribbean. Participants were recruited between February 2020 and July 2020 from the antenatal care service by research nurses. This pilot study sought to enroll approximately 200 participants who met the following eligibility criteria: pregnant women aged ≥16

years who planned to give birth at the UHWI and had access to a smartphone compatible with the mobile app. Following these criteria, the research nurses orally explained the study, invited these pregnant women in the waiting room of this antenatal clinic on their first medical clinic visit (when the general educational talks are delivered) to take part in the study, and enrolled those interested. Furthermore, some of the participants also told other pregnant women about the study and recommended their participation, after which these women would also individually approach the research nurses, who would follow-up with the standard enrollment procedures. Names and contact details were stored locally by the research nurses, but all study questionnaires were pseudonymized with the use of unique study identifiers. Enrollment visit procedures are presented in the *Study Implementation* section.

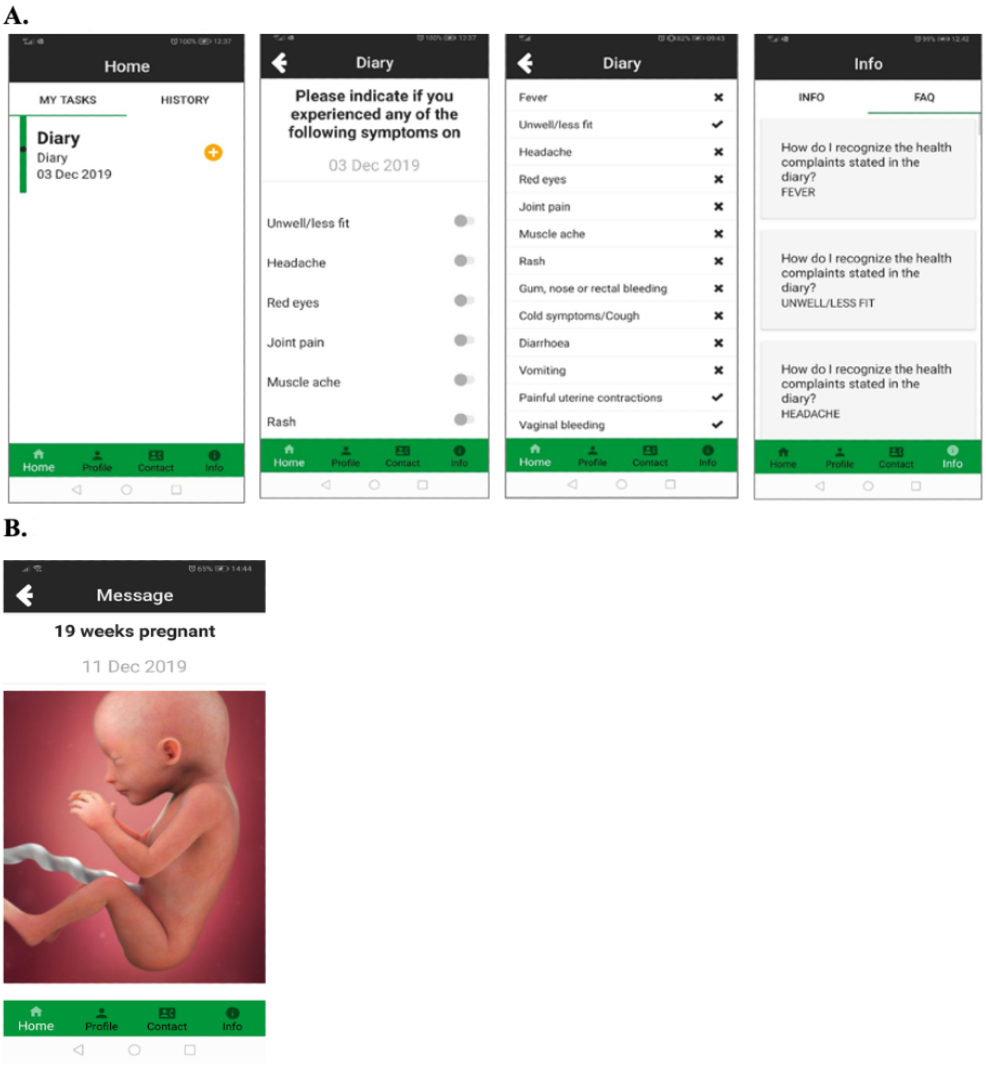
The ZIKApp Intervention

ZIKApp was developed by the University Medical Center Utrecht in partnership with Your Research, a company based in the Netherlands. ZIKApp is compatible with both Android and iOS platforms. The app included a daily symptom diary for reporting presence or absence of symptoms and provided users with information on potential symptoms. In addition, the app provided information about the study itself and included a *Frequently Asked Questions* section (Figure 1A).

The app was designed to provide regular informative messages related to the pregnancy such as "Your baby is now the size of a pear," similar to content provided by commercial pregnancy tracking apps (Figure 1B). This was to provide an additional incentive for women to adhere to app use.

Participants were sent daily reminders by email to complete their symptom diary. The daily diary entry could be completed on the same day or with a lag period of up to 7 days, after which it was no longer accessible. The symptoms included in the app and the start and end of arbovirus episode triggers are presented in Multimedia Appendix 1 and were selected to identify symptoms suggestive of arbovirus infection as well as 2 specific pregnancy complications (painful uterine contractions and vaginal bleeding). The app also included some short periodic questionnaires; for example, the first questionnaire requested information on the woman's last menstrual period date (to allow the app to send gestation-appropriate messages). Data recorded in the app were stored on a secure cloud-based portal (ResearchFollowApp portal). The portal also provided a user-friendly dashboard for the research nurses with an authorized log-in to monitor in real time the symptoms reported by participants.

Figure 1. ZIKApp screenshots. (A) Diary and symptom information. (B) Example of an informative message regarding pregnancy: "...your baby at week 19. The top of your womb (uterus) now reaches your belly button and will grow about a centimetre higher than this each week. Your baby measures up to 15.3 cm (6 in) from head to bottom (crown to rump) and weighs about 240 g (8.5 oz). He's the size of a big tomato...".



Study Implementation

A standard operating procedure (SOP) was developed to guide the research nurses, and in-person training was conducted at the UHWI in January 2020. The SOP detailed the different tasks that research nurses should fulfill during the recruitment once informed consent was obtained. It explained the process of setting up an account for the participant through the portal, assisting her in downloading the app from the Google Play Store or Apple App Store, guiding the initial app log-in, instructing her about the different app functions and notifications that she could receive (*potential arbovirus episode* or *possible pregnancy complication*), and guiding her to complete her first diary entry. Likewise, the SOP detailed the process to report any app or network issues to the coordinating University College London and University Medical Center Utrecht teams to resolve them and thus avoid early exits by participants and to improve app functionality.

At recruitment, a standardized form was used to collect additional sociodemographic, obstetric, and clinical information regarding the participant extracted from the medical chart. Data were entered and managed using REDCap (Research Electronic Data Capture; Vanderbilt University) hosted at Penta Foundation ONLUS [39,40]. REDCap is a secure, web-based software platform designed to support data capture for research studies. The ResearchFollowApp portal was also designed to facilitate the follow-up of participants by the research nurses and to send automatic emails and notifications. The research nurses implemented the use of WhatsApp messages to inform participants about their office hours and to advise those who received notifications of *potential arbovirus episode* or *possible pregnancy complication* to contact the health services using a provided list of phone numbers. All participants received monthly phone credits to enable internet access to allow the app to transmit the data collected to the portal.

Study participation ended when the participant gave birth or when the pregnancy came to an end for other reasons (eg, miscarriage), although the women could withdraw from the study at any time. At study exit, the research nurses guided participants to uninstall ZIKApp and invited them to complete an exit survey that included questions on their experience of using the app. In addition, a standardized questionnaire was used to collect data regarding the delivery as well as details of the newborn, which were entered in the study REDCap database.

A case note review was conducted for women with potential arbovirus triggers to identify any maternal diagnoses (eg, from the laboratory information systems). During the pilot study, surveillance for SARS-CoV-2 among the pregnant women was implemented at the UHWI, and a case note review was also carried out for enrolled women with positive SARS-CoV-2 tests.

Statistical Analysis

App data stored on the ResearchFollowApp portal was downloaded and merged with data from the REDCap database using unique study identifiers before analysis. Descriptive analyses of participant characteristics were conducted. Univariable comparisons of categorical variables were assessed using chi-square or Fisher exact tests. To assess participants' adherence to completing the daily symptom diary, every diary day for each woman (ie, for the total time they used the app) was coded into 1 of 4 categories: diary completed on the same day, 2 to 3 days later, 4 to 7 days later, and >7 days elapsed without diary entry (ie, not completed). Next, for each woman, 2 binary variables (0 and 1) were created: for *good adherence* and *poor adherence* for their entire period of enrollment. *Good adherence* was where a participant had completed her diary on the same day or 2 to 3 days later for at least 80% of the time between enrollment and pregnancy end. *Poor adherence* was where there was an uncompleted diary (ie, >7 days had elapsed without diary entry) 30% of the time.

Potential factors associated with *good* and *poor* adherence were assessed using logistic regression analysis to obtain the odds ratios with 95% CIs: participant age, education, number of children, having an income, previous adverse pregnancy outcome, comorbidities and chronic diseases, duration of app use, and whether a potential arbovirus episode was reported. Several factors were considered when selecting the final multivariable models: first, all variables that were significant with a *P* value of <.10 in univariable analysis were considered for inclusion; second, a backward stepwise selection approach was used to determine the final adjusted model. Stata software (version 16.1; StataCorp LLC) was used to conduct the analyses.

Ethics Approval

The protocol for this study was reviewed and approved by the University College London Research Committee on 27 September 2021 (Project ID 3715/005) and by the University of the West Indies Mona Campus Research Ethics Committee (project ID ECP 47, 19/20). All participants signed an informed consent form during the enrollment.

Results

Overview

A total of 173 pregnant women were enrolled in the study (Figure 2), with the last delivery occurring on January 5, 2021. Of these 173 women, 5 (2.9%) had no data recorded in the app portal, indicating that they never used the app (although, of these 5 women, for 2, 40%, this may have been a technical or connection issue as both reported completing the diary in their exit survey) and 1 (0.6%) had <1 week of study participation owing to a miscarriage 6 days after enrollment. These women were excluded from further analyses. Of the remaining 167 women, 157 (94%) used the app until they gave birth or the end of their pregnancy, whereas 10 (6%) chose to exit the study before they gave birth.

Figure 2. Study participant flow chart.

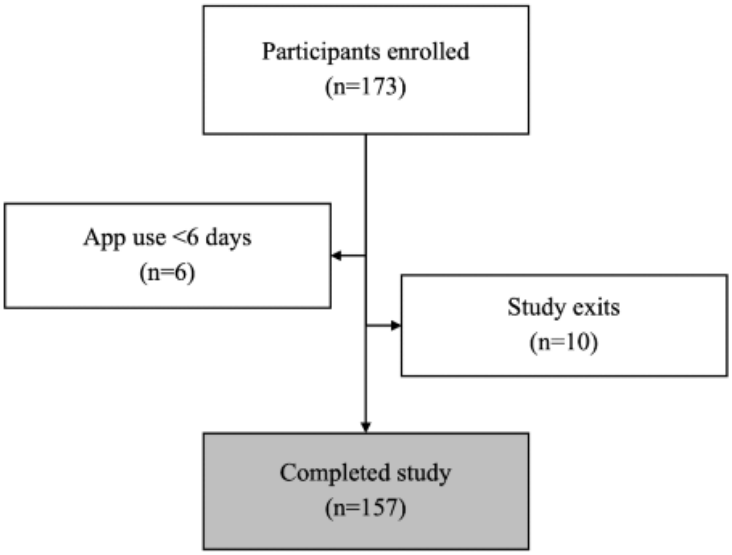


Table 1 presents the participants' baseline characteristics, stratified by study completion and study exit status. Overall, of the 167 women, 166 (99.4%) were born in Jamaica and 1 (0.6%) was born in the United Kingdom; most (137/167, 82%) of the women were enrolled in the second trimester of pregnancy. Of the 10 women who exited the study, 6 (60%) gave phone-related reasons for their study exit (eg, no longer having a smartphone or having changed their phone), 2 (20%) changed their mind about study participation, and 2 (20%) cited problems with phone credit reimbursement. There were no statistically significant differences between the women who exited the study and those who remained and used the app until they gave birth with respect to sociodemographic characteristics and comorbidities (data not shown).

Table 1. Participant sociodemographic, clinical, and app use characteristics by study completion and study exit status (N=167).

Characteristics	Completed study (n=157)	Exited study (n=10)
Age (years), median (IQR; range)	28 (24-32; 18-44)	26.5 (22-31; 21-36)
Marital status, n (%)		
Married	45 (28.7)	1 (10)
Cohabiting	47 (29.9)	3 (30)
Single	60 (38.2)	6 (60)
Divorced or separated	5 (3.2)	0 (0)
Highest level of education, n (%)		
Secondary	61 (38.9)	3 (30)
Tertiary	96 (61.2)	7 (70)
Employed or has regular income, n (%)		
No	77 (49)	3 (30)
Yes	80 (51)	7 (70)
Parity, n (%)		
Nulliparous	90 (57.3)	5 (50)
Primiparous	41 (26.1)	3 (30)
Multiparous	26 (16.6)	2 (20)
Chronic conditions, n (%)		
None	93 (59.2)	6 (60)
Hypertension[a]	26 (16.6)	0 (0)
Sickle cell disease[a]	11 (7)	0 (0)
Pre-existing or gestational diabetes[a]	7 (4.6)	0 (0)
Asthma[a]	24 (15.3)	1 (10)
Obesity[a]	6 (3.8)	1 (10)
Gestational age at enrollment (weeks), median (IQR; range)	18 (15-22; 11-38)	17 (16-18; 14-21)
Pregnancy outcome, n (%)		
Live birth	151 (96.2)	N/A[b]
Stillbirth	2 (1.3)	N/A
Miscarriage	3 (1.9)	N/A
Termination (abnormality)	1 (0.6)	N/A
Gestational age at delivery[c] (weeks), n (%)		
<34	4 (2.7)	N/A
34 to 36	10 (6.6)	N/A
≥37	136 (90.1)	N/A
Duration of app use (days), median (IQR; range)	155 (127-173; 25-235)	133 (104-172; 80-179)

[a]Multiple responses were possible.

[b]N/A: not applicable.

[c]Live births only (1 unknown).

Daily Symptom Diary Reporting and Factors Associated With Good and Poor Adherence

Overall, 78.68% (17,833/22,664) of the daily diaries were completed over the study period. The timing of symptom diary completion (ie, same day, 2-3 days later, 4-7 days later, or not completed) per successive 30-day periods between enrollment and pregnancy end is presented in Figure 3 and Multimedia Appendix 2. For each of these periods (up to 150 days), of these 157 women, between 121 (77.1%) and 129 (82.2%) completed

their daily symptom diary, with the proportion of diaries completed on the same day per 30-day period (up to 150 days) staying roughly constant at one-third (Figure 3; Multimedia Appendix 2).

The proportion of noncompleted diaries (ie, >7 days elapsed without diary completion) increased from the fifth month of app use. From day 181 onwards, the 7-day period within which a participant could complete her symptom diary retrospectively would have likely encompassed the date of delivery. Considering only diary days that were completed, there was also evidence of a reduced timeliness of reporting with increasing enrollment duration, with 41.69% (1024/2456), 37.67% (440/1168), and 32.4% (56/173) of the diaries completed on the same day at days 121-150, 151-180, and 181-210, respectively (P<.001).

To understand whether external events (eg, the Christmas holiday season) had an impact on the timeliness of diary completion, we compared the period from December 19, 2020, to January 2, 2021, and the period from December 1, 2020, to December 15, 2020, and found that the proportion of diaries completed on the same day was higher in the earlier period than during the Christmas period at 33.5% (190/568) versus 23.1% (78/337), respectively, although the proportion of noncompleted diaries was lower at 24.6% (140/568) versus 29.4% (99/337; P=.01), respectively.

Among the 157 women who used the app until the end of their pregnancy, there were 50 (31.9%) classified as having good adherence to diary completion (ie, completed same day or 2-3 days later at least 80% of the days). In univariable analysis (Table 2), women in the oldest age group (women aged >34 years vs those aged 25-29 years) and with tertiary education level (vs secondary education level) had higher odds of good adherence, although there was no association between good adherence and history of adverse pregnancy outcome, having a regular income, number of children, chronic disease status, duration of app use, or report of symptoms of arbovirus infection through the app. In the adjusted model (which included age and education only), both variables remained independently associated with good adherence, with the odds of good adherence being 3-fold higher for participants aged >34 years compared with those aged 25 to 29 years and 2-fold higher for women with a tertiary education compared with those who received a secondary education (Table 2).

In total, 24.8% (39/157) of the participants were classified as poor adherers to the symptom diary. Consistent with findings for good adherence, older women (those aged >34 years) had a significantly lower odds (adjusted odds ratio 0.10, 95% CI 0.01-0.80) of being poor adherers than participants aged 25 to 29 years. In addition, a short duration of app use (ie, <90 days) was associated with poor adherence (adjusted odds ratio 2.71, 95% CI 1.06-6.93).

Figure 3. Timing of symptom diary completion, by 30 day period from initiation of app use (n=157). The raw data corresponding to the percentages, and the number of women contributing to each 30 day period, are shown in Multimedia Appendix 2.

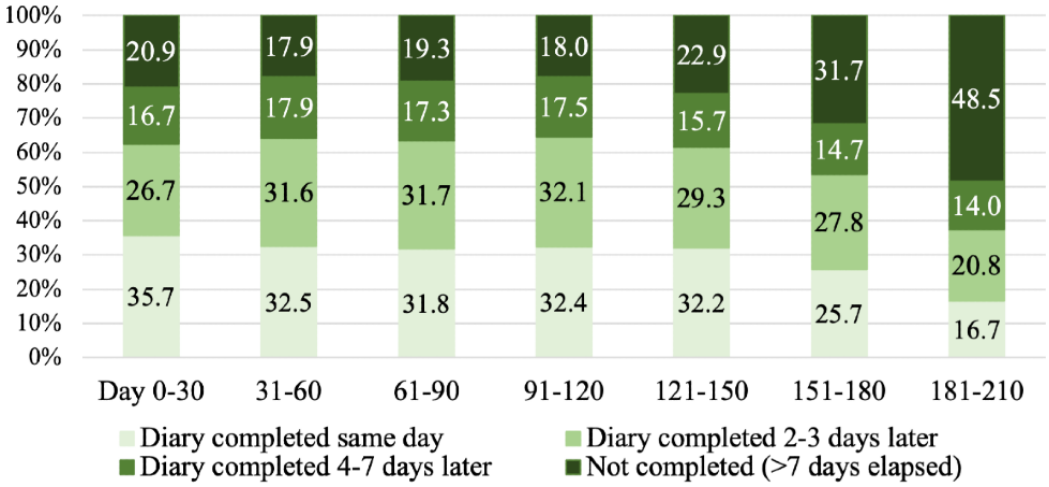


Table 2. Factors associated with good adherence to symptom diary reporting (N=157).

Explanatory variables	Values, n (%)	Good adherence, n (%)	Unadjusted OR[a] (95% CI)	*P* value	Adjusted OR (95% CI)	*P* value
Participant age (years)						
25 to 29	53 (33.8)	15 (28.3)	Reference	N/A[b]	Reference	N/A
<25	40 (25.5)	10 (25)	0.84 (0.33-2.15)	.72	0.97 (0.37-2.51)	.95
30 to 34	42 (26.8)	13 (31)	1.14 (0.47-2.75)	.78	1.22 (0.49-3.00)	.67
>34	22 (14)	12 (54.6)	3.04 (1.09-8.52)	.03	3.14 (1.10-8.98)	.03
Previous adverse pregnancy outcome						
No	110 (70.1)	35 (31.8)	Reference	N/A	N/A	N/A
Yes	47 (29.9)	15 (31.9)	1.00 (0.48-2.09)	.99	N/A	N/A
Regular income						
No	77 (49)	23 (29.9)	Reference	N/A	N/A	N/A
Yes	80 (51)	27 (33.8)	1.20 (0.61-2.34)	.60	N/A	N/A
Number of children						
0	90 (57.3)	26 (28.9)	Reference	N/A	N/A	N/A
1	41 (26.1)	14 (34.2)	1.28 (0.58-2.81)	.55	N/A	N/A
>1	26 (16.6)	10 (38.5)	1.54 (0.62-3.83)	.36	N/A	N/A
Education						
Secondary	61 (38.9)	13 (21.3)	Reference	N/A	Reference	N/A
Tertiary	96 (65.3)	37 (38.5)	2.32 (1.11-4.84)	.03	2.26 (1.06-4.83)	.03
Comorbidities and chronic diseases						
No	93 (59.2)	29 (31.2)	Reference	N/A	N/A	N/A
Yes	64 (40.8)	21 (32.8)	1.08 (0.55-2.13)	.83	N/A	N/A
Duration of app use (days)						
<90	25 (15.9)	7 (28)	Reference	N/A	N/A	N/A
≥90	132 (84.1)	43 (32.6)	1.24 (0.48-3.20)	.65	N/A	N/A
Symptoms of arbovirus infection reported						
No	148 (94.3)	49 (33.1)	Reference	N/A	N/A	N/A
Yes	9 (5.7)	1 (11.1)	0.25 (0.03-2.08)	.20	N/A	N/A

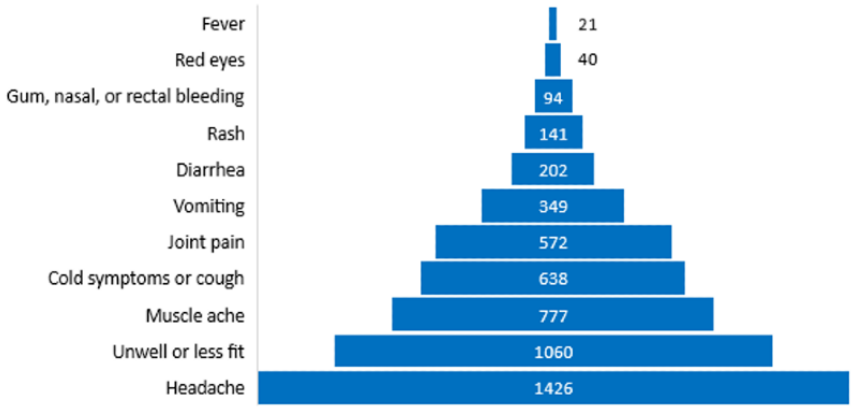
[a]OR: odds ratio.

[b]N/A: not applicable.

Reporting of Symptoms Overall

Overall, across the 17,883 completed diaries, there were 5454 (30.5%) individual reports of symptoms (n=5320, 29.75%, if vaginal bleeding and painful uterine contractions were excluded), including symptoms that may have been reported on the same day (eg, a headache and cold or cough). The most commonly reported symptoms were headache and feeling unwell or less fit (Figure 4). Of the 161 women who ever made a symptom diary entry (including n=4, 2.5%, who exited the study early), the median number of symptoms reported per woman was 17 (IQR 4-42; range 0-278); 17 (10.6%) women never reported a symptom.

Figure 4. Count of total reports of symptoms (excluding pregnancy complications).



Potential Arbovirus Infection Symptom Episodes

In total, 5.6% (9/161) of the participants reported a combination of symptoms that triggered a *potential arbovirus episode*, with 1 episode each at a median 26 (range 13-36) gestational weeks. The distribution of symptoms for these women, including those reported just before and after the episode, is presented in Multimedia Appendix 3. Of these 9 women, 4 (44%) reported some symptoms (headache, joint pain, or rash) the day before the notification of a potential arbovirus episode was triggered. In addition, of the 9 women, 2 (22%) had patterns of symptoms that resulted in both type 1 and 2 triggers and both reported additional symptoms (vomiting and cold symptoms or cough). Of the remaining 7 women, 5 (71%) reported one or more symptoms in the episode in addition to the relevant trigger symptoms, with cold symptoms or cough being the most common (reported by n=3, 43%; Multimedia Appendix 3).

No diagnosis of an arbovirus-related illness in relation to these episodes was made for any of the women (n=7) where detailed chart reviews were conducted. The women kept their regular antenatal appointments and had abdominopelvic ultrasound scans (anomaly scan and usually one more) to evaluate maternofetal health throughout pregnancy (which is the policy for evaluating any complications of arbovirus-related illness). Many (6/7, 86%) of the women had intercurrent illnesses in pregnancy, including cholestasis, urinary tract infections (several), ligamentous pain, otitis media, gastroesophageal reflux disease, and pre-eclampsia. All the women (n=9) delivered live births. None had SARS-CoV-2 detected in surveillance swabs.

COVID-19 Diagnoses

Of the 157 participants, 3 (1.9%) had laboratory-confirmed SARS-CoV-2 infection (nasopharyngeal swab positive for SARS-CoV-2 by polymerase chain reaction; all gave birth in late September and mid-October 2020). Of these 3 women, 1 (33%) reported headache over a 3-day period through the app, starting 8 days before the diagnosis; the case note review indicated additional symptoms not reported in the symptom diary (fever and cough), including anosmia and ageusia (not possible to report through the app); 1 (33%) reported headache on the day of the SARS-CoV-2 diagnosis and then 3 separate episodes of joint pain in the 10 days after the diagnosis; there was no record of these symptoms in her medical notes; and 1 (33%) reported feeling unwell or less fit for a 2-day period starting 4 days before the SARS-CoV-2 positive result; no symptoms were recorded in her medical notes.

Potential Pregnancy Complication Episodes

Of the 157 participants, 55 (35%) reported 114 pregnancy complication symptoms (painful uterine contractions or vaginal bleeding) up to, and including, the day of delivery or the end of pregnancy. Delivery occurred at term in 85% (47/55) of the women, at 34 to 36 weeks' gestation in 7% (4/55), and before 34 weeks in 4% (2/55). Of the 55 women, 2 (4%), miscarried. Among the 114 pregnancy complication symptoms, there were 59 (51.8%) reports of painful uterine contractions, which occurred at a median of 36 (IQR 5-108; range 0-199) days before delivery, and 39 (34.2%) reports of vaginal bleeding, which occurred at a median of 64 (IQR 6-127; range 0-174) days before delivery. Nearly a third of these symptom reports occurred within 7 days before delivery (35/114, 30.7%, for both). There were 8 episodes where both these symptoms were reported on the same day, in all cases in the 3 days before delivery.

Adverse Pregnancy Outcomes and Symptom Reporting

Of the 5 participants whose pregnancy ended in miscarriage or stillbirth, 2 (40%) did not complete their diary in the 2 weeks before the end of the pregnancy. Of the remaining 3 participants, 2 (67%) reported pregnancy complications during this period; in addition, all (3/3, 100%) reported one or more of the following symptoms: unwell or less fit, rash, headache, red eyes, muscle ache, cold symptoms or cough, and diarrhea (but without activating the potential arbovirus trigger).

Participants' Perceptions of the App

Of the 157 participants who used the app until delivery or the end of their pregnancy, 139 (89%) participated in the exit survey. Regarding their experience of participating in the study, 51.8% (71/137) rated it as excellent (5 on a scale of 1-5), with 35.8% (49/137), 11.7% (16/137), and 0% (0/137) giving ratings of 4, 3, and 2, respectively. Of the 137 respondents, only 1 (0.7%) reported her experience as disappointing (rating of 1). The women who reported an *excellent* experience were less likely to be poor adherers to symptom diary completion than other women, with 18% (13/71) having poor adherence compared with 33% (22/66) who had good adherence (P=.04).

Of the 139 respondents, 138 (99.3%) reported that the app was easy to understand and use. Overall, 52.5% (73/139) of the

women reported experiencing technical difficulties while using the app (ie, difficulties accessing Wi-Fi to complete their diary, difficulties accessing the internet on their mobile phone, and app freezing). However, experiencing technical difficulties was not associated with poor adherence (*P*=.90) or good adherence (*P*=.60; data not shown).

Of the 10 women who exited the study before delivery, 5 (50%) participated in the exit survey. With respect to their experience as participants in the study, of the 5 respondents, 2 (40%) rated their experience as excellent, 2 (40%) gave a rating of 4, and 1 (20%) gave a rating of 2. Of these 5 women, 5 (100%) reported that the app was easy to understand and 4 (80%) reported that it was easy to use. In addition, of the 4 women who answered the section on technical difficulties, 2 (50%) reported experiencing them while using the app.

Discussion

Principal Findings

This pilot study enrolled 173 pregnant women attending antenatal care at a university hospital clinic in Jamaica to evaluate their longitudinal engagement with a smartphone app to report symptoms potentially associated with arbovirus infections and pregnancy complications. Specifically, we wanted to understand whether it was feasible for pregnant women to report through the app the presence or absence of symptoms as well as symptom type on a daily basis over a prolonged period (from the second trimester to delivery).

This 1-year pilot study achieved a larger sample than pilot studies of other apps in pregnancy, for example, monitoring weight gain (2 studies recruiting ≤100 women [34,41]) or on reducing stress (29 women) [42], which is noteworthy given that our pilot overlapped with the beginning of the COVID-19 pandemic that forced many countries to stop or reduce ambulatory health care. We found that very few (10/173, 5.8%) women exited the study early and only 3.5% (6/173) never used the app after enrollment. This compares favorably with a study in Germany, which involved monthly web-based visits and surveys to digitally assess pregnancy-related symptoms and complications (including physical symptoms, depression, and anxiety), in which 7% of the women formally exited the study and 55% overall stopped using the app and made no further contact with the study team [43]. Thus, initial concerns that there might be a high rate of attrition because of the perceived burden of symptom reporting were not borne out in our pilot study.

ZIKApp was designed to *tolerate* delayed reporting of symptoms (ie, up to 7 days later), with the rationale that this cutoff would maximize reporting over the remainder of pregnancy while minimizing recall bias and allowing for any temporary problems with internet access. Adherence to symptom reporting through the app was good, with 79% (17,905/22,664) of all diaries completed. As shown in our figure of timing of symptom diary completion (Figure 3), overall patterns were fairly consistent with increasing duration of app use (considering successive 30-day periods) over the first 20 weeks after enrollment, with approximately a third of the diaries completed on the same day during this period (Multimedia Appendix 2). For the first 5 months of enrollment, only approximately 20% of the daily diaries were not completed at all (Multimedia Appendix 2).

However, patterns after 150 days of enrollment showed a decrease in diary completion overall, which may partly reflect women giving birth while they were within the 7-day window for daily diary completion. However, considering only those women who completed a diary, we showed that, as the duration of enrollment increased, there was a significant decline in the proportion completing their diary on the same day. The challenge of long-term adherence as a potential drawback of self-monitoring of symptoms through e-diaries has also been confirmed in other studies in pregnant and nonpregnant women that have reported less frequent diary reporting over calendar time [43-50].

A possible explanation for the good engagement of our participants could be the push notifications and reminders that were sent through our app as well as through email that could have been essential cues to action that encouraged our participants to complete their diaries or communicate with health care providers. This would be consistent with findings from a trial evaluating an mHealth intervention for healthy weight gain in pregnancy, which had good engagement in the intervention arm (only 9% attrition overall) in which SMS text messaging was a central feature [34]. The provision of phone credits in our study is also likely to have contributed because poor socioeconomic status has been associated with low engagement with an app for pregnant women elsewhere [51].

We found that older women (those aged >34 years) and women who had been to college or university were more likely to be good adherers to symptom diary reporting than younger women and those with less education, respectively. This finding supports other studies that have shown social variables such as older age and higher educational level to be significantly associated with sustained app use, including symptom diaries [43,44,47,52]. Other variables such as having a regular income, number of children, history of adverse pregnancy outcome, having a chronic condition, duration of app use, and reporting arbovirus infection symptoms were not found to be associated with good adherence in our study in contrast to others [44]. It was interesting to note that, although there was lower diary reporting overall and reduced timeliness of symptom reporting over the Christmas period (consistent with competing priorities during the busy holiday season), there was no association between other external factors such as technical difficulties using the app (eg, internet access and wireless connectivity) and adherence. In Jamaica, because the Christmas period is known to be a period of increased deliveries (including preterm births), the significant association with decreased app use is an important observation that should be considered for future implementation.

Of the 161 women who ever made a diary entry, 145 (90.1%) reported at least one symptom over a median of 22 weeks of app use. There was substantial heterogeneity in symptom reporting, with 1 in 10 women reporting no symptoms, whereas some (5/161, 3.1%) reported >200 symptoms. A potential challenge of using symptom diaries in pregnancy to identify signals of potential infections is the *noise* generated as a result

of common pregnancy-related symptoms. Triggers were selected to try to differentiate between *noise* and signals of true infections, but the pilot was not designed to evaluate the diagnostic performance of these triggers, which would require a different type of study to include diagnostic follow-up of all women with and without episode triggers. The most common symptoms reported in our study were headache, feeling unwell or less fit, and muscle ache, whereas the relatively low frequency of vomiting may reflect the fact that the first quartile of gestational age at enrollment was 15 weeks. It should also be noted that 45.5% (76/167) of the participants had at least one chronic condition, although participation in the study per se could have meant that the women were more sensitized to any physical symptoms they were experiencing and reported accordingly, whereas the normal ailments of pregnancy may have been exaggerated in other women, especially in light of the ongoing COVID-19 pandemic, potentially creating additional stressors for the pregnant woman.

This pilot study coincided with the start of the COVID-19 pandemic in Jamaica and as the study progressed, national surveillance for SARS-CoV-2 moved from passive to active surveillance, with pregnant women screened for symptomatic illness and asymptomatic involvement. Simultaneously, an unlinked serosurvey was performed in this antenatal population, which showed increasing SARS-CoV-2 seroprevalence, from 6.9% in September 2020 to 16.9% in October 2020 and 24% in November 2020; of the 37 pregnant women who tested SARS-CoV-2 immunoglobulin G antibody positive, only 3 were symptomatic [53]. The app development predated the emergence of SARS-CoV-2 but did capture some symptoms commonly reported with COVID-19. It was interesting to note that of the 3 pregnant women diagnosed with SARS-CoV-2, 2 (67%) reported symptoms approximately at the time of infection through the app that were not reported in the medical notes, whereas 1 (33%) did not report the fever and cough she experienced through the app (but delivered preterm within 4 days of experiencing symptoms and 2 days of testing positive).

The findings of this pilot suggest that despite most arbovirus infections being asymptomatic, screening pregnant women for relevant symptoms can improve case detection among those who are symptomatic [54], and that was part of the rationale for the development of ZIKApp. Our perspective is supported by recent evidence showing that a simple score based on clinical data and laboratory results provides a useful tool to help diagnose arbovirus infections [55].

To improve the usefulness of the app, it would possibly be more valuable to implement it during an epidemic period, rather than during a period of low prevalence of circulating arboviruses in the community, which was the case during this study. We also incorporated 2 symptoms (painful vaginal contractions and vaginal bleeding) that could signal important pregnancy complications (depending on timing).

We obtained feedback from participants about their experience of using the app in the exit survey, which had a high participation (139/157, 88.5%). More than half (71/137, 51.8%) of the responding women stated that their experience in the study was excellent and, consistently, they were less likely to be poor adherers to diary completion. Almost all (138/139, 99.3%) reported that the app was easy understand and use, despite a relatively high proportion experiencing technical difficulties at some point. This finding differs from a study in pregnant women with gestational diabetes that found that technological problems with the app had a negative impact on user satisfaction [56]. Overall, our exit survey results corroborate other mHealth intervention results that show ease of use and simplicity [33,44,56] and ease of navigation and ease of understanding [57,58] are key features of apps that can influence sustained adherence. Furthermore, our findings support previous evidence about the role of the perception of the product on intended app use by women [59].

Our finding of *long and strong* adherence to symptom diary reporting in pregnant women in this Jamaican setting provides important evidence to inform the potential applications of apps where symptom diaries and self-monitoring may be used as a tool for research (eg, to develop a better understanding of patient-reported outcomes), for surveillance and participatory epidemiology (eg, as seen for tracking COVID-19 or influenza), or for clinical purposes (eg, for remote health monitoring of low-risk pregnancies, the importance of which was highlighted by the disruption of traditional pathways for health care during the COVID-19 pandemic). Our experience in this pilot study showed that future implementation of this intervention for clinical use will require women to be linked directly to clinical care providers in different health services (eg, antenatal clinic, labor ward, emergency department, high-risk obstetric ward, and newborn services) rather than through research nurses to allow interpretation of symptoms and provision of appropriate clinical care in real time.

Limitations

We were unable to conduct the planned qualitative aspects of this study (ie, focus discussion groups with a sample of participants and staff) because of COVID-19–related restrictions, although we were able to obtain data on user perceptions through the exit survey. We were therefore unable to explore other potential facilitators (eg, the generic pregnancy-related messages embedded in the app, acceptability of answering questionnaires through the app, the role of the research nurses, and the perceived role of the app in the context of the COVID-19 pandemic) as well as barriers to app engagement (eg, whether feeling unwell had an impact on the timeliness of diary completion) or cues to action regarding health-seeking behaviors. Questions relating to facilitators and barriers to prolonged engagement with health-monitoring apps in pregnancy as well as linked health-related behaviors therefore require future research (eg, with mixed methods approaches).

Other limitations of this pilot study include the possibility of selection bias because only pregnant women who approached the nurses after the initial information provision in the clinic waiting room were recruited into the study. Consequently, our participants likely represent pregnant women using this antenatal clinic who were willing to take part in an mHealth intervention study. Likewise, women who were willing to participate were potentially more likely to adhere to diary reporting than nonparticipants, potentially resulting in social desirability

response bias, particularly with respect to the exit survey. In addition, the results of this pilot study may not be generalizable to all pregnant women in Jamaica because of potential sociodemographic differences between our study population and the general antenatal population; for example, there were higher proportions of women who were nulliparous and who had received tertiary education in our study than in another ZIKAction consortium study that enrolled pregnant women from across the Kingston, Jamaica, metropolitan area [60]. Further research could recruit women from community-based antenatal clinics to capture a more socioeconomically diverse sample.

Conclusions

We have demonstrated the feasibility and usability of ZIKApp in an arbovirus-endemic region, showing that most pregnant women were able to adhere to symptom reporting through the app for a prolonged period and supporting its future development to contribute to surveillance and diagnosis of, and communication about, arbovirus infections in pregnancy. The findings also indicate that such an app shows promise for future development and implementation by direct treatment and care teams to optimize obstetric care. For any of these potential uses, further research will be required, for example, to explore how app use could be linked to sampling (including self-sampling) and testing within a surveillance program while adapting the mobile app interface, features, and messages to the appropriate cultural context.

Acknowledgments

The authors thank all the women and their families who, the authors hope, would have benefited positively and holistically by their active participation in this study. The authors are also thankful to all the staff of the antenatal clinics, labor ward, and postnatal wards, as well as the medical records office, who were very helpful in implementing all aspects of this study, and Ms ShreeAnn Simon, the ZIKApp study administrator. Similarly, the authors want to acknowledge and express gratitude to Francesca Viero, Giorgia Dalla Valle, Rachel Knowles, Peter Tanghe, Andrea Oletto, Thomas Byrne, Jacqueline McLean, and Trudy-Ann Blair-Wallace. This work was supported by the European Union's Horizon 2020 research and innovation program (734857; ZIKAction).

The ZIKAction Consortium collaborators are as follows: Tiziana Grossele, Francesca Viero, Giorgia Dalla Valle, Andrea Oletto, Rebecca Lundin, Daniel Lang, Andrea Garbutt, Roxanne Melbourne-Chambers, Orville Morgan, Alexander Onyonyor, Paul Mitchell, Leroy Campbell, Marissa Mayne, Karen Webster-Kerr, Joshua Anzinger, Georgiana Gordon-Strachan, Sheree Mair, Keisha Francis, Yakima Phillips, Chadwic Mears, Tracia James-Powell, Yanique Brown, Clayton Kelly, Lizette Mowatt, Judy Tapper, Amanda Rousseau, Brittiana Brown, Sofier Scott, Marlyn Grindley, Donna Kamicka, John Lindo, ShreeAnn Simon, Thomas Byrne, Moira Spyer, Eleni Nastouli, AE Ades, Fabiana Gordon, J Glenn Morris Jr., Rigan Louis, Valery E M Beau De Rochars, John Lednicky, Maha Elbadry, Derek Adam Cummings, Cyrus Saleem, Mahbubul Alam, Jean Claude Cadet, Chamblain Richard, Sonese Chavannes, Medjina Syndic, Mathieu Kern, Merisier Delson, Maureen Long, Erika Schwarz, Ana Bispo, David Brown, Patricia Brasil, Antoni Soriano-Arandes, Ana Alarcón, Elena Sulleiro, Isadora Siqueira, Rosa Bologna, Griselda Berberian, Greta Mino, Francesco Bonfante, Eva Mazzetto, Federico Martinon-Torres, Irene Rivero Calle, Trudie Lang, Bonnie Baker, Liam Boggs, Marco Safadi, Flavia Almeda, Daniel Jarovsky, and Alessandra Soriano.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Symptoms included in ZIKApp and the start and end of arbovirus episode triggers.
[DOCX File , 21 KB - formative_v6i7e34423_app1.docx]

Multimedia Appendix 2
Timing of symptom diary completion, by 30-day period from app initiation.
[DOCX File , 19 KB - formative_v6i7e34423_app2.docx]

Multimedia Appendix 3
Type and duration of symptoms reported by women with arbovirus trigger episodes.
[DOCX File , 81 KB - formative_v6i7e34423_app3.docx]

References

1. Campos GS, Bandeira AC, Sardi SI. Zika virus outbreak, Bahia, Brazil. Emerg Infect Dis 2015 Oct;21(10):1885-1886 [FREE Full text] [doi: 10.3201/eid2110.150847] [Medline: 26401719]

2. Petersen LR, Jamieson DJ, Powers AM, Honein MA. Zika virus. N Engl J Med 2016 Apr 21;374(16):1552-1563. [doi: 10.1056/NEJMra1602113] [Medline: 27028561]
3. Christie CD, Melbourne-Chambers R, Ennevor J, Young-Peart S, Buchanan T, Scott-Brown P, et al. Chikungunya in Jamaica – public health effects and clinical features in children. West Indian Med J 2016 Nov 10;65(3):431-437. [doi: 10.7727/wimj.2016.529]
4. Paixao ES, Leong WY, Rodrigues LC, Wilder-Smith A. Asymptomatic prenatal Zika virus infection and congenital Zika syndrome. Open Forum Infect Dis 2018 Apr 7;5(4):ofy073 [FREE Full text] [doi: 10.1093/ofid/ofy073] [Medline: 29732381]
5. Thompson R, Martin Del Campo J, Constenla D. A review of the economic evidence of Aedes-borne arboviruses and Aedes-borne arboviral disease prevention and control strategies. Expert Rev Vaccines 2020 Feb;19(2):143-162. [doi: 10.1080/14760584.2020.1733419] [Medline: 32077343]
6. Rasmussen SA, Jamieson DJ, Honein MA, Petersen LR. Zika virus and birth defects--reviewing the evidence for causality. N Engl J Med 2016 May 19;374(20):1981-1987. [doi: 10.1056/NEJMsr1604338] [Medline: 27074377]
7. Charlier C, Beaudoin MC, Couderc T, Lortholary O, Lecuit M. Arboviruses and pregnancy: maternal, fetal, and neonatal effects. Lancet Child Adolesc Health 2017 Oct;1(2):134-146. [doi: 10.1016/S2352-4642(17)30021-4] [Medline: 30169203]
8. Paixão ES, Campbell OM, Teixeira MG, Costa MC, Harron K, Barreto ML, et al. Dengue during pregnancy and live birth outcomes: a cohort of linked data from Brazil. BMJ Open 2019 Jul 24;9(7):e023529 [FREE Full text] [doi: 10.1136/bmjopen-2018-023529] [Medline: 31345962]
9. Ades AE, Soriano-Arandes A, Alarcon A, Bonfante F, Thorne C, Peckham CS, et al. Vertical transmission of Zika virus and its outcomes: a Bayesian synthesis of prospective studies. Lancet Infect Dis 2021 Apr;21(4):537-545 [FREE Full text] [doi: 10.1016/S1473-3099(20)30432-1] [Medline: 33068528]
10. de St Maurice A, Ervin E, Chu A. Ebola, Dengue, Chikungunya, and Zika infections in neonates and infants. Clin Perinatol 2021 Jun;48(2):311-329. [doi: 10.1016/j.clp.2021.03.006] [Medline: 34030816]
11. Rankine-Mullings A, Reid ME, Moo Sang M, Richards-Dawson M, Knight-Madden JM. A retrospective analysis of the significance of haemoglobin SS and SC in disease outcome in patients with sickle cell disease and dengue fever. EBioMedicine 2015 Aug;2(8):937-941 [FREE Full text] [doi: 10.1016/j.ebiom.2015.07.002] [Medline: 26425701]
12. Davidson TD, Vickers I, Christie CD. Outcome of dengue in hospitalized Jamaican children. West Indian Med J 2017 Feb 10;65(3):442-449. [doi: 10.7727/wimj.2016.525]
13. Lue AM, Richards-Dawson MA, Gordon-Strachan G, Kodilinye SM, Dunkley-Thompson JA, James-Powell TD, et al. Severity and outcomes of dengue in hospitalized Jamaican children in 2018-2019 during an epidemic surge in the Americas. Front Med (Lausanne) 2022 Jun 21;9:889998 [FREE Full text] [doi: 10.3389/fmed.2022.889998] [Medline: 35801209]
14. Webster-Kerr K. Trends in dengue: The Jamaican Experience over 40 years. In: Proceedings of the 3rd International ZIKA and Aedes International Conference. 2020 Presented at: ZIKAction '20; February 13-16, 2020; Washington, DC, USA.
15. Duncan J, Gordon-Johnson KA, Tulloch-Reid MK, Cunningham-Myrie C, Ernst K, McMorris N, et al. Chikungunya: important lessons from the Jamaican experience. Rev Panam Salud Publica 2017 Aug 21;41:e60 [FREE Full text] [Medline: 28902273]
16. Evans-Gilbert T. Chikungunya and neonatal immunity: fatal vertically transmitted Chikungunya infection. Am J Trop Med Hyg 2017 Apr;96(4):913-915 [FREE Full text] [doi: 10.4269/ajtmh.16-0491] [Medline: 28167590]
17. Pham PN, Williams LT, Obot U, Padilla LA, Aung M, Akinyemiju TF, et al. Epidemiology of Chikungunya fever outbreak in Western Jamaica during July-December 2014. Res Rep Trop Med 2017;8:7-16 [FREE Full text] [doi: 10.2147/rrtm.s122032] [Medline: 29375245]
18. Freitas AR, Gérardin P, Kassar L, Donalisio MR. Excess deaths associated with the 2014 chikungunya epidemic in Jamaica. Pathog Glob Health 2019 Feb;113(1):27-31 [FREE Full text] [doi: 10.1080/20477724.2019.1574111] [Medline: 30714498]
19. Anzinger JJ, Mears CD, Ades A, Francis K, Phillips Y, Leys YE, ZIKAction Consortium1,2. Antenatal seroprevalence of zika and chikungunya viruses, Kingston metropolitan area, Jamaica, 2017-2019. Emerg Infect Dis 2022 Feb;28(2):473-475 [FREE Full text] [doi: 10.3201/eid2802.211849] [Medline: 35076369]
20. Christie C, Giaquinto C. Unravelling the paediatric and perinatal Zika virus epidemic through population-based research. West Indian Med J 2016 Sep 19;65(1):239-242. [doi: 10.7727/wimj.2016.454] [Medline: 28375541]
21. Melbourne-Chambers R, Christie CD, Greenaway E, Bullock R. Acute paralysis and neuro-inflammation in Jamaican children during Zika virus and Dengue epidemics of 2016. West Indian Med J 2016 Nov 10;65(3):425-430. [doi: 10.7727/wimj.2016.526]
22. Webster-Kerr KR, Christie C, Grant A, Chin D, Burrowes H, Clarke K, et al. Emergence of Zika virus epidemic and the national response in Jamaica. West Indian Med J 2016 Sep 26;65(1):243-249. [doi: 10.7727/wimj.2016.488] [Medline: 28375542]
23. James-Powell T, Brown Y, Christie CD, Melbourne-Chambers R, Moore JT, Morgan O, et al. Trends of microcephaly and severe arthrogryposis in three urban hospitals following the Zika, Chikungunya and Dengue fever epidemics of 2016 in Jamaica. West Indian Med J 2017 Jul 14;66:10-19. [doi: 10.7727/wimj.2017.124]
24. John Leon Singh H, Couch D, Yap K. Mobile health apps that help with COVID-19 management: scoping review. JMIR Nurs 2020 Aug 6;3(1):e20596 [FREE Full text] [doi: 10.2196/20596] [Medline: 32897271]

25. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med 2020 Jul;26(7):1037-1040 [FREE Full text] [doi: 10.1038/s41591-020-0916-2] [Medline: 32393804]
26. Smith AC, Thomas E, Snoswell CL, Haydon H, Mehrotra A, Clemensen J, et al. Telehealth for global emergencies: implications for coronavirus disease 2019 (COVID-19). J Telemed Telecare 2020 Jun;26(5):309-313 [FREE Full text] [doi: 10.1177/1357633X20916567] [Medline: 32196391]
27. Lichtman A, Greenblatt E, Malenfant J, Kuo A. Universal symptom monitoring to address presenteeism in healthcare workers. Am J Infect Control 2021 Aug;49(8):1021-1023 [FREE Full text] [doi: 10.1016/j.ajic.2021.02.009] [Medline: 34294381]
28. Nair P, Bhaskaran H. The emerging interface of healthcare system and mobile communication technologies. Health Technol 2014 Sep 20;4(4):337-343. [doi: 10.1007/s12553-014-0091-x]
29. Rodriguez-Valero N, Luengo Oroz M, Cuadrado Sanchez D, Vladimirov A, Espriu M, Vera I, et al. Mobile based surveillance platform for detecting Zika virus among Spanish Delegates attending the Rio de Janeiro Olympic Games. PLoS One 2018 Aug 22;13(8):e0201943 [FREE Full text] [doi: 10.1371/journal.pone.0201943] [Medline: 30133492]
30. Ocampo CB, Mina NJ, Echavarria MI, Acuña M, Caballero A, Navarro A, et al. VECTOS: an integrated system for monitoring risk factors associated with urban arbovirus transmission. Glob Health Sci Pract 2019 Mar 22;7(1):128-137 [FREE Full text] [doi: 10.9745/GHSP-D-18-00300] [Medline: 30926741]
31. Rodríguez S, Sanz AM, Llano G, Navarro A, Parra-Lara LG, Krystosik AR, et al. Acceptability and usability of a mobile application for management and surveillance of vector-borne diseases in Colombia: an implementation study. PLoS One 2020 May 29;15(5):e0233269 [FREE Full text] [doi: 10.1371/journal.pone.0233269] [Medline: 32469894]
32. Carrillo MA, Kroeger A, Cardenas Sanchez R, Diaz Monsalve S, Runge-Ranzinger S. The use of mobile phones for the prevention and control of arboviral diseases: a scoping review. BMC Public Health 2021 Jan 09;21(1):110 [FREE Full text] [doi: 10.1186/s12889-020-10126-4] [Medline: 33422034]
33. de Vries ST, Wong L, Sutcliffe A, Houÿez F, Ruiz CL, Mol PG, IMI Web-RADR Work Package 3b Consortium. Factors influencing the use of a mobile app for reporting adverse drug reactions and receiving safety information: a qualitative study. Drug Saf 2017 May;40(5):443-455 [FREE Full text] [doi: 10.1007/s40264-016-0494-x] [Medline: 28035492]
34. Willcox JC, Wilkinson SA, Lappas M, Ball K, Crawford D, McCarthy EA, et al. A mobile health intervention promoting healthy gestational weight gain for women entering pregnancy at a high body mass index: the txt4two pilot randomised controlled trial. BJOG 2017 Oct;124(11):1718-1728. [doi: 10.1111/1471-0528.14552] [Medline: 28220604]
35. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. 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]
36. Parsa S, Khajouei R, Baneshi MR, Aali BS. Improving the knowledge of pregnant women using a pre-eclampsia app: a controlled before and after study. Int J Med Inform 2019 May;125:86-90. [doi: 10.1016/j.ijmedinf.2019.03.001] [Medline: 30914185]
37. Saronga NJ, Burrows T, Collins CE, Ashman AM, Rollo ME. mHealth interventions targeting pregnancy intakes in low and lower-middle income countries: systematic review. Matern Child Nutr 2019 Apr;15(2):e12777 [FREE Full text] [doi: 10.1111/mcn.12777] [Medline: 30609297]
38. Sandborg J, Söderström E, Henriksson P, Bendtsen M, Henström M, Leppänen MJ, et al. Effectiveness of a smartphone app to promote healthy weight gain, diet, and physical activity during pregnancy (HealthyMoms): randomized controlled trial. JMIR Mhealth Uhealth 2021 Mar 11;9(3):e26091 [FREE Full text] [doi: 10.2196/26091] [Medline: 33704075]
39. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
40. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
41. Redman LM, Gilmore LA, Breaux J, Thomas DM, Elkind-Hirsch K, Stewart T, et al. Effectiveness of SmartMoms, a novel eHealth intervention for management of gestational weight gain: randomized controlled pilot trial. JMIR Mhealth Uhealth 2017 Sep 13;5(9):e133 [FREE Full text] [doi: 10.2196/mhealth.8228] [Medline: 28903892]
42. Dennis-Tiwary TA, Denefrio S, Gelber S. Salutary effects of an attention bias modification mobile application on biobehavioral measures of stress and anxiety during pregnancy. Biol Psychol 2017 Jul;127:148-156 [FREE Full text] [doi: 10.1016/j.biopsycho.2017.05.003] [Medline: 28478138]
43. Brusniak K, Arndt HM, Feisst M, Haßdenteufel K, Matthies LM, Deutsch TM, et al. Challenges in acceptance and compliance in digital health assessments during pregnancy: prospective cohort study. JMIR Mhealth Uhealth 2020 Oct 14;8(10):e17377 [FREE Full text] [doi: 10.2196/17377] [Medline: 33052134]

44. Mattila E, Orsama AL, Ahtinen A, Hopsu L, Leino T, Korhonen I. Personal health technologies in employee health promotion: usage activity, usefulness, and health-related outcomes in a 1-year randomized controlled trial. JMIR Mhealth Uhealth 2013 Jul 29;1(2):e16 [FREE Full text] [doi: 10.2196/mhealth.2557] [Medline: 25098385]
45. Semple JL, Sharpe S, Murnaghan ML, Theodoropoulos J, Metcalfe KA. Using a mobile app for monitoring post-operative quality of recovery of patients at home: a feasibility study. JMIR Mhealth Uhealth 2015 Feb 12;3(1):e18 [FREE Full text] [doi: 10.2196/mhealth.3929] [Medline: 25679749]
46. Ledford CJ, Womack JJ, Rider HA, Seehusen AB, Conner SJ, Lauters RA, et al. Unexpected effects of a system-distributed mobile application in maternity care: a randomized controlled trial. Health Educ Behav 2018 Jun;45(3):323-330. [doi: 10.1177/1090198117732110] [Medline: 28918669]
47. Seng EK, Prieto P, Boucher G, Vives-Mestres M. Anxiety, incentives, and adherence to self-monitoring on a mobile health platform: a naturalistic longitudinal cohort study in people with headache. Headache 2018 Nov;58(10):1541-1555 [FREE Full text] [doi: 10.1111/head.13422] [Medline: 30334248]
48. Carter J, Sandall J, Shennan AH, Tribe RM. Mobile phone apps for clinical decision support in pregnancy: a scoping review. BMC Med Inform Decis Mak 2019 Nov 12;19(1):219 [FREE Full text] [doi: 10.1186/s12911-019-0954-1] [Medline: 31718627]
49. Wang N, Deng Z, Wen LM, Ding Y, He G. Understanding the use of smartphone apps for health information among pregnant Chinese women: mixed methods study. JMIR Mhealth Uhealth 2019 Jun 18;7(6):e12631 [FREE Full text] [doi: 10.2196/12631] [Medline: 31215516]
50. Di Fraia M, Tripodi S, Arasi S, Dramburg S, Castelli S, Villalta D, et al. Adherence to prescribed e-diary recording by patients with seasonal allergic rhinitis: observational study. J Med Internet Res 2020 Mar 16;22(3):e16642 [FREE Full text] [doi: 10.2196/16642] [Medline: 32175909]
51. Dalton JA, Rodger D, Wilmore M, Humphreys S, Skuse A, Roberts CT, et al. The Health-e Babies app for antenatal education: feasibility for socially disadvantaged women. PLoS One 2018 May 16;13(5):e0194337 [FREE Full text] [doi: 10.1371/journal.pone.0194337] [Medline: 29768407]
52. Morren M, van Dulmen S, Ouwerkerk J, Bensing J. Compliance with momentary pain measurement using electronic diaries: a systematic review. Eur J Pain 2009 Apr;13(4):354-365. [doi: 10.1016/j.ejpain.2008.05.010] [Medline: 18603458]
53. Chisolm T, Walker JP, Leys Y, Butterfield TR, Medley C, Thompson TK, et al. SARS-CoV-2 seroprevalence among antenatal clinic attendees in Kingston, Jamaica, September-November 2020. medRxiv 2021 Feb 11. [doi: 10.1101/2021.02.08.21251367]
54. Naik S, Robinson ML, Alexander M, Chandanwale A, Sambarey P, Kinikar A, et al. Intensified short symptom screening program for dengue infection during pregnancy, India. Emerg Infect Dis 2020 Apr;26(4):738-743 [FREE Full text] [doi: 10.3201/eid2604.191476] [Medline: 32186485]
55. Vellere I, Lagi F, Spinicci M, Mantella A, Mantengoli E, Corti G, et al. Arbo-Score: a rapid score for early identification of patients with imported arbovirosis caused by Dengue, Chikungunya and Zika virus. Microorganisms 2020 Nov 04;8(11):1731 [FREE Full text] [doi: 10.3390/microorganisms8111731] [Medline: 33158274]
56. Varnfield M, Redd C, Stoney RM, Higgins L, Scolari N, Warwick R, et al. M♡THer, an mHealth system to support women with gestational diabetes mellitus: feasibility and acceptability study. Diabetes Technol Ther 2021 May;23(5):358-366 [FREE Full text] [doi: 10.1089/dia.2020.0509] [Medline: 33210954]
57. Goetz M, Müller M, Matthies LM, Hansen J, Doster A, Szabo A, et al. Perceptions of patient engagement applications during pregnancy: a qualitative assessment of the patient's perspective. JMIR Mhealth Uhealth 2017 May 26;5(5):e73 [FREE Full text] [doi: 10.2196/mhealth.7040] [Medline: 28550005]
58. Lau Y, Cheng LJ, Chi C, Tsai C, Ong KW, Ho-Lim SS, et al. Development of a healthy lifestyle mobile app for overweight pregnant women: qualitative study. JMIR Mhealth Uhealth 2018 Apr 23;6(4):e91 [FREE Full text] [doi: 10.2196/mhealth.9718] [Medline: 29685868]
59. Lin YH, Guo JL, Hsu HP, Yang LS, Fu YL, Huang CM. Does "hospital loyalty" matter? Factors related to the intention of using a mobile app. Patient Prefer Adherence 2019;13:1283-1294 [FREE Full text] [doi: 10.2147/PPA.S207031] [Medline: 31534315]
60. Christie CD. Overview and preliminary results of the Zika vertical transmission study in Jamaica. In: Oral Presentation at 3rd International Conference on Zika Virus and Aedes Related Infections. 2020 Feb 16 Presented at: ZIKAction '20; February 13-16, 2020; Washington, DC, USA.

Abbreviations

CHIKV: chikungunya virus
DENV: dengue virus
mHealth: mobile health
REDCap: Research Electronic Data Capture
SOP: standard operating procedure
UHWI: University Hospital of the West Indies

ZIKV: Zika virus

Edited by A Mavragani; submitted 22.10.21; peer-reviewed by P Cook, S Badawy; comments to author 10.02.22; revised version received 23.02.22; accepted 01.03.22; published 27.07.22.

Please cite as:
Ruiz-Burga E, Bruijning-Verhagen P, Palmer P, Sandcroft A, Fernandes G, de Hoog M, Bryan L, Pierre R, Bailey H, Giaquinto C, Thorne C, Christie CDC, ZIKAction Consortium
Detection of Potential Arbovirus Infections and Pregnancy Complications in Pregnant Women in Jamaica Using a Smartphone App (ZIKApp): Pilot Evaluation Study
JMIR Form Res 2022;6(7):e34423
URL: https://formative.jmir.org/2022/7/e34423
doi:10.2196/34423
PMID:35896029

Original Paper

Telemedicine Preparedness Among Older Adults With Chronic Illness: Survey of Primary Care Patients

Kelsey Ufholz[1], PhD; Amy Sheon[2], MPH, PhD; Daksh Bhargava[1], MSc; Goutham Rao[1,3], MD

[1]Department of Family Medicine and Community Health, School of Medicine, Case Western Reserve University, Cleveland, OH, United States

[2]Public Health Innovators LLC, Cleveland, OH, United States

[3]Department of Family Medicine, University Hospitals, Cleveland, OH, United States

Corresponding Author:
Kelsey Ufholz, PhD
Department of Family Medicine and Community Health
School of Medicine
Case Western Reserve University
Suite 1056
11100 Euclid Avenue
Cleveland, OH, 44016
United States
Phone: 1 2162862212
Email: keu8@case.edu

Abstract

Background: Older adults are a high priority for telemedicine given their elevated COVID-19 risk and need for frequent provider contact to manage chronic illnesses. It seems that many older adults now use smartphones but few studies have examined their overall readiness for telemedicine.

Objective: The aim of this study is to survey older primary care patients about their telemedicine preparedness, including internet usage, internet-capable devices, telemedicine experiences and concerns, and perceived barriers. Results were used to inform a telemedicine preparedness training program.

Methods: Community-dwelling older adult patients (aged 65-81 years; N=30) with a chronic health condition that could be managed remotely who were present at a family medicine clinic that primarily serves an urban African American population for a prescheduled in-person appointment were asked to complete a brief survey written for this study. Data were collected February-June 2021 at a large, urban, Midwestern hospital. To minimize patient burden, the survey was limited to 10 questions, focused on the most critical topics.

Results: Most participants (21/30, 70%) reported having a device that could be used for telemedicine and using the internet. However, about half had only a single connected device, and messaging and video calling were the most commonly used applications. Few used email and none used online shopping or banking. Only 7 patients had had telemedicine appointments. Telemedicine users were younger than nonusers and used more internet functions than nonusers. Only 2 people reported problems with their telemedicine visits (technology and privacy). Nearly all respondents recognized avoiding travel and COVID-19 exposure as telemedicine benefits. The most common concerns were loss of the doctor-patient connection and inability to be examined.

Conclusions: Most older adults reported having devices that could be used for telemedicine, but their internet use patterns did not confirm the adequacy of their devices or skills for telemedicine. Doctor-patient conversations could be helpful in addressing telemedicine concerns but device and skill gaps must be addressed as well.

(JMIR Form Res 2022;6(7):e35028) doi:10.2196/35028

KEYWORDS

telemedicine; seniors; primary care; chronic illness; health equities; telehealth; older adult; healthcare; health care; digital health; senior health

Introduction

Background

Since the COVID-19 pandemic began, telemedicine appointments have replaced many in-person health care visits [1,2]. However, older people are less likely to participate in telemedicine, preferring in-person care or foregoing care altogether [3-6]. With a high prevalence of chronic conditions and vulnerability to COVID-19 morbidity and mortality through exposure to others in health care environments [1-4], promoting telemedicine use among older adults should be a high priority.

Older Adults' Barriers to Telemedicine

Older adults face significant barriers to participation in telemedicine, including limited access to the internet and devices suitable for telemedicine [7]. Older adults may also lack digital skills or have visual, auditory, and tactile difficulties with telemedicine, or be uncertain about whether or when to use it. To inform our plans for offering telemedicine training to older adults presenting to an outpatient family medicine teaching clinic that serves predominantly African American, economically disadvantaged adults with chronic illness in Cleveland, Ohio, we administered a survey to learn about their telemedicine readiness, and telemedicine barriers and facilitators.

Methods

Participants

We sought to recruit 30 participants, the minimum recommended sample size for estimating univariate averages, and a number thought adequate to identify common patient journeys that would guide our plans for telemedicine training [8,9]. Participants were recommended to this convenience sample by primary care providers who were familiar with their medical history and the study criteria. Inclusion criteria included age ≥65 years and having a chronic health condition (diabetes, hypertension, arthritis, etc) that could be managed remotely. Patients with known dementia, residence in a long-term care facility, and presenting with an acute condition requiring in-person care (eg, fall or chronic obstructive pulmonary disease exacerbation) were ineligible.

Survey Instrument

Because existing surveys tend to lack the specificity needed to determine the adequacy of devices and skills for telemedicine, we designed and pretested a new survey instrument based on a review of the literature, and input from our primary care providers and a digital equity expert (Multimedia Appendix 1). Because we were not offering compensation, we minimized patient burden by limiting the survey to 10 questions. Topics included demographics, experience using telemedicine, problems and perceived barriers, ownership of telemedicine-ready devices(s), and use of various internet functions.

Procedures

Patients present at an in-person primary care visit for issues that could be accomplished remotely were approached by a research assistant to complete the survey between February and June 2021. Data were collected on paper, with a research assistant available to read the survey questions and record responses if needed. The research assistant entered anonymous responses into a REDCap database to protect patient privacy. Descriptive statistics were calculated to inform our telemedicine readiness training plans. Chi-square tests were used to test for statistical significance, α=.05.

Ethical Considerations

University Hospitals' Institutional Review Board determined the study (2021611) to be no more than minimal risk and granted expedited approval. Written informed consent was not required but prior to beginning the study, participants received written information informing them that they were invited to participate in a voluntary research study and were free to decline participation.

Results

Devices and Internet Usage

Of 30 respondents, 25 (83%) said they had devices that could be used for a telemedicine visit and that they went on the internet, but just 7 of 30 (23%) had had telemedicine visits. However, few patients had advanced devices (iPhones, desktops, laptops, or tablets) that are best suited to telemedicine. In addition, 14 of 30 respondents (47%) had only a single device that was not an iOS-based mobile device (Table 1) and may have had limited videoconferencing ability. All participants with devices said they used them for "messaging on the internet," but this was the only function used by 12 of 30 respondents (40%). No one used the internet for banking or shopping, and few used internet functions commonly needed for telemedicine (email: 7 respondents, 23%; video calling: 9 respondents, 30%) (Table 1).

Table 1. Survey participant demographics and telemedicine readiness.

Demographics and telemedicine readiness	Participants, n	Participants, %
Age (years)[a]		
65-74	24	80
75-80	5	17
80-89	1	3
Chronic conditions		
1	5	17
2	13	43
3	10	33
≥4	2	7
Hypertension	19	63
Diabetes	18	60
Device ownership		
iPhone	5	17
Desktop, tablet, laptop	6	20
Other smartphone only	14	47
0	5	17
1	21	70
≥2	4	13
Internet use		
Telemedicine visit	7	23
Video calls	9	30
Entertainment	5	17
Email	4	13
Messaging only	12	40
Work, banking, shopping	0	0
No internet functions	5	17
1 internet function	12	40
2 internet functions	8	27
3 internet functions	5	17
Telemedicine advantages		
No travel	29	97
Avoid COVID-19	25	83
Telemedicine disadvantages		
Doctor cannot examine me	7	23
Loss of personal connection	10	33
Inferior care quality	4	13
Lack of privacy	7	23
Other disadvantage	3	10

[a]Mean age 70.8 (SD 4.3) years; range 65-81 years.

Telemedicine Experiences and Perceptions

Of 30 respondents, 7 (23%) had had a telemedicine appointment. Participants who owned a computer or iPhone were more likely to have had a telemedicine visit than others (Figure 1A; χ^2_1=9.5; *P*=.002), as were participants who had used the internet for email or functions other than messaging (Figure 1B; χ^2_1=11.9; *P*<.001). All but one respondent who had a telemedicine visit had an iPhone or a computer and used internet functions other than messaging. Participants with iPhones or computers used their devices for a broader range of tasks (Table 2; χ^2_3=18.0; *P*<.001), endorsed fewer telemedicine disadvantages (χ^2_3=11.9; *P*=.008), and were more likely to indicate interest in future telemedicine visits (χ^2_1=5.7; *P*=.02) than were patients with other types of mobile devices or no devices at all. Telemedicine attitudes of patients who used email or other internet functions were similar to those with advanced devices. Loss of connection with their doctor was the most commonly endorsed telemedicine disadvantage (10/30, 33%) followed by concerns about exam privacy and quality (7/30, 23%). Patients who were aged 65-70 years were more likely to have an iPhone or other computer (χ^2_1=10.5; *P*=.001; Figure 2A), and were more likely to have had a telemedicine visit (χ^2_1=6.7; *P*=.01; Figure 2B) and to have used internet functions other than messaging (χ^2_1=15.9; *P*<.001; Figure 2C) compared with patients aged 70 years and older.

Figure 1. Comparisons of likelihood to have had a telemedicine visit by device ownership and device usage.

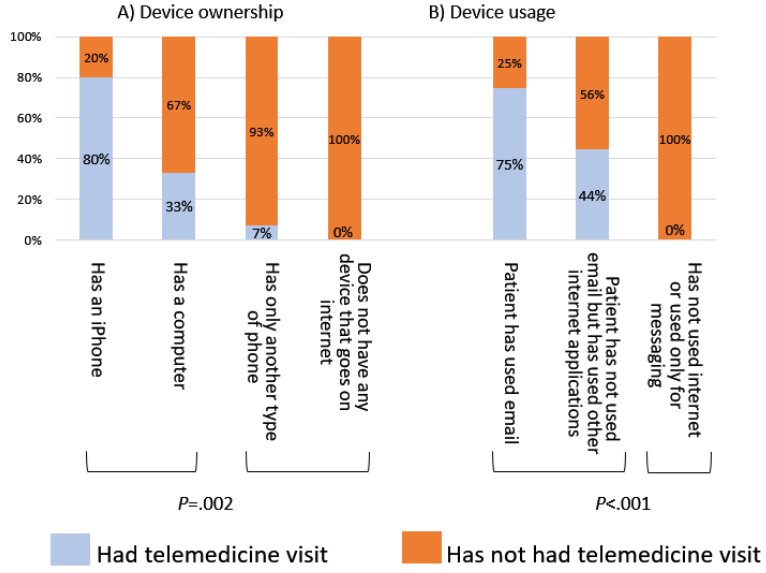
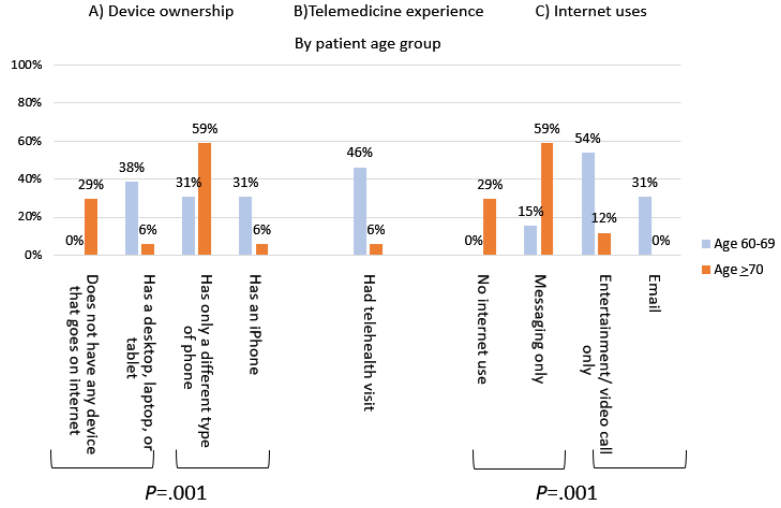


Table 2. Internet uses and telemedicine attitudes by device type.

Devices and functions used	Internet uses		Telemedicine attitudes	
	Mean number of ways participants use the internet	Messaging only, n (%)	Mean number of telemedicine disadvantages	Interest in future telemedicine visit, n (%)
Type of device				
iPhone	2.4	0 (0)	0.4	5 (100)
iPad or computer	2.3	1 (8)	0.7	4 (67)
Other mobile only	1.2	11 (92)	1.4	6 (43)
None	0.0	0 (0)	1.2	1 (20)
Internet functions used				
Used email	N/A[a]	N/A	0.5	4 (100)
No email but used entertainment or video calling	N/A	N/A	0.7	7 (77.8)
Used messaging only	N/A	N/A	1.4	4 (33.3)
No internet use	N/A	N/A	1.2	1 (20)

[a]N/A: not applicable.

Figure 2. Comparison of device ownership, telemedicine experience, and internet uses by age group.



Discussion

Principal Findings

This small survey revealed significant gaps in telemedicine readiness among older adults who said they had devices that could be used for telemedicine and that they went online. No patients used key internet functions needed for staying safe during the COVID-19 pandemic, and few used internet applications that required the skills needed for accessing telemedicine. Few patients had devices that are optimal for older adults using telemedicine. Patients with more advanced devices used more internet functions and had more telemedicine experience and more favorable attitudes than others. Our results confirm previous studies [10-12] showing generally lower technological proficiency among older adults and some concerns about participating in telemedicine. However, our study is novel in pointing to subtle dimensions of telemedicine readiness that warrant further study—device capacity and use of internet in ways that build skills needed for telemedicine such as email and video calling. Before training older adults to use telemedicine, it is important to ensure that they have the devices, basic digital skills, and connectivity needed for telemedicine. Screening for readiness may require nuanced assessment regarding specific device capacity and skills.

Limitations and Future Directions

Because of the survey's limited nature, other important topics, such as home internet access and interest in digital skills training, could not be addressed. Results may not be generalizable to other contexts, such as specialty clinics or rural areas. Participants present in the clinic may be different from those not seeking care, which could bias our results. Larger studies are needed to confirm our results and apply multivariate analysis to understand the relationships among age, device quality, internet skills, and telemedicine attitudes. Development of validated scales of telemedicine readiness as well as telemedicine training to complement in-person care can help health systems offer precision-matched interventions to address barriers, facilitate increased adoption, and generally improve patients' overall access to primary care and engagement with their primary care provider.

Acknowledgments

The authors would like to thank the Family Medicine department staff and our patients for their assistance with completion of this project.

Conflicts of Interest

The Case Western Reserve University Department of Family Medicine provided funds to AS for expert consultation around digital health equity. The other authors declared no other conflicts of interest to report.

Multimedia Appendix 1
Survey questions.
[PNG File , 241 KB - formative_v6i7e35028_app1.png]

References

1. Wosik J, Fudim M, Cameron B, Gellad ZF, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. J Am Med Inform Assoc 2020 Jun 01;27(6):957-962 [FREE Full text] [doi: 10.1093/jamia/ocaa067] [Medline: 32311034]

2. Calton B, Abedini N, Fratkin M. Telemedicine in the time of coronavirus. J Pain Symptom Manage 2020 Jul;60(1):e12-e14 [FREE Full text] [doi: 10.1016/j.jpainsymman.2020.03.019] [Medline: 32240756]
3. Patel SY, Mehrotra A, Huskamp HA, Uscher-Pines L, Ganguli I, Barnett ML. Variation in telemedicine use and outpatient care during the COVID-19 pandemic in the United States. Health Aff (Millwood) 2021 Feb;40(2):349-358 [FREE Full text] [doi: 10.1377/hlthaff.2020.01786] [Medline: 33523745]
4. Chang JE, Lai AY, Gupta A, Nguyen AM, Berry CA, Shelley DR. Rapid Transition to Telehealth and the Digital Divide: Implications for Primary Care Access and Equity in a Post-COVID Era. Milbank Q 2021 Jun;99(2):340-368 [FREE Full text] [doi: 10.1111/1468-0009.12509] [Medline: 34075622]
5. Pierce RP, Stevermer JJ. Disparities in use of telehealth at the onset of the COVID-19 public health emergency. J Telemed Telecare 2020 Oct 21:1357633X20963893 [FREE Full text] [doi: 10.1177/1357633X20963893] [Medline: 33081595]
6. Whaley CM, Pera MF, Cantor J, Chang J, Velasco J, Hagg HK, et al. Changes in health services use among commercially insured US populations during the COVID-19 pandemic. JAMA Netw Open 2020 Nov 02;3(11):e2024984 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.24984] [Medline: 33151319]
7. Zhang K, Liu W, Locatis C, Ackerman M. Mobile videoconferencing apps for telemedicine. Telemed J E Health 2016 Jan;22(1):56-62 [FREE Full text] [doi: 10.1089/tmj.2015.0027] [Medline: 26204322]
8. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/bf03195514] [Medline: 14587545]
9. Macefield R. How to specify the participant group size for usability studies: a practitioner's guide. Journal of Usability Studies 2009;5(1):34-45 [FREE Full text]
10. Champagne K, Boot WR. Exploring predictors of mobile device proficiency among older adults. 2017 Presented at: International Conference on Human-Computer Interaction; 2017; Vancouver, Canada p. 162-171. [doi: 10.1007/978-3-319-58077-7_13]
11. Williams S, Orsega-Smith E, Ruggiero L. Exploring technology perceptions and intentions to use in older adults. Innovation in Aging 2020;4(Suppl 1):194. [doi: 10.1093/geroni/igaa057.627]
12. Lyles CR, Wachter RM, Sarkar U. Focusing on digital health equity. JAMA 2021 Nov 09;326(18):1795-1796. [doi: 10.1001/jama.2021.18459] [Medline: 34677577]

Edited by A Mavragani; submitted 17.11.21; peer-reviewed by N Fareed, K Freeman-Costin; comments to author 31.01.22; revised version received 16.03.22; accepted 03.06.22; published 27.07.22.

Please cite as:
Ufholz K, Sheon A, Bhargava D, Rao G
Telemedicine Preparedness Among Older Adults With Chronic Illness: Survey of Primary Care Patients
JMIR Form Res 2022;6(7):e35028
URL: https://formative.jmir.org/2022/7/e35028
doi:10.2196/35028
PMID:35896013

Protocol

Study Protocol and Preliminary Results of the Impact of Occupational Health Workers' Activities on Their Health: Nationwide Prospective Internet-Based Survey

Kazunori Ikegami[1], MD, PhD; Yasuro Yoshimoto[1], MD; Hiroka Baba[1], MD; Shingo Sekoguchi[1], MD; Hajime Ando[1], MD, MOH, PhD; Akira Ogami[1], MD, PhD

Department of Work Systems and Health, Institute of Industrial Ecological Sciences, University of Occupational and Environmental Health, Fukuoka, Japan

Corresponding Author:
Kazunori Ikegami, MD, PhD
Department of Work Systems and Health
Institute of Industrial Ecological Sciences
University of Occupational and Environmental Health
1-1, Iseigaoka, Yahata-nishi-ku
Kitakyushu
Fukuoka, 807-8555
Japan
Phone: 81 93 603 1611
Fax: 81 93 601 2667
Email: kikegami@med.uoeh-u.ac.jp

Abstract

Background: Owing to the impact of the COVID-19 pandemic, work environments and systems, as well as occupational health measures or activities that fall within our research field, are constantly changing. It is necessary to assess the impact of these changes on the physical and mental health of workers.

Objective: To assess how occupational health measures affect the health of workers, we conducted a baseline, longitudinal internet-based survey among Japanese workers in October 2021 and additionally scheduled 2 follow-up surveys for 2022 and 2023. We describe the details of the protocol of the work systems and health internet research (WSHIR) study, provide an overview of the results of the baseline survey, and discuss the study procedures and data used in the study.

Methods: This prospective cohort study was conducted online among internet monitors. The baseline survey was conducted from October 1 to 7, 2021. This study targeted those who were working and between the ages of 20 and 69 years. A total of 5111 respondents who passed the screening survey and proceeded to the main survey were enrolled according to collection units organized by sex and age. For the screening and main surveys, the questionnaire consisted of 9 and 33 items with 9 and 55 questions, respectively. Consistency and completeness checks were performed after the questionnaires were submitted. We compared basic characteristics, such as sex, age group, educational background, and marital status, among all participants, including those who withdrew from the analysis.

Results: Of the 5111 initial survey respondents, 571 (11.2%) were considered fraudulent. The data of the remaining 4540 (88.8%) participants (2273, 50.1%, males; 2267, 49.9%, females) included in the analysis were well balanced across participant sex and age groups according to the sampling plan because there was no significant difference by sex and age group using the chi-square test for checking the distribution bias of the participants (P=.84). Compared to female participants, male participants tended to be more likely to be managers and supervisors (323, 14.2%, males; 86, 3.8%, females), to work in a secondary industry (742, 32.6%, males; 357, 15.7%, females), and to have an annual income of ≥5 million yen (976, 42.9%, males; 429, 18.9%, females). For the evaluation of a psychological indicator, Kessler 6 (K6) score, by sex and age group, the characteristics of the score distribution of the included participants were similar to those reported in previous studies.

Conclusions: This study presents a protocol and overview of the results of an internet-based occupational health survey of workers. Using the results of this survey, we hope to evaluate the changes in occupational health activities and their impact on workers' health while controlling for the COVID-19 pandemic.

(JMIR Form Res 2022;6(7):e35290) doi:10.2196/35290

KEYWORDS

Japan; occupational health; worker; internet surveys; questionnaires; cohort study; COVID-19; mental health; online health

Introduction

The global outbreak of COVID-19 in 2020 had a profound impact on the economy, daily and working life, and medical practice in Japan [1-3]. The Japanese government repeatedly announced a state of emergency, asking the public to exercise voluntary restraint, such as refraining from going out and traveling to distant places, curtailing corporate business activities, and refraining from dinners and other socializing opportunities [4]. In the occupational field, several COVID-19 infection control guidelines were developed by various industries and organizations [5]. The introduction of telecommuting [6] and the implementation of COVID-19 vaccination or antigen testing in workplaces have also been recommended. These changes brought about by the COVID-19 pandemic have resulted in dramatic changes in the work environment, work systems, and occupational health activities [7,8].

In Japan, the occupational health system is defined using occupational safety and health laws, and occupational health services are implemented in many companies [9,10]. Companies with 50 or more employees are required to appoint at least 1 industrial physician and 1 health manager [9]. Occupational physicians have been among the leaders in promoting infection and prevention measures in workplaces during the COVID-19 pandemic [5]. Occupational physicians have often played a central role in workplace COVID-19 vaccination programs as well as in awareness-raising activities for countermeasures against the COVID-19 pandemic in occupational fields. They provide health support for employees affected by COVID-19 and health management for all employees in view of the ever-changing landscape and the impact of the COVID-19 pandemic. We believe that the COVID-19 pandemic brought a more vivid focus on occupational physicians by the public and that it served as the turning point for the promotion of occupational health activities.

In Japan, since the lifting of the government's emergency restrictions at the end of October 1, 2021, until now (end of June 2022), no restrictions have been in place, and more than 60% of the population has completed the third vaccination against COVID-19. It is possible that COVID-19 will once again become prevalent in Japan and that countermeasures will have to be taken on a case-by-case basis. However, it is unlikely that the situation will change as dramatically as it did between early 2020 and September 2021. Looking ahead to the post–COVID-19 pandemic era, it is important to monitor future occupational health activities and assess how they will affect the work environment and systems or the health of workers.

We consider it necessary to focus on future challenges and issues regarding occupational health fields by looking ahead at the post–COVID-19 pandemic period. Such future challenges and issues include cooperation between health management and practice; workers' health management, including annual health checkups, countermeasures against communicable diseases in the workplace, and fitness to work; and the actual status of occupational health services and occupational physician activities, in addition to longitudinally assessing how these affect the health of workers. Therefore, we conducted a longitudinal study, that is, a work systems and health internet research (WSHIR) study, among workers from October 2021. In addition, we scheduled 2 follow-up surveys for 2022 and 2023.s

The aim of this paper is to present details of the WSHIR study protocol. Moreover, it provides an overview of the results of the baseline survey and includes a discussion of the study procedures and improvements to the quality of the data used in the study. We plan to use the data from this study to inform various research themes focused on occupational health issues, such as the impact of occupational health services and activities on workers, changes, and new challenges for occupational health in the workplace in the post–COVID-19 state.

Methods

Study Design

This survey was a prospective cohort study conducted online among internet monitors registered with Cross Marketing Inc. (Tokyo, Japan), which is a Japanese internet research contractor with 4.7 million registered monitors. We sent participation information to the registered monitors, so this was not an open survey.

The baseline survey was conducted from October 1 to 7, 2021. Two follow-up surveys are scheduled for 2022 and 2023. The study targeted those who were working and between the ages of 20 and 69 years in the baseline survey.

A document describing the time required to complete the survey, storage location, period of the survey data, the investigator, and the purpose of the study are available on the website of the Department of Work System and Health, Institute of Industrial Ecological Sciences, University of Occupational and Environmental Health, Japan (in Japanese).

Ethical Considerations

All participants provided informed consent online to participate in this study. The study was approved by the ethics committee of the University of Occupational and Environmental Health, Japan (reference no. R3-037).

Sample Size

The statistical method was not determined, because this study was mainly exploratory. However, the sample size was calculated by adapting the following conditions, which are most likely to be assumed: in sectional or cohort studies, 2-sided significance level $(1 - \alpha)=95$; power $(1 - \beta$, percentage chance of detecting$)=80$; ratio of sample size, unexposed/exposed=1; percentage of unexposed with the outcome=5; odds ratio=1.5;

risk/prevalence ratio=1.5; and risk/prevalence difference=2.3. Under these conditions, the required total sample size was calculated to be 3380 [11]. Hence, we set the sample size to 5000 to account for data exclusion.

Sampling Plan

Only people registered with Cross Marketing Inc. (Tokyo, Japan) could complete the survey. First, the questionnaire for the screening survey confirmed informed consent to participate in the study; the respondents were regular workers, we excluded temporary and part-time workers, and the age was ≥20 years. Only respondents who met these conditions proceeded to the main survey, which consisted of 10 collection units organized by both sex and age group, with 500 respondents per collection unit, for a total sample size of 5000. Third, each collection unit was designed to be closed once it reached 520 respondents; thereafter, respondents could not proceed from the screening stage to the main survey stage.

Personal information was not collected from a series of surveys. The respondents were contracted according to the privacy policy of Cross Marketing, Inc.

Recruitment Process for Participants

As instructed, Cross Marketing Inc. sent emails that included a link to the website, along with an introduction to the survey and an entry to the questionnaire page. There is an automatic method for capturing the responses on a website. Completion of this survey was voluntary, and as an incentive to complete the survey, the respondents earned points that could be exchanged for various products.

We counted the number of participants in this survey by counting the monitor IDs assigned to the respondents when they accessed the survey system. For the first survey, participation invitations were sent via email to approximately 59,000 monitors randomly selected by Cross Marketing Inc. from among more than 5 million registered monitors. This survey was started on October 1, 2021, and all sample collections were completed on October 5, 2021.

Measurements

For the screening and main surveys, the questionnaire consisted of 9 and 33 items with 9 and 55 questions, respectively. The questionnaire consisted of 33 pages, with 1 item per page. The survey consisted of the following 3 major categories: basic and socioeconomic characteristics and health status, a psychological questionnaire that was already validated, and questions pertaining to occupational medicine and health.

Questions relating to basic and socioeconomic characteristics included sex, age, marital status, income, educational background, area of residence, and work-related factors, such as occupation, number of employees at workplaces (branches, factories, and sales offices), type of industry where the participants worked, and average working hours. In addition, single items regarding the participants' health status included present medical history, presence of current physical and psychological problems and their causes, and the number of days of sick leave absence.

The psychological questionnaires included the Brief Job Stress Questionnaire [12], the Japanese version of the 3-item Utrecht Work Engagement Scale [13,14], and Quantity and Quality [15] as an evaluation index of presenteeism. Other psychological questionnaires were the Patient Health Questionnaire-2 [16] and the Kessler 6 (K6) score [17,18]. The reliability and validity of these questionnaires on psychological scales have been confirmed in previous studies. The Japanese versions of these psychological scales were used without modification.

For questions related to occupational medicine and health, we surveyed the perceived workplace health support [19], the actual state of health management for workers (eg, health checkups, countermeasures for return to work), fitness to work, countermeasures against communicable diseases in the workplace, provision of occupational health services, health consultation services, and management of workers' health information. Each question consisted of 1 or 2 original items and was not evaluated by calculating the scale scores.

All the aforementioned questionnaires were created or selected by 3 experts certified by senior occupational health physicians, who were certified by the Japan Society for Occupational Health, the body that discusses current issues regarding occupational health in Japan. After generating the questionnaires, we requested 3 other occupational physicians to respond and review the drafts.

Additionally, we verified that these questions could be answered without any problems on a web-based system before conducting the main survey. We also checked for inappropriate expressions, ease of answering, typographical errors, and other issues that were used to revise the questionnaires.

Completeness Check

Consistency and completeness checks were performed after the questionnaires were submitted. We detected fraudulent respondents based on the 3 types of algorithms designed in this survey:

- Respondents who failed to correctly answer 2 basic knowledge questions unrelated to the survey. Specifically, one was to correctly select odd numbers among five 2-digit numbers, and the other was to correctly select multiples of 3 out of five 2-digit numbers.
- Respondents who provided contradictory answers to the 3 predetermined questions (the 3 questions were likely to be contradictory if the respondents did not answer them carefully). Specifically, we designed a 2-option (yes/no) question on whether the respondent had undergone a health checkup within the past year, followed by a 6-option question on how to obtain, store, and use the results of the health checkup. In 1 instance, the respondent answered, "I had not had a health checkup within a year," but the respondent also answered, "I kept the results of the health checkup within a year."
- Respondents whose response time was <3 minutes. To exclude questionnaires submitted too soon, we excluded those with a response time of less than 3 minutes. We set these cut-off points based on our actual response time to the questionnaires and the fact that Fujino et al [20] set the

cut-off points at 6 minutes for the questionnaire consisting of 55 items and 160 questions in their study.

Respondents who met the exclusion criteria based on these consistency or completeness checks were considered withdrawn; otherwise, they were considered enrolled.

Statistical Analysis

We compared basic characteristics, such as sex, age group, educational background, and marital status, between enrolled and withdrawn participants in the analysis using the chi-square test. In addition, we used the chi-square test to compare the characteristics between respondents who provided contradictory answers to the 3 questions and those who did not and between those whose response time was <3 minutes and those whose response time was ≥3 minutes. The comparisons between those who answered the 2 basic knowledge questions correctly and incorrectly were not analyzed because of incorrect answers.

We analyzed the educational background, marital status, occupation, industrial classification, number of employees in the business unit where the participants worked, number of employees in the company where the participants worked, and annual income (yen) by sex or age group (20-29, 30-39, 40-49, 50-59, and 60-69 years) using the chi-square test. We analyzed the K6 score by sex using the Mann-Whitney *U* test or by age group using the Kruskal-Wallis test.

Results

Participant Details

As shown in Figure 1, the survey invitation was sent to approximately 59,000 registrants, and 7300 (12.4%) responded to the screening survey. A total of 2189 (30%) respondents were excluded from the screening survey stage, and 5111 (70%) completed the main survey (completion rate=ratio of users who completed the survey to users who agreed to participate). The distribution of the enrolled participants by sex and age group (ie, by 10 collection units) is shown in Table 1. The number of enrolled participants by sex and age group was evaluated using the chi-square test to check for distribution bias; there was no significant difference (*P*=.99). Of the 5111 respondents, 571 (11.2%) withdrew. Of those with duplication, 434 (8.5%) provided contradictory answers to the 3 questions, 161 (3.2%) had a response time of less than 3 minutes, and none had incorrect answers to the 2 basic knowledge questions. The final number of participants enrolled in the analysis was 4540 (Figure 1). The number of enrolled participants by sex and age group was evaluated using the chi-square test to check for distribution bias; there were no significant differences (*P*=.84).

According to residence (prefecture), the 4540 participants were distributed across all 47 prefectures in Japan. The highest and lowest proportion of respondents per 100,000 population was 5.4 (Tokyo) and 1.5 (Miyazaki Prefecture), with a 47-prefecture median (quartile) of 3.0 (2.5-3.6). Among 9 (19.1%) of the 47 prefectures, with a population of more than 5 million per prefecture, 7 (77.8%) were among the top 10 with the highest proportion of respondents per 100,000 population.

We compared the basic characteristics of the enrolled and withdrawn participants (Table 2). There were no significant differences in sex, educational background, or marital status among groups. However, a higher proportion of younger participants withdrew.

We also compared the basic characteristics of the respondents who provided contradictory answers to the 3 questions (Table 2). The number of respondents who provided contradictory answers was 434 (8.5%) of 5111. They tended to be younger and less educated, and 24 (5.5%) had a response time of <3 minutes; in addition, they were significantly more in number (n=434, 8.5%) than those who provided no contradictory responses (n=137, 2.9%).

Next, we compared the basic characteristics of respondents whose response times were ≥3 and <3 minutes (Table 3). The number of respondents with a response time of <3 minutes was 161 (3.2%). The median, 25th, and 75th percentiles were 6 minutes 49 seconds, 4 minutes 56 seconds , and 9 minutes 50 seconds, respectively. Those with a response time of <3 minutes tended to be younger, better educated, and unmarried. Of those with a response time of <3 minutes, those with contradictory answers were significantly more in number (n=24, 14.9%) than those with a response time of ≥3 minutes (n=410, 8.3%).

We compared the basic and work-related characteristics of the enrolled participants between the sexes and among the 5 age groups (Tables 4 and 5). Male participants tended to be more likely to be married and have a university or graduate school degree than female participants. In terms of work-related characteristics, male participants tended to be more likely to be managers and supervisors, work in a secondary industry or in large-size workplaces or enterprises, and have an annual income of ≥5 million yen (US $36,183.51) than the female participants. Female participants were significantly more likely to be in a third industry, to work in small workplaces or enterprises, and to have an annual income of <2.99 million yen (US $20,986.44) than the male participants.

The 20-29-year age group tended to be less likely to be married compared to those in other age groups. Regarding educational background, younger participants tended to be more likely to have a university or graduate school degree, be regular employees, and be employed in large workplaces or enterprises. With the increasing age of the respondents, an increased proportion was found among those who were engaged in workplaces or enterprises (49 employees or fewer), as well as those who were engaged in the secondary industry.

The K6 score was higher in female participants than in male participants; in addition, the K6 score tended to be higher in those aged 20-39 years (Table 6).

Figure 1. Flowchart of participant selection. WSHIR: work systems and health internet research.

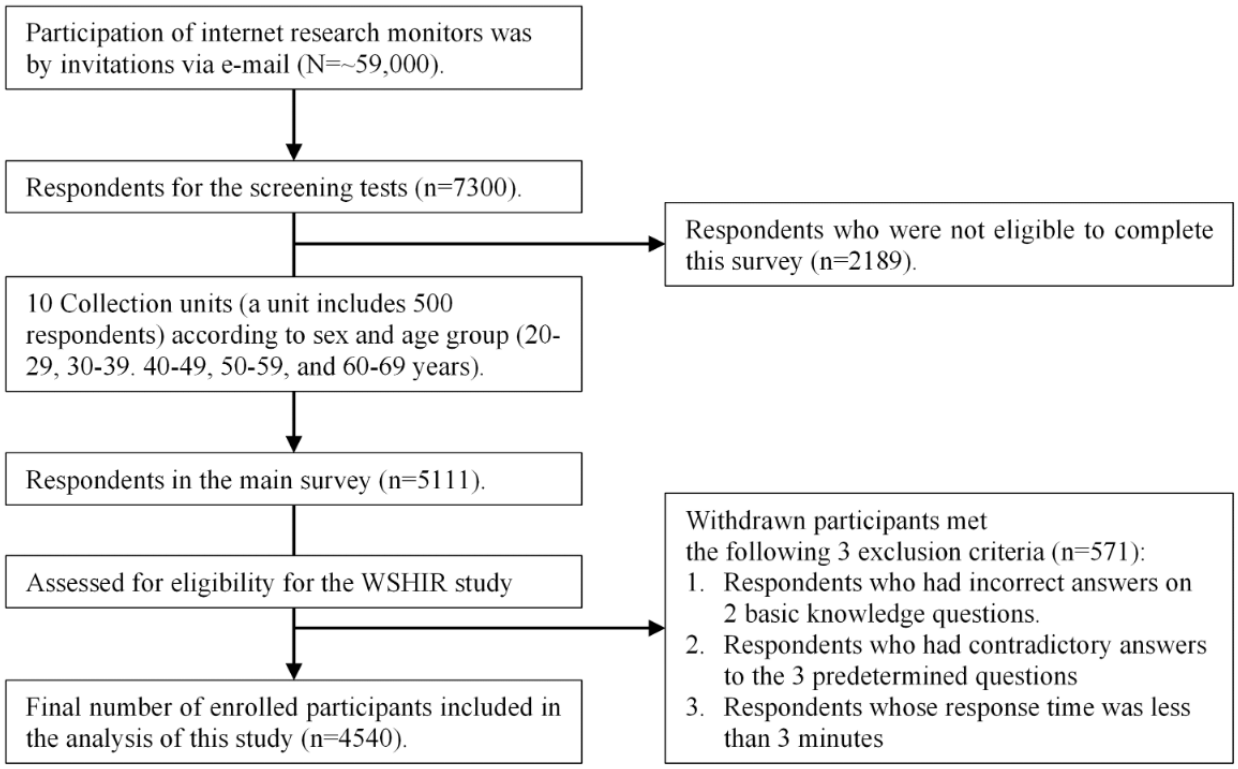


Table 1. Respondents and enrolled participants per collection unit by age group and sex.

Age group (years)	All participants (N=5111), *P*=.99		Enrolled participants (N=4540), *P*=.84	
	Male participants (N=2551), n (%)	Female participants (N=2560), n (%)	Male participants (N=2273), n (%)	Female participants (N=2267), n (%)
20-29	507 (19.9)	513 (20.0)	425 (18.7)	418 (18.4)
30-39	505 (19.8)	509 (19.9)	420 (18.5)	448 (19.8)
40-49	514 (20.1)	500 (19.5)	467 (20.5)	448 (19.8)
50-59	511 (20.0)	518 (20.2)	480 (21.1)	481 (21.2)
60-69	514 (20.1)	520 (20.3)	481 (21.2)	472 (20.8)

Table 2. Comparison of enrolled and withdrawn participants.

Characteristics	Total (N=5111), n (%)	Participants		Contradictory answers[a]	
		Enrolled (N=4540), n (%)	Withdrawn (N=571), n (%)	No (N=4677), n (%)	Yes (N=434), n (%)
Sex; participants *P*=.53, contradictory answers *P*=.33					
Male	2551 (49.9)	2273 (50.1)	278 (48.7)	2344 (50.1)	207 (47.7)
Female	2560 (50.1)	2267 (49.9)	293 (51.3)	2333 (49.9)	227 (52.3)
Age group (years); participants *P*<.001, contradictory answers *P*<.001					
20-29	1020 (20.0)	843 (18.6)	177 (31.0)	898 (19.2)	122 (28.1)
30-39	1014 (19.8)	868 (19.1)	146 (25.6)	914 (19.5)	100 (23.0)
40-49	1014 (19.8)	915 (20.2)	99 (17.3)	941 (20.1)	73 (16.8)
50-59	1029 (20.1)	961 (21.2)	68 (11.9)	967 (20.7)	62 (14.3)
60-69	1034 (20.2)	953 (21.0)	81 (14.2)	957 (20.5)	77 (17.7)
Educational background; participants *P*=.83, contradictory answers *P*=.04					
Junior high school or high school	1104 (21.6)	976 (21.5)	128 (22.4)	995 (21.3)	109 (25.1)
Technical college or junior college	1086 (21.2)	963 (21.2)	123 (21.5)	984 (21.0)	102 (23.5)
University or graduate school	2921 (57.2)	2601 (57.3)	320 (56.0)	2698 (57.7)	223 (51.4)
Marital status; participants *P*=.11, contradictory answers *P*=.62					
Unmarried	2486 (48.6)	2190 (48.2)	296 (51.8)	2270 (48.5)	216 (49.8)
Married	2625 (51.4)	2350 (51.8)	275 (48.2)	2407 (51.5)	218 (50.2)
Contradictory answers[a]					
No	4677 (91.5)	4540 (100.0)	137 (24.0)	N/A[b]	N/A
Yes	434 (8.5)	0	434 (76.0)	N/A	N/A
Response time (minutes; contradictory answers *P*=.003)					
≥3	4950 (96.8)	4540 (100.0)	410 (71.8)	4540 (97.1)	410 (94.5)
<3	161 (3.2)	0	161 (28.2)	137 (2.9)	24 (5.5)

[a]Respondents who provided contradictory answers to the 3 predetermined questions (yes) and those who did not (no).

[b]N/A: not applicable.

Table 3. Basic characteristics of respondents (N=5111) whose response times were ≥3 and <3 minutes.

Characteristics	Total, n (%)	Response time (minutes)	
		≥3 (N=4950), n (%)	<3 (N=161), n (%)
Sex; *P*=.56			
Male	2551 (49.9)	2467 (49.8)	84 (52.2)
Female	2560 (50.1)	2483 (50.2)	77 (47.8)
Age group (years); *P*<.001			
20-29	1020 (20.0)	957 (19.3)	63 (39.1)
30-39	1014 (19.8)	960 (19.4)	54 (33.5)
40-49	1014 (19.8)	983 (19.9)	31 (19.3)
50-59	1029 (20.1)	1021 (20.6)	8 (5.0)
60-69	1034 (20.2)	1029 (20.8)	5 (3.1)
Educational background; *P*=.004			
Junior high school or high school	1104 (21.6)	1082 (21.9)	22 (13.7)
Technical college or junior college	1086 (21.2)	1059 (21.4)	27 (16.8)
University or graduate school	2921 (57.2)	2809 (56.7)	112 (69.6)
Marital status; *P*=.02			
Unmarried	2486 (48.6)	2393 (48.3)	93 (57.8)
Married	2625 (51.4)	2557 (51.7)	68 (42.2)
Contradictory answers[a]; *P*>=.003			
No	4677 (91.5)	4540 (91.7)	137 (85.1)
Yes	434 (8.5)	410 (8.3)	24 (14.9)

[a]Respondents who provided contradictory answers to the 3 predetermined questions (yes) and those who did not (no).

Table 4. Basic characteristics of enrolled participants by sex.

Characteristics	Male participants (N=2273), n (%)	Female participants (N=2267), n (%)
Educational background; *P*<.001		
Junior high school or high school	501 (22.0)	475 (21.0)
Technical college or junior college	276 (12.1)	687 (30.3)
University or graduate school	1496 (65.8)	1105 (48.7)
Marital status; *P*<.001		
Unmarried	877 (38.6)	1313 (57.9)
Married	1396 (61.4)	954 (42.1)
Occupation; *P*<.001		
Regular employees	1257 (55.3)	1354 (59.7)
Managers	323 (14.2)	86 (3.8)
Others	693 (30.5)	827 (36.5)
Industrial classification; *P*<.001		
Primary industry	8 (0.4)	3 (0.1)
Secondary industry	742 (32.6)	357 (15.7)
Third industry	1523 (67.0)	1907 (84.1)
Number of employees of business units where the participants worked; *P*<.001		
1-49	809 (35.6)	1028 (45.3)
50-999	831 (36.6)	726 (32.0)
≥1000	561 (24.7)	375 (16.5)
Unclear	72 (3.2)	138 (6.1)
Number of employees of companies where the participants worked; *P*<.001		
1-49	643 (28.3)	818 (36.1)
50-299	457 (20.1)	445(19.6)
300-999	292 (12.8)	269 (11.9)
1000-9999	466 (20.5)	290 (12.8)
≥10,000	274 (12.1)	191 (8.4)
Unclear	141 (6.2)	254 (11.2)
Annual income (yen); *P*<.001		
<3 million (<US $21,710.11[a])	281 (12.4)	713 (31.5)
3-4.9 million (US $21,710.11-$35,459.84)	735 (32.3)	720 (31.8)
5-9.9 million (US $36,183.51-$71,643.35)	787(34.6)	354 (15.6)
≥10 million (US $72,367.02)	189 (8.3)	75 (3.3)
Unclear	281 (12.4)	405 (17.9)

[a]An exchange rate of 1 Japanese yen=US $0.0072 has been applied.

Table 5. Basic characteristics of enrolled participants by age group.

Characteristics	20-29 years (N=843), n (%)	30-39 years (N=868), n (%)	40-49 years (N=915), n (%)	50-59 years (N=961), n (%)	60-69 years (N=953), n (%)
Educational background; *P*<.001					
Junior high school or high school	145 (17.2)	147 (16.9)	197 (21.5)	281 (29.2)	206(21.6)
Technical college or junior college	144 (17.1)	150 (17.3)	221 (24.2)	251 (26.1)	197 (20.7)
University or graduate school	554 (65.7)	571 (65.8)	497 (54.3)	429 (44.6)	550 (57.7)
Marital status; *P*<.001					
Unmarried	632 (75.0)	436 (50.2)	443 (48.4)	389 (40.5)	290 (30.4)
Married	211 (25.0)	432 (49.8)	472 (51.6)	572 (59.5)	663 (69.6)
Occupation; *P*<.001					
Regular employees	635 (75.3)	621 (71.5)	563 (61.5)	451 (46.9)	341 (35.8)
Managers	9 (1.1)	38 (4.4)	111 (12.1)	159 (16.5)	92 (9.7)
Others	199 (23.6)	209 (24.1)	241 (26.3)	351 (36.5)	520 (54.6)
Industrial classification; *P*=.01[a]					
Primary industry	4 (0.5)	3 (0.3)	2 (0.2)	2 (0.2)	0
Secondary industry	202 (24.0)	224 (25.8)	232 (25.4)	256 (26.6)	185 (19.4)
Third industry	637 (75.6)	641 (73.8)	681 (74.4)	703 (73.2)	768 (80.6)
Number of employees in business units where the participants worked; *P*<.001					
1-49	223 (26.5)	286 (32.9)	331 (36.2)	445 (46.3)	552 (57.9)
50-999	357 (42.3)	347 (40.0)	346 (37.8)	283 (29.4)	224 (23.5)
≥1000	197 (23.4)	186 (21.4)	199 (21.7)	202 (21.0)	152 (15.9)
Unclear	66 (7.8)	49 (5.6)	39 (4.3)	31 (3.2)	25 (2.6)
Number of employees in companies where the participants worked; *P*<.001					
1-49	132 (15.7)	209 (24.1)	268 (29.3)	375 (39.0)	477 (50.1)
50-299	211 (25.0)	199 (22.9)	181 (19.8)	179 (18.6)	132 (13.9)
300-999	135 (16.0)	127 (14.6)	126 (13.8)	91 (9.5)	82 (8.6)
1000-9999	166 (19.7)	136 (15.7)	174 (19.0)	156 (16.2)	124 (13.0)
≥10,000	93 (11.0)	111 (12.8)	86 (9.4)	95 (9.9)	80 (8.4)
Unclear	106 (12.6)	86 (9.9)	80 (8.7)	65 (6.8)	58 (6.1)
Annual income (yen); *P*<.001					
<3 million (<US $21,710.11[b])	219 (26.0)	178 (20.5)	160 (17.5)	205 (21.3)	232 (24.3)
3-4.9 million (US $21,710.11-$35,459.84)	406 (48.2)	312 (35.9)	258 (28.2)	218 (22.7)	261 (27.4)
5-9.9 million (US $36,183.51-$71,643.35)	97 (11.5)	234 (27.0)	310 (33.9)	279 (29.0)	221 (23.2)
≥10 million (US $72,367.02)	11 (1.3)	36 (4.1)	54 (5.9)	82 (8.5)	81 (8.5)
Unclear	110 (13.0)	108 (12.4)	133 (14.5)	177 (18.4)	158 (16.6)

[a]In 33% of the cells, the expected frequencies are <5; therefore, this *P* value is not accurate.

[b]An exchange rate of 1 Japanese yen=US $0.0072 has been applied.

Table 6. K6[a] scores of enrolled participants by sex and age group (P<.001).

Characteristics	K6 score	Participants, n (%)
Sex		
Male (N=2273)	0-7	2 (0.1)
Female (N=2267)	0-7	3 (0.1)
Age group (years)		
20-29 (N=843)	0-8	8 (0.9)
30-39 (N=868)	0-9	9 (1.0)
40-49 (N=915)	0-7	7 (0.7)
50-59 (N=961)	0-6	6 (0.6)
60-69 (N=953)	0-4	4 (0.4)

[a]K6: Kessler 6.

Discussion

Principal Findings

In October 2021, after the fifth wave of the COVID-19 pandemic had subsided in Japan, we conducted an internet-based occupational health survey among workers. Internet surveys have become more common in recent years in the fields of public health and epidemiology, health care services, and even medicine because of the potential to collect relatively large amounts of data in a short period [21]. Compared to conventional population and workplace surveys, internet surveys have the advantages of making it easier to achieve the target sample size, incorporating many conditions, surveying in a short period, and making it easier to conduct surveys targeting workers. We believe that our data obtained using the WSHIR study will be valuable for future research on the working conditions and health status of workers post–COVID-19.

One of the problems with internet surveys is that respondents may provide fraudulent answers [22,23]. Many private internet survey companies set up incentives, such as points that can be exchanged for products, to increase the number of registrants and encourage them to completely respond to various surveys. There is a possibility that some respondents may answer the questionnaires inappropriately without understanding the aims of the survey or the questionnaire just to obtain these incentives. Therefore, internet surveys must be designed to detect fraudulent respondents.

In this study, we used 3 algorithms to detect fraudulent respondents. The first was the setting of 2 basic knowledge questions that were not related to the main survey. However, all the respondents answered these questions correctly. It is possible that in many surveys, questions have already been prepared to detect fraudulent respondents or that many respondents are aware that the questions are designed to detect fraudulent practices; thus, respondents decide to respond correctly. However, we speculate that ensuring that the respondents are aware that the questionnaire contains algorithms to detect fraudulent respondents may have a deterrent effect on fraudulent responses.

Second, the way the 3 predetermined questions were set could lead to contradictory answers if they were not answered carefully. This algorithm is complicated because the respondent must be consistent among the 3 predetermined questions. In fact, 434 of 5111 respondents provided contradictory answers, and all were treated as fraudulent respondents.

Finally, the cut-off points were set to exclude premature response times to the questionnaire, which were recorded using a questionnaire system by Cross Marketing Inc. In this survey, the median, 25th, 75th, and 5th percentiles were 6 minutes 49 seconds, 4 minutes 56 seconds, 9 minutes 50 seconds, and 3 minutes 16 seconds, respectively. Therefore, we believe it was reasonable to exclude 161 respondents whose answers were within 3 minutes. Regarding the 2 conditions for detecting fraudulent respondents, we found that those who met one of the conditions were significantly more likely to meet the other condition.

This study found that respondents with extremely short response times tend to provide contradictory responses. In addition, this tendency was observed in younger participants. We speculate that 1 of the reasons for this is the possibility that there are a certain number of internet monitors who are only interested in obtaining incentives. When conducting a questionnaire survey with slightly more difficult content via the internet, as in this study regarding occupational health fields, the researchers propose that a procedure is needed to validate the data set provided by internet research contractors.

In this survey, the respondents' residences were not added to the collection unit. However, we could enroll participants from all 47 prefectures in Japan, although the proportion of respondents tended to be higher in metropolitan areas. In a conventional occupational health survey, where the researcher directly asks for research cooperation from companies, it is almost impossible to adjust for this kind of regional distribution of participants, and we believe that the data are well balanced.

By examining the relationships among the variables, we confirmed the authenticity of the data. For example, compared to female participants, male participants had a higher proportion of managers and those engaged in the secondary industry as well as a higher annual income. In addition, regarding

psychological indicators, it has been reported that the K6 scores of women are higher than those of men [24,25], and a similar trend was observed in this study. Regarding age groups, the younger age group was more likely to be unmarried, and the proportion of managers was higher in the 40-49- and 50-59-year age groups and lower in the 60-69-year age group. In Japan, most companies have a mandatory retirement age of 60 years, and those aged ≥60 years are often rehired without job positions, on contract, or in fixed-term employment. As for psychological indicators, previous studies have reported that K6 scores tend to be higher in younger persons, and similar results were observed in this study [26].

Selection bias is unavoidable in internet surveys. For example, those who use the internet and are willing to answer the questionnaire will inevitably be selected. Respondents were simply those who were registered with an internet research company and did not represent a particular population. *Measurement error* refers to errors caused by differences in the approach used in responding, even by the same respondent, depending on whether the answers are recorded using other methods (eg, by telephone or interview) or are self-recorded (eg, by the respondent self-completing individually) and whether the answers are shown once the questionnaire paper has been filled out or shown on an internet screen [22,27,28]. To improve the validity of this study, we adhered to the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) statement [21]. In addition, it is important to clarify the characteristics of the survey population by comparing various factors with those in previous studies. This study focused on occupational health, and we confirmed work-related factors, such as occupation, industry classification, size of the workplaces or enterprises where the participants worked, annual income, working hours, occupational health system, and activity status in as much detail as possible. We also examined several health-related factors of workers and psychosocial indicators related to work that have been used in many previous studies, which can be compared to the results of this study.

This study is intended to be conducted over a 3-year period, starting in October 2021, when several people in Japan have been vaccinated against COVID-19, with the number of infected people remaining low even after the government lifted the state of emergency. Of course, there is a possibility of a recurrent epidemic in the future. However, the government and the Japan Federation of Economic Organizations (Nippon Keidanren) are making efforts to resume or strengthen economic activities, since the situation may be approaching a possible control of the COVID-19 pandemic [21,22]. Once controlled, this study could provide an overview of changes in occupational health in Japan and how COVID-19 has affected workers.

Limitations

We mentioned some research limitations before, but there are several other limitations to this study. First, the sampling plan was not specifically designed to consider the unit of workplace characteristics, such as workplace size, location, and type of industry. Therefore, it should be noted that the results of this study generalize to the whole working population in Japan. Second, some of the participants belonged to the same workplace. According to the Statistics Bureau of the Ministry of Internal Affairs and Communications, the number of workplaces in Japan is approximately 578,000. The likelihood that all participants had different workplaces was approximately 16.8% by a simple calculation. Because of the possibility that participants may belong to the same workplace, we need to be careful when analyzing and evaluating the study data. However, as we plan to continue this study in the future, we believe that the quality of the data can be improved by obtaining data such as the zip codes of the workplaces in a follow-up survey. Third, all data in this survey were based on parameters self-reported by individual workers. Respondents of this study might be unaware of or might not correctly understand the health-related measures implemented in their workplaces, depending on their position or status. In addition, this study did not use objective indicators of mental or physical health. Such research should utilize many indicators that have been commonly used in previous studies and should be examined with reference to previous studies.

Conclusion

We commenced an internet-based occupational health survey focusing on occupational health activities and workers' health in October 2021 when approximately 80% of the population aged ≥12 years had been vaccinated and the fifth wave of the COVID-19 pandemic was under control in Japan. This paper presents the protocol of this study and provides an overview of the data from the baseline survey, the study procedures, and the quality of the data in this survey. Using the data of this survey, we aimed to evaluate the changes in occupational health activities and their impact on workers' health after the COVID-19 pandemic was controlled. We plan to analyze the data from multiple perspectives and present new findings regarding occupational health fields sequentially.

Acknowledgments

We thank Cross Marketing Inc. (Tokyo, Japan) for conducting this internet survey and Editage for English language editing.

Conflicts of Interest

None declared.

References

1. Jimi H, Hashimoto G. Challenges of COVID-19 outbreak on the cruise ship Diamond Princess docked at Yokohama, Japan: a real-world story. Glob Health Med 2020 Apr 30;2(2):63-65 [FREE Full text] [doi: 10.35772/ghm.2020.01038] [Medline: 33330779]
2. Sayeed UB, Hossain A. How Japan managed to curb the pandemic early on: lessons learned from the first eight months of COVID-19. J Glob Health 2020 Dec;10(2):020390 [FREE Full text] [doi: 10.7189/jogh.10.020390] [Medline: 33282218]
3. World Health Organization. Coronavirus Disease (COVID-19) Situation Reports. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports [accessed 2021-10-20]
4. Ministry of Health Labour and Welfare, Japan. Checklist for the Prevention of COVID-19 Spreading at Workplaces. URL: https://www.mhlw.go.jp/content/11303000/000616869.pdf [accessed 2022-06-02]
5. Suzuki H, Miyamoto T, Hamada A, Nakano A, Okoshi H, Yamasawa F, Members of Occupational Health Committee on JSTH (Japanese Society of TravelHealth). A guide for businesses and employers responding to novel coronavirus disease (COVID-19): 4th edition. J Occup Health 2021 Apr 05;63(1):e12225 [FREE Full text] [doi: 10.1002/1348-9585.12225] [Medline: 34713533]
6. Nagata T, Nagata M, Ikegami K, Hino A, Tateishi S, Tsuji M, CORoNaWork project. Intensity of home-based telework and work engagement during the COVID-19 pandemic. J Occup Environ Med 2021 Nov 01;63(11):907-912 [FREE Full text] [doi: 10.1097/JOM.0000000000002299] [Medline: 34334780]
7. Kikuchi S, Kitao S, Mikoshiba M. Who suffers from the COVID-19 shocks? Labor market heterogeneity and welfare consequences in Japan. J Jpn Int Econ 2021 Mar;59:101117 [FREE Full text] [doi: 10.1016/j.jjie.2020.101117]
8. Sinclair RR, Allen T, Barber L, Bergman M, Britt T, Butler A, et al. Occupational health science in the time of COVID-19: now more than ever. Occup Health Sci 2020;4(1-2):1-22 [FREE Full text] [doi: 10.1007/s41542-020-00064-3] [Medline: 32838031]
9. e-Gov. Industrial Safety and Health Law. URL: https://elaws.e-gov.go.jp/document?lawid=347AC0000000057 [accessed 2022-01-27]
10. Jain A, Hassard J, Leka S, Di Tecco C, Iavicoli S. The role of occupational health services in psychosocial risk management and the promotion of mental health and well-being at work. Int J Environ Res Public Health 2021 Mar 31;18(7):3632 [FREE Full text] [doi: 10.3390/ijerph18073632] [Medline: 33807352]
11. Sullivan KM, Dean A, Soe MM. OpenEpi: a web-based epidemiologic and statistical calculator for public health. Public Health Rep 2009;124(3):471-474 [FREE Full text] [doi: 10.1177/003335490912400320] [Medline: 19445426]
12. Inoue A, Kawakami N, Shimomitsu T, Tsutsumi A, Haratani T, Yoshikawa T, et al. Development of a short version of the new brief job stress questionnaire. Ind Health 2014;52(6):535-540 [FREE Full text] [doi: 10.2486/indhealth.2014-0114] [Medline: 24975108]
13. Shimazu A, Schaufeli W, Kosugi S, Suzuki A, Nashiwa H, Kato A, et al. Work engagement in Japan: validation of the Japanese version of the Utrecht Work Engagement Scale. Appl Psychol 2008 Jul;57(3):510-523. [doi: 10.1111/j.1464-0597.2008.00333.x]
14. Schaufeli WB, Shimazu A, Hakanen J, Salanova M, De Witte H. An ultra-short measure for work engagement. Eur J Psychol Assess 2019 Jul;35(4):577-591. [doi: 10.1027/1015-5759/a000430]
15. Brouwer WB, Koopmanschap MA, Rutten FF. Productivity losses without absence: measurement validation and empirical evidence. Health Policy 1999 Jul;48(1):13-27. [doi: 10.1016/s0168-8510(99)00028-7] [Medline: 10539583]
16. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
17. Furukawa TA, Kawakami N, Saitoh M, Ono Y, Nakane Y, Nakamura Y, et al. The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. Int J Methods Psychiatr Res 2008;17(3):152-158 [FREE Full text] [doi: 10.1002/mpr.257] [Medline: 18763695]
18. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SLT, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002 Aug;32(6):959-976. [doi: 10.1017/s0033291702006074] [Medline: 12214795]
19. Chen L, Hannon PA, Laing SS, Kohn MJ, Clark K, Pritchard S, et al. Perceived workplace health support is associated with employee productivity. Am J Health Promot 2015;29(3):139-146. [doi: 10.4278/ajhp.131216-QUAN-645] [Medline: 25559250]
20. Fujino Y, Ishimaru T, Eguchi H, Tsuji M, Tateishi S, Ogami A, et al. Protocol for a nationwide internet-based health survey of workers during the COVID-19 pandemic in 2020. J UOEH 2021;43(2):217-225 [FREE Full text] [doi: 10.7888/juoeh.43.217] [Medline: 34092766]
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. Greenacre ZA. The importance of selection bias in internet surveys. Open J Stats 2016;06(03):397-404. [doi: 10.4236/ojs.2016.63035]
23. Liu M, Wronski L. Trap questions in online surveys: results from three web survey experiments. Int J Market Res 2018 Feb 05;60(1):32-49. [doi: 10.1177/1470785317744856]

24. Prochaska JJ, Sung H, Max W, Shi Y, Ong M. Validity study of the K6 scale as a measure of moderate mental distress based on mental health treatment need and utilization. Int J Methods Psychiatr Res 2012 Jun 20;21(2):88-97 [FREE Full text] [doi: 10.1002/mpr.1349] [Medline: 22351472]
25. Nagasu M, Muto K, Yamamoto I. Impacts of anxiety and socioeconomic factors on mental health in the early phases of the COVID-19 pandemic in the general population in Japan: a web-based survey. PLoS One 2021;16(3):e0247705 [FREE Full text] [doi: 10.1371/journal.pone.0247705] [Medline: 33730044]
26. Sunderland M, Hobbs MJ, Anderson TM, Andrews G. Psychological distress across the lifespan: examining age-related item bias in the Kessler 6 Psychological Distress Scale. Int Psychogeriatr 2011 Sep 21;24(2):231-242. [doi: 10.1017/s1041610211001852]
27. Bethlehem J. Selection bias in web surveys. Int Stats Rev 2010;78(2):161-188. [doi: 10.1111/j.1751-5823.2010.00112.x] [Medline: 32870582]
28. Cantuaria ML, Blanes-Vidal V. Self-reported data in environmental health studies: mail vs. web-based surveys. BMC Med Res Methodol 2019 Dec 12;19(1):238 [FREE Full text] [doi: 10.1186/s12874-019-0882-x] [Medline: 31830906]

Abbreviations

K6: Kessler 6
WSHIR: work systems and health internet research

Edited by A Mavragani; submitted 29.11.21; peer-reviewed by G Liu, H Ayatollahi, S Kanamori; comments to author 20.06.22; revised version received 05.07.22; accepted 17.07.22; published 28.07.22.

Please cite as:
Ikegami K, Yoshimoto Y, Baba H, Sekoguchi S, Ando H, Ogami A
Study Protocol and Preliminary Results of the Impact of Occupational Health Workers' Activities on Their Health: Nationwide Prospective Internet-Based Survey
JMIR Form Res 2022;6(7):e35290
URL: https://formative.jmir.org/2022/7/e35290
doi:10.2196/35290
PMID:35900807

Original Paper

Predicting Participant Engagement in a Social Media–Delivered Lifestyle Intervention Using Microlevel Conversational Data: Secondary Analysis of Data From a Pilot Randomized Controlled Trial

Ran Xu[1], PhD; Joseph Divito[1], BS; Richard Bannor[1], MPH; Matthew Schroeder[2], MS; Sherry Pagoto[1], PhD

[1]Department of Allied Health Sciences, Institute for Collaboration in Health, Interventions, and Policy, University of Connecticut, Storrs, CT, United States

[2]Center for Aging Research, Indiana University, Indianapolis, IN, United States

Corresponding Author:
Ran Xu, PhD
Department of Allied Health Sciences
Institute for Collaboration in Health, Interventions, and Policy
University of Connecticut
Koons Hall, Room 326
Storrs, CT, 06269
United States
Phone: 1 860 486 2945
Email: ran.2.xu@uconn.edu

Abstract

Background: Social media–delivered lifestyle interventions have shown promising outcomes, often generating modest but significant weight loss. Participant engagement appears to be an important predictor of weight loss outcomes; however, engagement generally declines over time and is highly variable both within and across studies. Research on factors that influence participant engagement remains scant in the context of social media–delivered lifestyle interventions.

Objective: This study aimed to identify predictors of participant engagement from the content generated during a social media–delivered lifestyle intervention, including characteristics of the posts, the conversation that followed the post, and participants' previous engagement patterns.

Methods: We performed secondary analyses using data from a pilot randomized trial that delivered 2 lifestyle interventions via Facebook. We analyzed 80 participants' engagement data over a 16-week intervention period and linked them to predictors, including characteristics of the posts, conversations that followed the post, and participants' previous engagement, using a mixed-effects model. We also performed machine learning–based classification to confirm the importance of the significant predictors previously identified and explore how well these measures can predict whether participants will engage with a specific post.

Results: The probability of participants' engagement with each post decreased by 0.28% each week (P<.001; 95% CI 0.16%-0.4%). The probability of participants engaging with posts generated by interventionists was 6.3% (P<.001; 95% CI 5.1%-7.5%) higher than posts generated by other participants. Participants also had a 6.5% (P<.001; 95% CI 4.9%-8.1%) and 6.1% (P<.001; 95% CI 4.1%-8.1%) higher probability of engaging with posts that directly mentioned weight and goals, respectively, than other types of posts. Participants were 44.8% (P<.001; 95% CI 42.8%-46.9%) and 46% (P<.001; 95% CI 44.1%-48.0%) more likely to engage with a post when they were replied to by other participants and by interventionists, respectively. A 1 SD decrease in the sentiment of the conversation on a specific post was associated with a 5.4% (P<.001; 95% CI 4.9%-5.9%) increase in the probability of participants' subsequent engagement with the post. Participants' engagement in previous posts was also a predictor of engagement in subsequent posts (P<.001; 95% CI 0.74%-0.79%). Moreover, using a machine learning approach, we confirmed the importance of the predictors previously identified and achieved an accuracy of 90.9% in terms of predicting participants' engagement using a balanced testing sample with 1600 observations.

Conclusions: Findings revealed several predictors of engagement derived from the content generated by interventionists and other participants. Results have implications for increasing engagement in asynchronous, remotely delivered lifestyle interventions,

which could improve outcomes. Our results also point to the potential of data science and natural language processing to analyze microlevel conversational data and identify factors influencing participant engagement. Future studies should validate these results in larger trials.

Trial Registration: ClinicalTrials.gov NCT02656680; https://clinicaltrials.gov/ct2/show/NCT02656680

(JMIR Form Res 2022;6(7):e38068) doi:10.2196/38068

KEYWORDS

weight loss; social media intervention; engagement; data science; natural language processing; NLP; social media; lifestyle; machine learning; mobile phone

Introduction

Background

Obesity is prevalent in the United States and is a known risk factor for cardiovascular disease, type 2 diabetes [1,2], and cancer. Although lifestyle interventions are effective for weight loss and diabetes prevention [3], they require numerous clinic visits for up to a year, which is burdensome for many people. Technology-delivered lifestyle interventions, by not requiring visits, are less burdensome for participants and have shown promising weight loss outcomes [4]. Some technology-based interventions use popular commercial social media platforms such as Facebook in an effort to meet people where they are [5,6]. Many social media users already use these platforms to discuss their health experiences [7,8]. Community-building features on social media platforms, such as private groups [9,10], make them particularly amenable to delivering group-based lifestyle interventions.

Systematic reviews and meta-analyses show support for the efficacy of social media–delivered lifestyle interventions [4,11]; however, this area of research is still nascent. Participant engagement, defined as posts in the group, replies to posts, "likes," and votes in polls, appears to be an important predictor of outcomes [12-15]. For example, a study found that every 10 posts by participants corresponded to −0.5% weight loss [16]. Another study found only certain types of engagement predicted weight loss [17]. Interestingly, the degree of participant engagement reported in studies of social media–delivered interventions is highly variable, ranging from an average of once during the entire intervention period to 11 times per week [17-19]. Engagement also generally declines over time in these programs [16]. Our understanding of the factors that influence participant engagement in these interventions is limited. Emerging evidence in the web-based communication literature shows the importance of multilevel factors influencing participant engagement, such as the characteristics of posts (eg, post length and topic and popularity of the poster), characteristics of the conversation thread in response to the post (eg, sentiment and reciprocity behavior), and participant characteristics (eg, motivation and habits) [14,20-27]. However, research is scant in the context of social media–delivered behavioral interventions [28]. Furthermore, characteristics of the conversation thread (ie, other peoples' replies and comments) preceding each participant's engagement is often ignored, which could be valuable in terms of providing microlevel contextual data that influences each participant's decision to engage.

A promising approach to increase our understanding of the factors influencing participants' engagement in social media–based behavioral interventions is to study the content and interactions generated by the interventionists and participants during the intervention using natural language processing (NLP). Data collected directly from web-based platforms (eg, Facebook) can provide detailed, real-time behavioral information over the course of intervention programs. NLP can handle a large quantity of text, generate reliable qualitative coding [29], be leveraged to derive various real-time microlevel insights concerning the characteristics of the posts (eg, topics) and conversations that followed the posts (eg, sentiment), and understand how they affect participants' decisions of engagement independently and aggregately. This has potential implications for the design and implementation of future interventions to increase participant engagement, which could lead to more favorable weight loss outcomes.

Objective

Using data from a 16-week pilot feasibility randomized weight loss trial that delivered lifestyle interventions via Facebook, drawing on multilevel factors influencing participant engagement identified by previous web-based communication literature, we derived various factors from the content generated by participants and interventionists over the course of the intervention, including characteristics of the posts (eg, poster, time, and topic), conversations that followed the post (eg, sentiment and receiving replies), and participants' previous engagement behaviors, and assessed how well these factors predict participant engagement individually and all together in the context of a social media–delivered lifestyle intervention.

Methods

Study Design, Settings, and Participants

In a pilot feasibility randomized trial, we randomized 80 participants who were overweight or obese into 1 of 2 remotely delivered lifestyle interventions. We recruited people interested in losing weight via web-based advertisements at the University of Connecticut on ResearchMatch and in yard sale or neighborhood Facebook groups in 37 states across the United States between June and October of 2019. Inclusion criteria included BMI between 27 and 45 kg/m², smartphone ownership, active Facebook user (ie, comments or posts more than once a week), aged 18 to 65 years, and having daily internet access. Exclusion criteria included pregnancy or planning to become pregnant during the study, bariatric surgery or plans for bariatric surgery during the study, ≥5% weight loss in the past 3 months,

pre-existing conditions that preclude physical activity or dietary changes, taking medications affecting weight, incapable of walking one-fourth of a mile unaided without stopping, type 1 or type 2 diabetes, and participation in prior weight loss studies under the principal investigator.

Participants completed an orientation webinar before randomization to learn more about the study, and those still interested in participating were mailed a Wi-Fi scale (FitBit Aria, FitBit Inc) and asked to provide the staff with their log-in information for the scale so that weights could be recorded for the assessments. We randomized 80 participants to the 2 conditions.

Intervention Conditions

Overview

Participants were randomized to either a Facebook group in which new participants were continually enrolled during weeks 1 to 8 (open enrollment) or a Facebook group that included only the original 40 randomized participants (closed enrollment). In the open enrollment condition, 54 additional participants were enrolled between weeks 1 to 8 for a final group size of 94. However, we only included the original 80 randomized participants in this study to ensure all participants had an equal amount of time to engage in all 16 weeks of the intervention.

Facebook-Delivered Lifestyle Intervention

Both conditions received the identical 16-week lifestyle intervention based on the Diabetes Prevention Program (DPP) but modified to be delivered in a private Facebook group where twice-daily posts guided participants through the program, which was led by a dietitian (counselor) who was assisted by a student counselor. We adapted the DPP content to be appropriate for a web-based setting, as described elsewhere [30]. We gave each participant an individualized calorie goal that would facilitate a 1 lbs to 2 lbs weight loss weekly and asked them to use MyFitnessPal to track their calories daily. In addition, we asked participants to have the counselor review at least 2 weeks of their MyFitnessPal records, although they could request more as needed. The Facebook group was private such that only those invited by the study team could join the group and view the intervention content. We gave participants diet and exercise goals for the week each Monday and asked them to report progress on their goals in a conversation thread on Sunday and report their weight change for the week in a conversation thread each Friday. In between, intervention posts addressed the learning objectives of each module of the DPP. The dietitian leading the group was instructed to reply to all participant posts and comments that merited a reply and otherwise hit a reaction (eg, like and laugh) button to acknowledge the participants' comments. Participants were encouraged to post to the group and reply to each other. The details of the intervention, study procedures, and primary study results can be found elsewhere [31].

Ethics Approval

This pilot feasibility randomized trial was approved by the University of Connecticut Institutional Review Board (H17-215) in October 2017.

Measures

Overview

We included all posts and comments or replies within posts from interventionists and randomized and nonrandomized participants to construct the measures. Posts without text (approximately 6% of the posts were excluded) and polls were excluded, which resulted in 761 posts and 9396 comments or replies across the 2 intervention arms.

The outcome of interest was on the postparticipant pair level; that is, whether each participant had engaged with (ie, commented or replied to) a post in the Facebook group (1 if yes and 0 if no). Comments are in response to the original post, whereas replies are responses to comments made by others on a post. We focused on comments and replies as these activities are active forms of engagement rather than passive types of engagement such as views and reactions (eg, "likes") and have been shown to positively predict weight loss [16,32,33]. We extracted engagement data from the private Facebook groups using the Grytics app [34]. The Grytics app allowed us to download all the content posted in each Facebook group into Microsoft Excel sheets along with its unique Facebook ID number (post ID, comment ID, and parent comment ID), time stamp, reaction data, and author. Participants were asked to allow the Grytics app to access their Facebook account name so that the content from the group was identified (ie, post or comment author name was included in export).

Post Characteristics

The post characteristics described in Textbox 1 were included.

Textbox 1. Post characteristics.

Original poster

- We used a binary variable to indicate whether the focal post was created by the interventionist (1) or participant (0).

Post length

- We measured the number of words in each post.

Content sentiment

- We measured the average sentiment (text polarity) of each post's content. Text polarity measures the valence and emotion in the text and ranges on a continuous spectrum from negative (lower value) to positive (higher value). We standardized the measure of sentiment for the analysis. Sentiment analysis was performed using the *sentimentr* package in R (version 3.6.1).

Topics

- We used natural language processing to identify the topics that appeared in each post, comment, and reply. The content was preprocessed to remove emojis and non-English characters. Topics were detected using *Top2Vec* in Python 3.10.0, a deep learning–based sentence embedding algorithm that detects topics in the documents. It detected 117 unique topics along with their top words (see Multimedia Appendix 1 for a random sample of 35 topics with their top 8 words) from 10,157 pieces of content (posts, comments, and replies), which were further consolidated and coded under 7 topics based on the top 20 words of each original topic: exercise, diet, weight, MyFitnessPal app use, expressing emotion, sleep, and goals or plans. Here we focused on the topics of each post and created 7 binary variables to represent whether each post involves each of the aforementioned topics. Each post could include multiple topics; for example, a post mentioning a specific dietary goal would be categorized under both diet and goals or plans.

Time of the post

- We collected the time (number of days from day 1 of the intervention) and day of the week when the post was created.

Reply or Comment Characteristics

We constructed a series of variables representing the characteristics of replies or comments on each post. To reflect the content of conversations before each participant's engagement, if the participant engaged with the post, we calculated these variables based on all previous comments or replies under the post before their engagement for each unique postparticipant pair; if the participant did not engage with the post, we calculated these measures based on all the comments or replies under the post. The characteristics are described in Textbox 2.

Textbox 2. Reply or comment characteristics.

Tags or mentions

- We created two binary variables to represent whether each participant had been tagged or mentioned by (1) interventionists or (2) other participants in the previous replies or comments within the same post. It is worth mentioning that most tags or mentions in our data were generated automatically by Facebook (eg, when participant A comments or replies to participant B's content, Facebook automatically generates a tag on B in A's reply or comment). Thus, most tags or mentions in our data represent reply or comment relationships. In very few instances, interventionists deliberately tagged previously disengaged participants; however, the sample size was too small to test their effects separately.

Reply or comment content sentiment

- We measured the average sentiment of all replies or comments for each postparticipant pair. The measure was standardized for the analysis.

Participants' Characteristics

The included participant characteristics were as follows:

- Percentage of previous posts commented or replied: For each post, we calculated the percentage of previous posts each participant has commented or replied to.
- Baseline and sociodemographic characteristics: Although these variables were not the focus of our analysis, we included baseline characteristics for each participant, including treatment condition (open vs closed), baseline weight, BMI, age, race, sex, education, marital status, number of people in the household, and employment status, as covariates in the analyses.

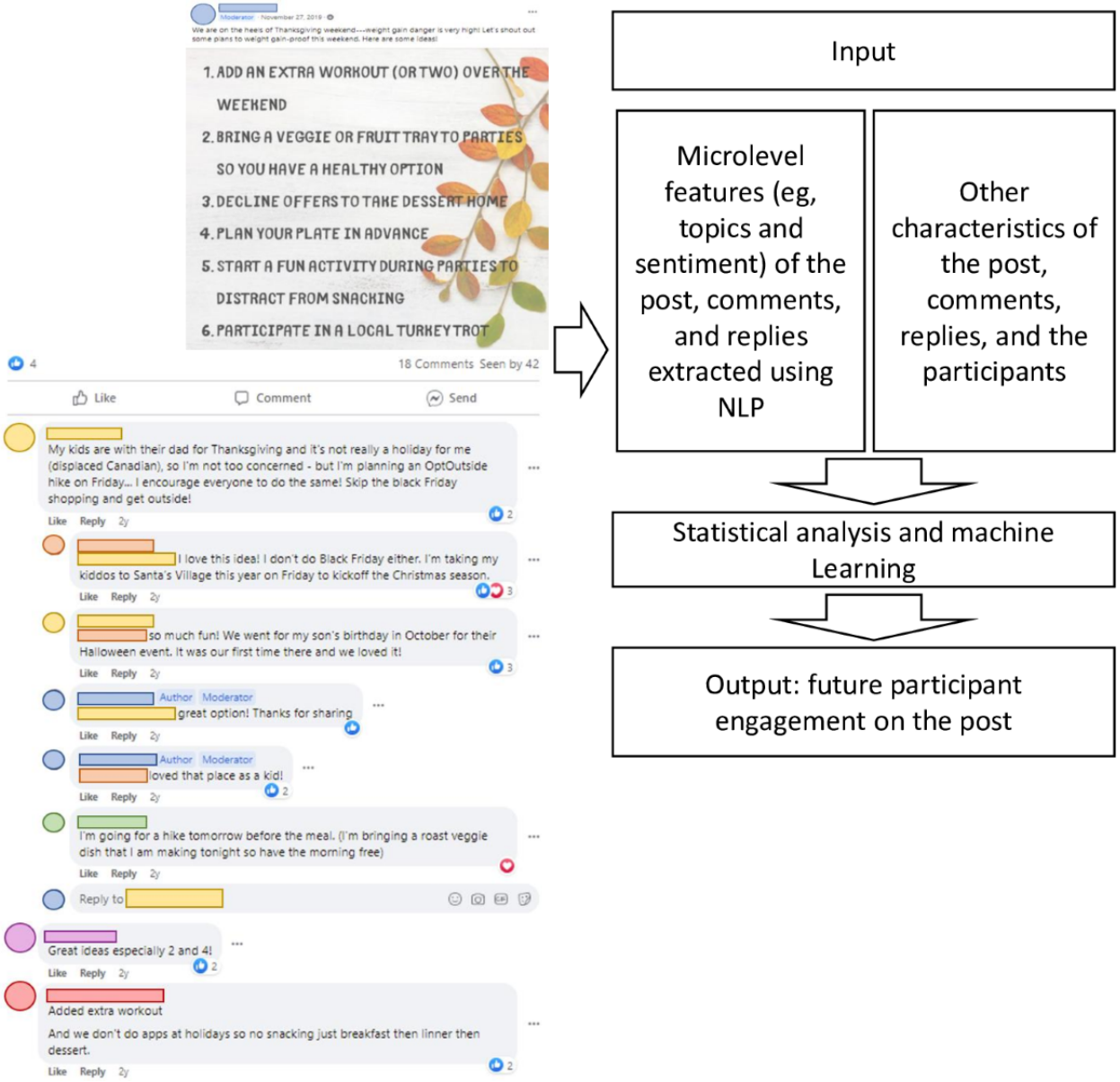
Statistical Analysis

We focused our analysis on whether each randomized participant (N=80) had engaged with each post, as randomized participants had access to the Facebook group the entire length of the intervention (it should be noted that each post was only available in a particular treatment arm and, thus, can only be seen by 40 randomized participants). To examine what predicts participant engagement with each post, analyses were performed on the postparticipant pair level (ie, whether each participant engaged with each post). This allowed us to construct measures that accurately reflect the content (ie, posts and conversations) before each participant's engagement. We included all possible engagements (ie, instances where participants engaged and instances where they did not engage) from the 80 randomized

participants with each of the 761 posts, which resulted in a final sample of 31,968 observations (participants engaged in 4462 instances and did not engage in 27,506 instances) for our analysis.

The overall analysis framework is depicted in Figure 1. Data were screened for deviations from assumptions required for the used statistical analyses. We calculated descriptive statistics for the outcome and key independent variables for each treatment condition. To account for the fact that engagement was nested within each post, we performed a mixed-effects regression with postlevel random effects using participants' engagement as the primary outcome, along with all aforementioned key predictors (ie, characteristics of the post, reply or comment, and the participants' previous engagement behavior) as independent variables, with participants' baseline and sociodemographic characteristics as covariates. We also included participant-level fixed effects as an alternative specification to account for possible omitted variable bias. As a robustness check, we also conducted a mixed-effects logistic regression with the same variables (Multimedia Appendix 1). To identify the important predictors of participant engagement, we reported the coefficient, 95% CI, and the associated *P* value for each predictor. All analyses were performed in STATA standard edition (version 16).

Figure 1. Analysis framework to identify important predictors of participant engagement. Left panel: an example of the intervention post and the comments or replies following it. Right panel: flow chart of the analysis. NLP: natural language processing.



Although regression analyses are useful to identify the statistical significance of linear relationships, some of the relationships might be much more complex (eg, nonlinear or moderated by other variables). To confirm the importance of significant predictors that we previously identified and to investigate how well these variables as a whole can predict participants' engagement with a particular post, we included all aforementioned predictors in machine learning algorithms, including gradient boosting machines, deep learning models, and an ensemble of them [35,36], and examined the performance of these models by calculating key metrics, including (1) area under the curve (ranging from 0.5 to 1, with 1 being the best)

using a 5-fold cross-validation, (2) variable importance across different models, and (3) out-of-sample prediction accuracies on a balanced sample with 1600 observations (800 random observations with engagement and 800 observations without). In the machine learning model, variable importance was determined by calculating the relative influence of each variable: in the tree-based model, it was calculated based on whether that variable was selected or included during the tree-building process and how much it improves the model fit. In other non–tree-based models, it was calculated as the magnitude of the weight or coefficient connecting a specific input or variable to the output [37]. We also evaluated the variable importance using an alternative approach called Shapley Additive Explanations contribution in one of the top performing models, which measures how much the average model prediction would change with and without a specific feature or variable [38], as shown in Multimedia Appendix 1. All the analyses were performed using the *h2o* package [39] in R (R Foundation for Statistical Computing; version 3.6.1).

Results

Table 1 presents the sociodemographic characteristics and key measures of engagement for the participants (N=80). The mean age of participants was 40.2 (SD 11.2) years with a mean baseline BMI of 34.4 (SD 5.0) kg/m^2. Participants were predominantly female (68/80, 85%) and mostly White (72/80, 90%) White, with most (58/80, 73%) reporting completing college or graduate school. The baseline characteristics of the participants in the 2 treatment conditions were similar, and we did not observe significant differences in these variables between the groups. Over the 16-week intervention, participants in the closed group commented or replied to 9.9% (37/374; SD 10.4%) of the posts on average, whereas participants in the open group commented or replied to 8.8% (34/387; SD 9.7%) of the posts on average.

Table 2 presents the key characteristics of the posts and comments or replies generated over the 16-week intervention across the 2 treatment arms. Post length was 33.4 words on average (SD 23.81) and 57.3% (436/761) of the posts were created by the interventionists. It should be noted participants in the 2 groups were exposed to identical program posts (whereas the number of self-generated posts by interventionists could be different). Topic modeling results showed that diet was the most popular topic across all posts (310/761, 40.7%), followed by exercise (163/761, 21.4%), goal or plan (152/761, 19.9%), and weight (138/761, 18.1%). We did not observe significant differences in post or comment or reply characteristics across groups, except that the percentage of replies or comments directed at randomized participants was significantly higher in the closed group than the open group (P<.001), possibly because there were 54 nonrandomized participants in the open group.

Table 1. Participant characteristics (N=80).

Participant characteristics	Closed enrollment (n=40)	Open enrollment (n=40)
Age (years), mean (SD)	40.4 (11.8)	40.0 (10.6)
Female, n (%)	34 (85)	34 (85)
Baseline BMI (kg/m^2), mean (SD)	34.8 (5.4)	34.0 (4.6)
Hispanic or Latino, n (%)	3 (8)	1 (3)
Race, n (%)		
White	36 (90)	36 (90)
Black or African American	3 (8)	3 (8)
Asian	0 (0)	0 (0)
Native Hawaiian or other Pacific Islander	0 (0)	0 (0)
American Indian or Alaska Native	0 (0)	0 (0)
Multiethnic	0 (0)	1 (3)
Unknown	1 (3)	0 (0)
Marital status, n (%)		
Married or living with partner but not married	29 (73)	30 (75)
Single	8 (20)	6 (15)
Widowed, divorced, or separated	3 (8)	4 (10)
Education, n (%)		
Less than high school, high school degree, GED[a], equivalent	1 (3)	2 (5)
Trade, technical, some college, associates	8 (20)	11 (28)
Bachelor's degree or some graduate school	21 (53)	17 (43)
Graduate degree	10 (25)	10 (25)
Employment status, n (%)		
Employed full-time	28 (70)	27 (68)
Employed part-time	7 (18)	4 (10)
Student	2 (5)	2 (5)
Unemployed, retired, disabled, or homemaker	3 (8)	6 (15)

[a]GED: General Educational Development.

Table 2. Post and reply or comment characteristics over the 16-week intervention.

Post characteristics	Closed enrollment (n=374)	Open enrollment (n=387)
Content sentiment, mean (SD)	0.134 (0.197)	0.133 (0.195)
Number of words, mean (SD)	33.78 (24.63)	33.28 (23.12)
Created by interventionists, n (%)	225 (60.2)	211 (54.5)
Topic, n (%)		
Exercise	80 (21.4)	83 (21.4)
Diet	158 (42.2)	152 (39.3)
Weight	64 (17.1)	74 (19.1)
MyFitnessPal app	61 (16.3)	67 (17.3)
Expressing emotion	28 (7.5)	28 (7.2)
Sleep	6 (1.6)	5 (1.3)
Goals or plans	80 (21.4)	72 (18.6)
Reply or comment characteristics[a]		
Content sentiment, mean (SD)	0.171 (0.255)	0.156 (0.234)
Participant's reply to other participants, n (%)	750 (23.8)	803 (12.9)
Interventionist reply to a participant, n (%)	1018 (32.3)	1195 (19.1)

[a]Closed enrollment n=3152 and open enrollment n=6244.

Table 3 shows the results from mixed-effects regression models on how well each variable predicted participants' engagement with a specific post. Regarding post characteristics, we found that the overall probability of participants' engagement with each post decreased by 0.04% each day (P<.001; 95% CI 0.02%-0.06%). Participants had a 6.3% (P<.001; 95% CI 5.1%-7.5%) higher probability of engaging with posts generated by the interventionists than with posts created by other participants. Post length also mattered—one additional word in a post's content was associated with a 0.05% (P<.001; 95% CI 0.03%-0.08%) increase in participants' probability of engagement. Participants also had a 6.5% (P<.001; 95% CI 4.9%-8.1%) and 6.1% (P<.001; 95% CI 4.1%-8.1%) higher probability of engaging with posts if the post content was related to weight and goals or plans, respectively. Regarding reply or comment characteristics, participants were 44.8% (P<.001; 95% CI 42.8%-46.9%) more likely to engage with a post when they received replies from other participants in the conversation or 46% (P<.001; 95% CI 44.1%-48.0%) more likely to engage if they received replies by interventionists. A 1 SD decrease in the sentiment in the previous replies or comments was associated with a 5.4% (P<.001; 95% CI 4.9%-5.9%) increase in the probability of participant engagement. Participants' engagement in previous posts was a strong predictor of future engagement—a 1% increase in participants' previous engagement was associated with a 0.76% (P<.001; 95% CI 0.74%-0.79%) increase in their probability to engage with a subsequent post. Robustness analyses showed that these results were largely consistent with the results from (1) multivariate linear regression results and mixed-effects regression with participant-level fixed effects and (2) mixed-effects logistic regression with or without postlevel random effects and participant fixed effects. Details of the results from these additional regression analyses can be found in Multimedia Appendix 1.

To confirm the importance of previously identified predictors and test how well the aforementioned variables can predict the probability of a participant engaging with a post, we performed a variety of machine learning–based classification algorithms with all the aforementioned predictors as the input and participant engagement as the outcome. Of the 32 models we tested, the ensemble approach of gradient boosting machine learning–based and deep learning–based classification algorithms performed the best, with an average area under the curve of 0.963 using 5-fold cross-validation (see more results in Multimedia Appendix 1). Figure 2 shows the variable importance across 20 machine learning models (excluding ensemble models) we tested, with those indicated by yellow and red being more important variables. The results show that receiving a reply from other participants and interventionists, percentage of previous posts participants had engaged in, average sentiment in previous replies or comments, time of the post, and day of the week were the most important variables across models, which were consistent with the results from regression analyses. Finally, we performed out-of-sample predictions on a balanced sample with 1600 observations (800 observations with engagement and 800 observations without) and achieved 90.9% accuracy and 0.908 F_1 score at maximum.

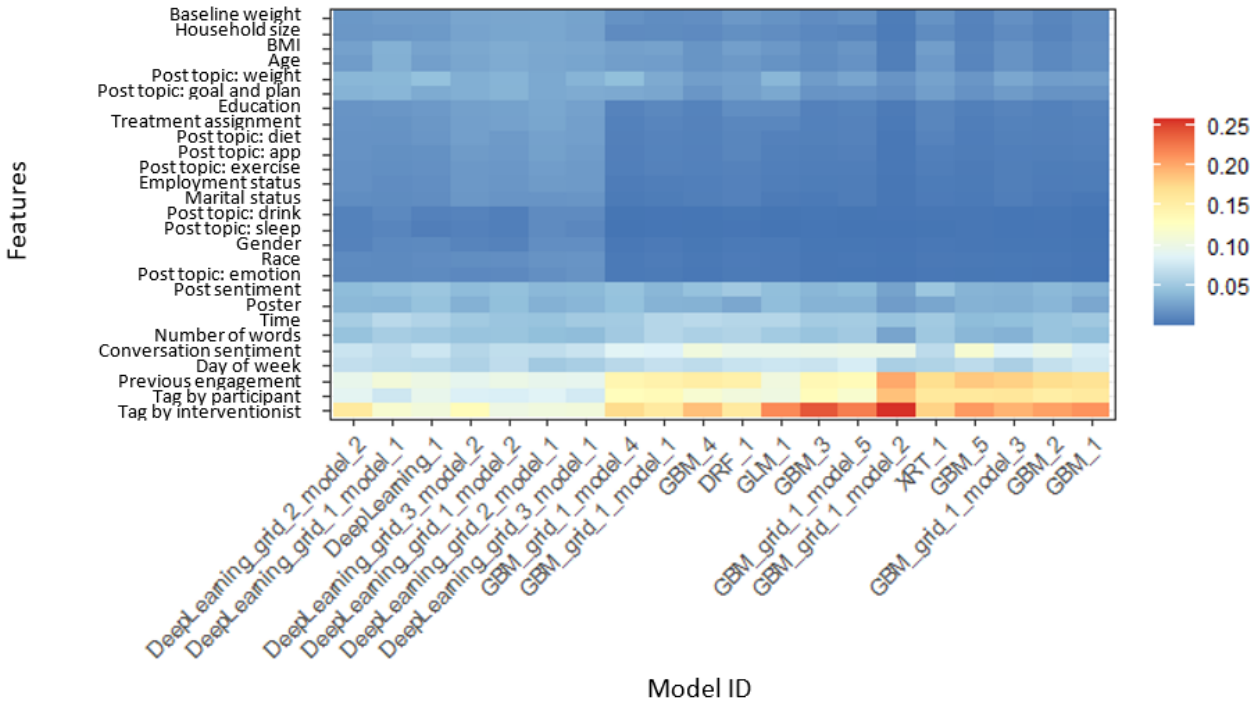
Table 3. Mixed-effects regression results predicting participants' engagement (N=31,968)[a].

Mixed-effects regression	Values, mean (SD; range)	Coefficient (95% CI)	*P* value
Outcome: participants' engagement	0.140 (0.347; 0 to 1)	—[b]	—
Post characteristics			
Created by interventionists	0.583 (0.493; 0 to 1)	0.0627 (0.0507 to 0.0746)	<.001
Number of words	33.44 (23.80; 1 to 107)	0.0005 (0.0003 to 0.0008)	<.001
Content sentiment (standardized)	0 (1; −3.67 to 4.74)	−.0042 (−0.097 to 0.0012)	.13
Topics			
Exercise	0.212 (0.409; 0 to 1)	0.0096 (−0.0075 to 0.0266)	.27
Diet	0.389 (0.487; 0 to 1)	−0.0085 (−0.0249 to 0.0078)	.31
Weight	0.191 (0.392; 0 to 1)	0.0654 (0.0494 to 0.0814)	<.001
MyFitnessPal app	0.163 (0.369; 0 to 1)	−0.0377 (−0.0534 to −0.0219)	<.001
Expressing emotion	0.071 (0.257; 0 to 1)	0.0083 (−0.01558 to 0.0321)	.50
Sleep	0.0141 (0.117; 0 to 1)	−0.0587 (−0.1070 to −0.0103)	.02
Goals or plans	0.209 (0.407; 0 to 1)	0.0612 (0.0414 to 0.0811)	<.001
Time of the post	46.61 (32.94; 1 to 112)	−0.0004 (−0.0006 to −0.0002)	<.001
Reply or comment characteristics			
Content sentiment (standardized)	0 (1; −7.84 to 5.91)	−0.0539 (−0.0589 to −0.0488)	<.001
Replied by other participants	0.026 (0.161; 0 to 1)	0.4484 (0.4279 to 0.4690)	<.001
Replied by interventionists	0.029 (0.167; 0 to 1)	0.4604 (0.4409 to 0.4798)	<.001
Participants characteristics			
Percentage previous posts commented or replied	13.16 (13.68; 0 to 100)	0.0076 (0.0074 to 0.0079)	<.001

[a]The model included postlevel random effects and also controlled for day of the week when the post was created and other baseline and sociodemographic characteristics of the participants, including treatment assignment, race, marital status, education, employment, number of people in the household, age, gender, baseline BMI, and weight.

[b]Not available.

Figure 2. Variable importance of predicting participant engagement across 20 machine learning models. The x-axis shows different model names, and variables from top to bottom on the y-axis are baseline weight, number of people in the household, baseline BMI, age, post topic weight, post topic goal or plan, education, treatment assignment, post topic diet, post topic MyFitnessPal app, post topic exercise, employment status, marital status, post topic drink, post topic sleep, gender, race, post topic expressing emotion, post sentiment, whether the post is created by interventionists, day of the intervention when the post is created, word count of the post, replies or comments sentiment, day of the week when the post is created, percentage of previous posts engaged, and whether replied by other participants or by interventionists.



Discussion

Principal Findings

In this study, we conducted secondary analyses using data from a 2-arm pilot feasibility randomized controlled trial that delivered lifestyle interventions via Facebook. We analyzed commenting or replying behavior from 80 participants in response to each of the 761 posts generated by counselors and participants over the 16-week intervention period and linked them to predictors, including post characteristics (eg, time, post length, and post topic), conversation characteristics (eg, sentiment of the conversation and participants being replied to), and participant characteristics (eg, sociodemographics and previous commenting or replying behavior). Our findings suggest that although participants' comments or replies decreased over time, important characteristics of the post, the conversation attached to that post, and the participants' engagement patterns predicted whether a participant engaged with a specific post. For example, we found that participants who engaged more with prior posts were more likely to engage with future posts. Posts that were longer (with the maximum number of words not exceeding 107), were created by interventionists, or had content related to weight (eg, weigh-in posts) and goal setting are more likely to attract engagement. The latter is consistent with the design of the intervention—participants were asked to set diet and exercise goals for the week each Monday, report their progress on their goals on Sunday, and report their weight change for the week each Friday. This is also encouraging because goal setting [40] and frequent self-weighing [41] are key behavioral weight loss strategies. Furthermore, posts with replies or comments that directly mentioned the focal participant were much more likely to attract subsequent replies from that participant. Moreover, posts with replies or comments that contained negative sentiments were more likely to attract subsequent comments. This is likely because participants who share struggles, problems, and challenges are naturally more negative in sentiment (see Multimedia Appendix 1 for examples of replies or comments with negative sentiments), and such content often attracts support and brainstorming from other participants who may feel called to help when others are struggling. These results were robust to multiple alternative specifications. Machine learning results also show that together, these characteristics can predict participant engagement with a high accuracy of 90.9%.

Implications

In this study, we demonstrate the potential of using NLP tools to analyze microlevel conversational data and identify factors influencing participants' commenting or replying behavior in a social media–delivered weight loss intervention. Our findings shed light on some important microlevel characteristics of the participants, posts, and conversations, which can shape participants' experiences during the intervention and predict their future engagement. These results have implications for the design and implementation of social media–delivered behavioral interventions in ways that maximize participant engagement. We previously reported a strong association between participant engagement and weight loss [31], which suggests that engagement-enhancing strategies could lead to more favorable outcomes. For example, enhancing engagement

early on may help with sustained engagement. Furthermore, we found that receiving replies appears to stimulate further engagement from the participant. This may also be a function of whether a participant shares something substantive about themselves in a comment. For example, if a participant's comment is "thanks!" the interventionist may just hit the "like" button; however, if a participant's comment is a question or sharing of a goal, the interventionist and other participants are more likely to reply to continue that conversation. Further research should explore the type of comments (eg, questions, sharing a problem, and setting a goal) that are most likely to elicit replies from others. Program content should be designed in ways that nudge participants to post more often and interact more with each other (eg, use open-ended questions and encourage peer-to-peer support). Although in this study we found that participants were more likely to respond to interventionists, greater peer-to-peer engagement could also facilitate stronger group cohesion, thereby further enhancing participant engagement. Posts in which participants share struggles, problems, and challenges they have encountered during the weight loss process may draw more participants into the conversation, which may generate richer brainstorming and social support, both of which could also enhance group cohesion. These implications could be applicable not only to social media–delivered weight loss interventions but also to other digital health interventions more generally.

Comparison With Prior Work

Although many studies have tested social media–delivered weight loss interventions or emphasized the importance of participant engagement in web-based communities [12-15,42], only a handful of studies have identified the factors that can influence participant engagement in digital health interventions [43,44], and most of them focused on macrolevel factors such as post type and participants' characteristics (eg, age and gender) [9,14]. Few studies have examined microlevel factors such as conversation dynamics. Several studies recognized this limitation and called for more research to identify all relevant factors that can predict participant engagement [28,43,45,46]. This study contributes to this line of research in two ways: (1) we demonstrate how content generated by interventionists and participants during the course of a digital intervention can be leveraged and combined with data science and NLP tools to identify microlevel predictors of participant engagement, and (2) we have identified many microlevel factors that influence participant engagement, which, to the best of our knowledge, have not been studied in previous social media–based behavioral interventions. This has practical implications for future intervention designs that can maximize participant engagement.

Similar to previous studies, we found that participant engagement is highly variable [17-19], and it generally declines over time [16,47]. Although many factors identified in this study have not been studied in the social media–based behavioral intervention context, our findings are consistent with psychological and sociological theories, as well as several empirical research on web-based communication. For example, our finding that posts created by interventionists are more popular is consistent with other studies on web-based communities, showing that important users or those with designated roles are more likely to draw responses from other users [26], possibly because of preferential attachment [48]. Our finding that replies or comments with negative sentiments draw more engagement implies that participants are more likely to reply or comment when they see others sharing their struggles and challenges. This could be possibly explained by social support processes, which have substantial empirical support across various web-based settings [22-25,49]. Finally, the importance of being replied to by interventionists and other participants can be explained by the preference for reciprocity [50], which has been found to be an important driver for communication in many other web-based settings [20,21].

Limitations and Future Work

This study has several limitations that point to avenues for future research. First, our sample size was small (80 participants; 10,157 total posts, comments, and replies) and our participants were predominantly White (72/80, 90%) and female (68/80, 85%). This limits the generalizability of our results, following a long-standing pattern in weight loss studies of difficulty recruiting male participants [51]. Similarly, given that 96% (77/80) of our participants reported attaining a college degree or advanced degree, we cannot generalize our results to individuals with lower levels of education. Future studies should devise recruitment strategies that attract more male participants and participants with low levels of education to further explore the individual heterogeneity across people from different backgrounds. Second, this study did not fully tease out all possible confounding factors and thus cannot establish causality. For instance, participants who are more successful in losing weight might also be more likely to comment simply because they are paying more attention to the group and have more to say as they are applying the knowledge and strategies they are learning. Future studies should include larger trials, surveys with more longitudinal measures (eg, physical activity and diet tracking, mental health, and other behaviors), and qualitative studies to establish the possible bidirectional and causal relationships between engagement and these factors. Third, we focused on replies and comments and did not explore other types of engagement such as reactions and views of posts and comments. Although comments and replies have been considered more substantive than other engagements (eg, likes and "lurking"), other engagements potentially comprise a substantial proportion of social media use and thus warrant careful consideration in future studies [33]. Similarly, although we included participants' posts in our analysis, we did not include posts only with images or videos or investigate the factors influencing participants' decisions to create posts, which is another important form of engagement. In addition, although we included tags or mentions relationships in our study, most tags were automatically generated by Facebook during replies or comments. Future studies should consider whether deliberate tagging can nudge disengaged participants to re-engage with the program. Fourth, we assumed all participants had an opportunity to engage with every post in the group, and we considered all replies or comments when constructing the predictors for participants who did not engage with a certain post, which might not necessarily be the case if the participant did not see the post or the previous comments or replies. Future

studies could take additional information into account, such as what participants viewed and the time a participant spends on Facebook. This will allow researchers to construct more refined measures to reflect the condition under which participants make a decision of whether to engage. Finally, although we observed that longer posts are more likely to draw engagement from participants, it should be noted the posts in this intervention were generally short by design (mean 33.4, maximum 107 words). Social media marketing reports reveal that Facebook posts that have <50 characters receive the highest level of engagement relative to longer posts [52]. A/B testing of a wide range of post lengths and different types of posts (eg, goal setting vs problem sharing) is needed to determine the ideal length of posts to maximize engagement in behavioral interventions.

Conclusions

In this study, we performed secondary analyses using data from a pilot feasibility randomized weight loss trial that delivered a lifestyle intervention via Facebook and linked participants' engagement with several important predictors, including characteristics of the posts, replies or comments, and participants. Our results point to the potential of using data science and NLP tools to analyze microlevel behavioral or conversational data and identify factors influencing participants' engagement during the social media weight loss intervention, which have implications for the design and implementation of future interventions that could lead to more favorable weight loss outcomes. Future studies are warranted to validate our results and further explore these relationships in similar and larger trials.

Acknowledgments

This project was supported by the National Institutes of Health grant K24HL124366 (principal investigator: SP).

Conflicts of Interest

SP has been a paid advisor for WW (formerly Weight Watchers) and FitBit.

Multimedia Appendix 1
Supplementary analyses and results.
[DOCX File , 91 KB - formative_v6i7e38068_app1.docx]

References

1. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity and severe obesity among adults: United States, 2017-2018. NCHS Data Brief 2020 Feb(360):1-8 [FREE Full text] [Medline: 32487284]
2. Wolin KY, Carson K, Colditz GA. Obesity and cancer. Oncologist 2010;15(6):556-565 [FREE Full text] [doi: 10.1634/theoncologist.2009-0285] [Medline: 20507889]
3. Pagoto S. The current state of lifestyle intervention implementation research: where do we go next? Transl Behav Med 2011 Sep;1(3):401-405 [FREE Full text] [doi: 10.1007/s13142-011-0071-x] [Medline: 24073065]
4. Beleigoli AM, Andrade AQ, Cançado AG, Paulo MN, Diniz MD, Ribeiro AL. Web-based digital health interventions for weight loss and lifestyle habit changes in overweight and obese adults: systematic review and meta-analysis. J Med Internet Res 2019 Jan 08;21(1):e298 [FREE Full text] [doi: 10.2196/jmir.9609] [Medline: 30622090]
5. Pagoto S, Waring ME, May CN, Ding EY, Kunz WH, Hayes R, et al. Adapting behavioral interventions for social media delivery. J Med Internet Res 2016 Jan 29;18(1):e24 [FREE Full text] [doi: 10.2196/jmir.5086] [Medline: 26825969]
6. Waring ME, Jake-Schoffman DE, Holovatska MM, Mejia C, Williams JC, Pagoto SL. Social media and obesity in adults: a review of recent research and future directions. Curr Diab Rep 2018 Apr 18;18(6):34. [doi: 10.1007/s11892-018-1001-9] [Medline: 29671135]
7. Fox S. The social life of health information. Pew Research Center. URL: https://www.pewresearch.org/fact-tank/2014/01/15/the-social-life-of-health-information/ [accessed 2022-06-22]
8. Heo Y, Park J, Kim J, Park H. The emerging viewertariat in South Korea: The Seoul mayoral TV debate on Twitter, Facebook, and blogs. Telematic Inform 2016 May;33(2):570-583. [doi: 10.1016/j.tele.2015.08.003]
9. Rus HM, Cameron LD. Health communication in social media: message features predicting user engagement on diabetes-related facebook pages. Ann Behav Med 2016 Oct;50(5):678-689. [doi: 10.1007/s12160-016-9793-9] [Medline: 27059761]
10. Nam Y, Lee Y, Park HW. Measuring web ecology by Facebook, Twitter, blogs and online news: 2012 general election in South Korea. Qual Quant 2014 Mar 16;49(2):675-689. [doi: 10.1007/s11135-014-0016-9]
11. An R, Ji M, Zhang S. Effectiveness of social media-based interventions on weight-related behaviors and body weight status: review and meta-analysis. Am J Health Behav 2017 Nov 01;41(6):670-682. [doi: 10.5993/AJHB.41.6.1] [Medline: 29025495]
12. Pagoto SL, Waring ME, Schneider KL, Oleski JL, Olendzki E, Hayes RB, et al. Twitter-delivered behavioral weight-loss interventions: a pilot series. JMIR Res Protoc 2015 Oct 23;4(4):e123 [FREE Full text] [doi: 10.2196/resprot.4864] [Medline: 26500186]

13. Turner-McGrievy GM, Tate DF. Weight loss social support in 140 characters or less: use of an online social network in a remotely delivered weight loss intervention. Transl Behav Med 2013 Sep;3(3):287-294 [FREE Full text] [doi: 10.1007/s13142-012-0183-y] [Medline: 24073180]
14. Hales SB, Davidson C, Turner-McGrievy GM. Varying social media post types differentially impacts engagement in a behavioral weight loss intervention. Transl Behav Med 2014 Dec;4(4):355-362 [FREE Full text] [doi: 10.1007/s13142-014-0274-z] [Medline: 25584084]
15. Pagoto S, Waring M, Olendzki E, Oleski J, May C, Evans M. The feasibility of incentivizing participation in an online social network weight loss program. In: Proceedings of the Hawaii International Conference on System Sciences. 2017 Presented at: Hawaii International Conference on System Sciences; Jan 4-7, 2017; Hawaii, USA.
16. Turner-McGrievy G, Tate D. Tweets, apps, and pods: results of the 6-month mobile pounds off digitally (mobile POD) randomized weight-loss intervention among adults. J Med Internet Res 2011 Dec 20;13(4):e120 [FREE Full text] [doi: 10.2196/jmir.1841] [Medline: 22186428]
17. Pagoto S, Waring M, Jake-Schoffman D. What type of engagement predicts success in a Facebook weight loss group? In: Proceedings of the Hawaii International Conference on System Sciences. 2018 Presented at: Hawaii International Conference on System Sciences; Jan 3 – 6, 2018; Hawaii, USA. [doi: 10.24251/hicss.2018.419]
18. West DS, Monroe CM, Turner-McGrievy G, Sundstrom B, Larsen C, Magradey K, et al. A technology-mediated behavioral weight gain prevention intervention for college students: controlled, quasi-experimental study. J Med Internet Res 2016 Jun 13;18(6):e133 [FREE Full text] [doi: 10.2196/jmir.5474] [Medline: 27296086]
19. Herring SJ, Cruice JF, Bennett GG, Rose MZ, Davey A, Foster GD. Preventing excessive gestational weight gain among African American women: a randomized clinical trial. Obesity (Silver Spring) 2016 Jan;24(1):30-36 [FREE Full text] [doi: 10.1002/oby.21240] [Medline: 26592857]
20. Constant D, Sproull L, Kiesler S. The kindness of strangers: the usefulness of electronic weak ties for technical advice. Org Sci 1996 Apr;7(2):119-135. [doi: 10.1287/orsc.7.2.119]
21. How open source software works: “free” user-to-user assistance. In: Produktentwicklung mit virtuellen Communities. Wiesbaden, Germany: Springer Verlag; 2004.
22. Kramer AD, 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]
23. Kramer A. The spread of emotion via Facebook. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12: CHI Conference on Human Factors in Computing Systems; May 5 - 10, 2012; Austin Texas USA. [doi: 10.1145/2207676.2207787]
24. Song Y, Lin Q, Kwon KH, Choy CH, Xu R. Contagion of offensive speech online: an interactional analysis of political swearing. Comput Human Behav 2022 Feb;127:107046. [doi: 10.1016/j.chb.2021.107046]
25. Centola D. The spread of behavior in an online social network experiment. Science 2010 Sep 03;329(5996):1194-1197. [doi: 10.1126/science.1185231] [Medline: 20813952]
26. Johnson SL, Faraj S, Kudaravalli S. Emergence of power laws in online communities: the role of social mechanisms and preferential attachment. MIS Q 2014 3;38(3):795-808. [doi: 10.25300/misq/2014/38.3.08]
27. Li J, Zheng H, Duan X. Factors influencing the popularity of a health-related answer on a Chinese question-and-answer website: case study. J Med Internet Res 2021 Sep 28;23(9):e29885 [FREE Full text] [doi: 10.2196/29885] [Medline: 34581675]
28. Pagoto S, Waring ME. A call for a science of engagement: comment on rus and cameron. Ann Behav Med 2016 Oct;50(5):690-691. [doi: 10.1007/s12160-016-9839-z] [Medline: 27663577]
29. Guetterman TC, Chang T, DeJonckheere M, Basu T, Scruggs E, Vydiswaran VG. Augmenting qualitative text analysis with natural language processing: methodological study. J Med Internet Res 2018 Jun 29;20(6):e231 [FREE Full text] [doi: 10.2196/jmir.9702] [Medline: 29959110]
30. Wang ML, Waring ME, Jake-Schoffman DE, Oleski JL, Michaels Z, Goetz JM, et al. Clinic versus online social network-delivered lifestyle interventions: protocol for the get social noninferiority randomized controlled trial. JMIR Res Protoc 2017 Dec 11;6(12):e243 [FREE Full text] [doi: 10.2196/resprot.8068] [Medline: 29229591]
31. Pagoto SL, Schroeder MW, Xu R, Waring ME, Groshon L, Goetz JM, et al. A Facebook-delivered weight loss intervention using open enrollment: randomized pilot feasibility trial. JMIR Form Res 2022 May 06;6(5):e33663 [FREE Full text] [doi: 10.2196/33663] [Medline: 35522466]
32. Cavallo DN, Martinez R, Webb Hooper M, Flocke S. Feasibility of a social media-based weight loss intervention designed for low-SES adults. Transl Behav Med 2021 Apr 26;11(4):981-992 [FREE Full text] [doi: 10.1093/tbm/ibaa070] [Medline: 32716040]
33. Edelmann N. What is lurking? A literature review of research on lurking. In: The Psychology of Social Networking Vol.1. Warsaw, Poland: De Gruyter Open Poland; 2016.
34. Grytics. URL: https://grytics.com/ [accessed 2022-07-13]
35. Natekin A, Knoll A. Gradient boosting machines, a tutorial. Front Neurorobot 2013;7:21 [FREE Full text] [doi: 10.3389/fnbot.2013.00021] [Medline: 24409142]

36. Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge, Massachusetts, United States: MIT Press; 2016.
37. Rifkin R, Klautau A. In defense of one-vs-all classification. J Mach Learn Res 2004;5:101-141. [doi: 10.5555/1005332.1005336]
38. Lundberg S, Lee S. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017 Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; Dec 4 - 9, 2017; Long Beach California USA.
39. R Interface for 'H2O'. Package 'h2o'. 2018. URL: https://cran.microsoft.com/snapshot/2018-10-04/web/packages/h2o/h2o.pdf [accessed 2022-07-11]
40. Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, IMAGE Study Group. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health 2011 Feb 18;11:119 [FREE Full text] [doi: 10.1186/1471-2458-11-119] [Medline: 21333011]
41. Steinberg DM, Bennett GG, Askew S, Tate DF. Weighing every day matters: daily weighing improves weight loss and adoption of weight control behaviors. J Acad Nutr Diet 2015 Apr;115(4):511-518 [FREE Full text] [doi: 10.1016/j.jand.2014.12.011] [Medline: 25683820]
42. Cavallo DN, Sisneros JA, Ronay AA, Robbins CL, Jilcott Pitts SB, Keyserling TC, et al. Assessing the feasibility of a web-based weight loss intervention for low-income women of reproductive age: a pilot study. JMIR Res Protoc 2016 Feb 26;5(1):e30 [FREE Full text] [doi: 10.2196/resprot.4865] [Medline: 26920252]
43. Escobar-Viera CG, Melcher EM, Miller RS, Whitfield DL, Jacobson-López D, Gordon JD, et al. A systematic review of the engagement with social media-delivered interventions for improving health outcomes among sexual and gender minorities. Internet Interv 2021 Sep;25:100428 [FREE Full text] [doi: 10.1016/j.invent.2021.100428] [Medline: 34401387]
44. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
45. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J Am Med Inform Assoc 2015 Jan;22(1):243-256 [FREE Full text] [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
46. Hale TM, Pathipati AS, Zan S, Jethwani K. Representation of health conditions on Facebook: content analysis and evaluation of user engagement. J Med Internet Res 2014 Aug 04;16(8):e182 [FREE Full text] [doi: 10.2196/jmir.3275] [Medline: 25092386]
47. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
48. Barabasi AL, Albert R. Emergence of scaling in random networks. Science 1999 Oct 15;286(5439):509-512. [doi: 10.1126/science.286.5439.509] [Medline: 10521342]
49. White M, Dorman SM. Receiving social support online: implications for health education. Health Educ Res 2001 Dec;16(6):693-707. [doi: 10.1093/her/16.6.693] [Medline: 11780708]
50. Exchange and Power in Social Life. Milton Park, Abingdon-on-Thames, Oxfordshire United Kingdom: Taylor & Francis; 1964.
51. Franz MJ, VanWormer JJ, Crain AL, Boucher JL, Histon T, Caplan W, et al. Weight-loss outcomes: a systematic review and meta-analysis of weight-loss clinical trials with a minimum 1-year follow-up. J Am Diet Assoc 2007 Oct;107(10):1755-1767. [doi: 10.1016/j.jada.2007.07.017] [Medline: 17904936]
52. Zote J. How long should social posts be? Try this social media character counter. SproutSocial. URL: https://sproutsocial.com/insights/social-media-character-counter/ [accessed 2022-06-22]

Abbreviations

DPP: Diabetes Prevention Program
NLP: natural language processing

Edited by A Mavragani; submitted 17.03.22; peer-reviewed by H Alhoori, N Maglaveras, P Kaur; comments to author 23.05.22; revised version received 22.06.22; accepted 23.06.22; published 28.07.22.

Please cite as:
Xu R, Divito J, Bannor R, Schroeder M, Pagoto S
Predicting Participant Engagement in a Social Media–Delivered Lifestyle Intervention Using Microlevel Conversational Data: Secondary Analysis of Data From a Pilot Randomized Controlled Trial
JMIR Form Res 2022;6(7):e38068
URL: https://formative.jmir.org/2022/7/e38068
doi:10.2196/38068
PMID:35900824

Original Paper

Developing a Health Care Transition Intervention With Young People With Spinal Cord Injuries: Co-design Approach

Emily Alice Bray[1], BCom, Grad Dip (Psych), MSW; Bronwyn Everett[1,2], RN, MSc (Hons), PhD; Ajesh George[1,2,3], BDS, MPH, PhD; Yenna Salamonson[1,2,4], BSc, RN, MA, PhD; Lucie M Ramjan[1,2], BN (Hons), RN, PhD

[1]School of Nursing and Midwifery, Western Sydney University, Penrith, Australia

[2]Ingham Institute Applied Medical Research, Liverpool, Australia

[3]School of Dentistry, Faculty of Medicine and Health, University of Sydney, Camperdown, Australia

[4]School of Nursing, Faculty of Science, Medicine and Health, University of Wollongong, Wollongong, Australia

Corresponding Author:
Emily Alice Bray, BCom, Grad Dip (Psych), MSW
School of Nursing and Midwifery
Western Sydney University
Locked Bag 1797
Penrith, 2751
Australia
Phone: 61 02 9685 9032
Email: 16251104@student.westernsydney.edu.au

Abstract

Background: Successful transition from pediatric to adult health care settings supports long-term health management and better overall outcomes in all domains. However, young people with spinal cord injuries (SCIs) continue to report challenges and unmet needs during the transition process. Including end users in health care research and intervention design is paramount as interventions designed in this way better meet their specific needs and are often more innovative. Although studies have reported involving young people with chronic conditions in the development of health care transition (HCT) interventions, few details have been provided as to how this was achieved.

Objective: This study outlined the co-design and development of an HCT intervention to support young people with SCIs. It contextualized the co-design process, methods, materials used, and steps implemented from defining the problem to conceiving and designing the solution. This was accomplished by understanding and listening to end users' needs and recommendations for HCT.

Methods: Using participatory methods, this qualitative study reports the co-design of an HCT intervention to support young people with SCIs and parents or caregivers. Two co-design workshops were conducted: one with young people with SCIs and one with parents and caregivers. Categories were defined through a hybrid deductive and inductive qualitative content analysis process that was informed by the Care Transitions Framework and guided the development of the HCT intervention. Following the creation of a prototype intervention, young people with SCIs, parents and caregivers, and key pediatric SCI stakeholders provided feedback on the intervention content and design in focus groups. Similar to the workshops, the focus groups were analyzed using a hybrid deductive and inductive qualitative content analysis process informed by the Care Transitions Framework. The Enhancing the Quality and Transparency of Health Research guidelines for qualitative research (Consolidated Criteria for Reporting Qualitative Research) were applied.

Results: Overall, 4 young people and 4 parents or caregivers participated in the co-design workshops. Key recommendations for the HCT intervention were that participants wanted a "one-stop shop" for all their transition information needs and an editable portable medical summary to take with them to appointments. On the basis of the analysis of participants' recommendations from the workshops, it was determined that a website would be an appropriate hosting platform for the interventions. The focus group feedback on the design and content of the prototype website was extremely positive, with minor recommendations for improvement.

Conclusions: This is the first study to co-design and develop an HCT intervention in partnership with young people with SCIs and parents and caregivers. Although the study sample was small, it has shown that it is possible to meaningfully engage and empower young people with SCIs and parents and caregivers in the co-design of an HCT intervention.

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2021-053212

(JMIR Form Res 2022;6(7):e38616) doi:10.2196/38616

KEYWORDS

co-design; participatory action research; health care transition; pediatric health care; adult health care; spinal cord injury

Introduction

Background

A key goal in the rehabilitation of children with a spinal cord injury (SCI) is to facilitate the attainment of a productive and satisfying life by addressing developmental milestones and the provision of education on managing the complex health issues that arise because of aging and living with an SCI [1]. Importantly, a seamless transition from the pediatric to the adult health care systems, termed health care transition (HCT), is a significant and critical factor in supporting the fulfillment of goals while fostering independence and improved health outcomes [2].

Supporting a seamless HCT has been on the international health care agenda for >3 decades [3-5], yet young people with chronic conditions continue to report facing a multitude of barriers to their move [6,7]. These barriers include fear of losing established relationships with pediatric providers and forming new ones with adult providers, inadequate preparation for and information on the adult health care system, lack of self-management skills and disease knowledge, and poor communication between the pediatric and adult health care providers [6,7]. Consequently, these barriers result in poor health outcomes such as nonadherence to treatment and medication, loss to follow-up, increased hospital admissions, and patient dissatisfaction [7]. As such, there is an opportunity to reduce the difficulties faced by young people with chronic conditions through the development of HCT interventions that prepare them for the move and improve the transfer process.

In recent years, there has been significant progress made toward improving HCT for young people with chronic conditions. Evidence suggests that a structured HCT process, including planning for transition, transfer assistance, and integration into adult services, can improve outcomes for young people with chronic conditions, such as patient satisfaction, population health, and the use of health care services [2]. However, to date, HCT research has been disease-specific, and studies are typically characterized at the lowest evidence level, making them difficult to apply in various contexts [2,8].

Similar to other young people with chronic conditions, young people with SCIs face many barriers and facilitators in their transition to adulthood with regard to both health care services and normative life milestones such as education, employment, social participation, and independent living [9]. A study conducted in the United States on 23 young people with SCIs and their caregivers identified processes within health care that acted as both a barrier to and facilitator of the transition to adulthood [9]. Facilitators included health care support comprising the transfer of medical records, clear communication of transition timelines and expectations, referrals to adult services, and collaboration between the pediatric and adult settings. Health care barriers to transition included complex adult services, limited resources, and minimal previous exposure to the adult health care setting. In particular, the study indicated that, when it came to health care, there was a need for more condition-specific education for local, nonspecialized health care providers; better communication among health care providers; and an accessible, concise, and comprehensive medical history [9]. Evidence on the availability of HCT interventions for young people with SCIs that attempt to address these transition needs is scarce, with only 1 Australian article offering an explanation of their HCT efforts [10].

As the importance of end-user involvement in health care research and intervention design has been increasingly recognized, there has been a shift away from end users being passive participants in research, where research is conducted on them, to active participation, where research is conducted with them [11,12]. One such research approach, participatory action research (PAR), is particularly useful in co-design. PAR involves researchers collaborating with service users and key stakeholders in a collective and reflective inquiry to understand and improve practices and situations [13]. This process acknowledges that participants have knowledge and expertise to share, which is particularly important for the disability community whose voices have too often been silenced [14].

Collaboration within the PAR process takes place through iterative cycles of “planning, acting, and review” [15], and co-design can be used to facilitate the action stages. The co-design process actively involves all stakeholders in identifying solutions to local problems using their experience and expertise to explore the current needs of service users and develop and test concepts before improving the prototype in an iterative process [12]. Interventions designed in this way better meet the specific needs of end users and are often more innovative [11].

Although studies have reported involving young people with chronic conditions in the development of HCT interventions, few details have been provided as to how this was achieved [16]. As such, more transparency is needed regarding the process, methods and materials used to include young people with chronic conditions in intervention development.

The aim of this qualitative study was to fill this knowledge gap by providing a detailed explanation of the process involved in understanding the needs and recommendations for HCT as part of the co-design and development of an HCT intervention with young people with SCIs. This is the first study to co-design and develop an HCT intervention in partnership with young people with SCIs and parents or caregivers. Please note hereafter and unless otherwise specified, the term caregivers will be used to denote parents or caregivers. Addressing the needs of individuals and the current gap in services has the potential to improve transition outcomes and the quality of life of children and young people with an SCI.

Research Context and Conceptual Framework

The work presented in this paper forms part of a wider 3-year study informed by a PAR approach that aimed to co-design, develop, implement, and evaluate an HCT intervention to support young people with SCIs in New South Wales (NSW), Australia. Further details on the current SCI services in NSW can be found in an evidence series by the Agency for Clinical Innovation [17].

The study protocol has been published elsewhere [18]. In summary, the 3 study phases were informed by the Care Transitions Framework [19]. The Care Transitions Framework is an adaptation of the Consolidated Framework for Implementation Research, an established conceptual framework in implementation science [19]. The framework guides the research and evaluation of care transition interventions within a variety of settings and can be used in parts or as a whole. Organized into 8 domains, the Care Transitions Framework provides a comprehensive guide to potential questions that can be explored depending on the nature of the research and its goals [19]. Five of the domains for the overall study are presented in Multimedia Appendix 1. The 3 remaining domains (external context, organizational characteristics, and characteristics and roles of providers) were addressed as part of the prestudy consultation process.

Methods

Study Design

The overarching 3-year PAR study consisted of 3 phases (Multimedia Appendix 1). Phase 1 used semistructured interviews to explore the experiences with HCT of young people with SCIs and caregivers and has been published elsewhere (Bray et al, under review). Briefly, these interviews revealed that young people with SCIs and caregivers faced barriers and had unmet needs in their transition to adult health care services (Bray et al, under review). During the phase 1 interviews, young people with SCIs and caregivers identified the need for a coordinated and streamlined handover from pediatric to adult health care providers and a “one-stop shop” for transition information, such as how it occurs, who to call for ongoing support and advice, and tips on how to transition successfully (Bray et al, under review).

This paper presents the findings from phase 2 of the PAR study, which consisted of 2 parts. Phase 2a built on the findings of the phase 1 interviews and engaged both young people with SCIs and caregivers as individuals with “lived experience” in co-design workshops to inform the development of a prototype HCT intervention. In continuation of the co-design process, phase 2b used focus groups to gather feedback on the prototype HCT intervention, leading to further refinement before phase 3 evaluation. As part of the PAR approach, the participatory methods used included collaboration, active engagement, and reflection that occurred through iterative cycles of “planning, acting, and review” [13,15]. This phase (phase 2) of the PAR study formed part of the acting phase of PAR and supported working collaboratively with the community of interest to identify actions necessary to achieve the desired outcomes.

This study has followed the Standards for Reporting Qualitative Research [20].

Ethics Approval

This study received ethics approval from the Western Sydney University Human Research and Ethics Committee (H14029) and was registered with the Australian New Zealand Clinical Trials Registry (ACTRN12621000500853). Participation in this study was voluntary. Written consent was obtained at the beginning of the study, and verbal consent was obtained at the beginning of each workshop or focus group. All young people aged <16 years also required written consent from a caregiver. Transcripts were deidentified, and participants were assigned a pseudonym to maintain confidentiality. Participants received an Aus $30 (US $20.30) e–gift card each as a *thank you* for their time.

Participants

Owing to the limited number of pediatric-onset SCI cases in Australia [21,22], participants were recruited from both metropolitan and rural areas of NSW. To be eligible for inclusion, participants needed to be young persons aged between 14 and 25 years and have sustained a pediatric-onset traumatic or nontraumatic SCI (at or before the age of 16 years) or be a parent or caregiver of the same. Individuals had to be preparing for or have made the transition from pediatric to adult health care services (including tertiary hospitals and community services). Sufficient English-language proficiency was another requirement to ensure that all participants could fully engage in the conversation. Exclusion criteria were individuals who were receiving rehabilitation treatment for an injury sustained in the previous 12 months. This exclusion criterion was implemented so as not to burden the individual or their family with the demands of participating in research or risk causing any additional emotional distress during this tremendous period of adjustment. Individuals with neural tube defects such as spina bifida were also excluded. Although the authors acknowledge that individuals with neural tube defects share many of the same clinical characteristics and complications as those with SCIs, these individuals also demonstrate distinct features [23]. Research on HCT also typically focuses purely on SCIs [9] or spina bifida [24,25], with children and young people with spina bifida reported as a separate group supported by their own services. As such, this study focused specifically on young people with pediatric-onset traumatic or nontraumatic SCIs.

Recruitment

Young people with SCIs and caregivers were recruited from the individuals who had participated in the previous phase of the overall study. We contacted all 9 participants first via email to determine availability. We then contacted the participants by phone to confirm their availability and book the workshop or focus group. Each participant was sent an SMS text message the day before the workshop or focus group to confirm attendance.

Rigor and Reflexivity

To maintain rigor within the study, credibility was ensured by reporting verbatim excerpts, tracking coding and category decisions, and confirming these through researcher triangulation

[26,27]. Although the study findings and the tailored intervention may not be generalizable to populations outside the study setting, providing a comprehensive description of the co-design process through detailed reports, thick descriptions, and analysis of contextual details, as described by Ponterotto [28], may allow for the transferability of the research method across contexts and medical conditions [29]. A comprehensive commentary reflecting on and cataloging the progress, obstacles, and successes of the research process increased dependability and confirmability by providing an audit trail for the study [29].

The primary researcher in this study (EAB) was an individual with an SCI, and a crucial step in their personal reflexivity [30] involved reflecting on how their position and perspective affected the study, in particular how being a member of the SCI community and having shared experiences but also common contacts allowed for the development of relationships with young people and caregivers. Developing these relationships at the beginning of the study and allowing time for non–research-focused conversations created a safe space in which young people could express their needs. Nevertheless, as the researcher sustained their SCI at the age of 22 years and did not use any pediatric services, the young people were the experts in this area.

Co-design Workshops (Phase 2a)

Methods

Overview

Two workshops were run on the web (owing to the COVID-19 pandemic) via videoconference (Zoom Video Communications): one for young people with SCIs and one for caregivers. Two researchers facilitated the workshops, one taking on the role of lead facilitator (EAB) and the second acting as cofacilitator and note-taker (LMR). Each workshop was run as a single group (all participants together all the time), with each participant invited to contribute their thoughts during each activity. Discussion prompters (eg, Ideaflip [Biggerflip Ltd] and Microsoft PowerPoint) were used in the workshops to organize material from discussions, enhance feedback, and guide intervention development. The workshops were recorded and transcribed with the participants' permission.

Preworkshop Preparation

A reference group was consulted to provide expert advice on the appropriateness of the co-design workshop activities, identify any issues or barriers that could impede the success of the workshop, and provide advice on how to resolve these issues or barriers. This reference group consisted of a young person with an SCI and 3 health care professionals (n=1, 33% clinical nurse consultants and n=2, 67% occupational therapists) from 3 different pediatric SCI health care service providers.

Approximately 1 week before the workshops, the facilitators met to clarify roles and review the workshop schedule and timing of activities. Participants were also sent an information booklet (Multimedia Appendices 2 and 3) and a link to the videoconference meeting by email.

Workshop Warm-ups

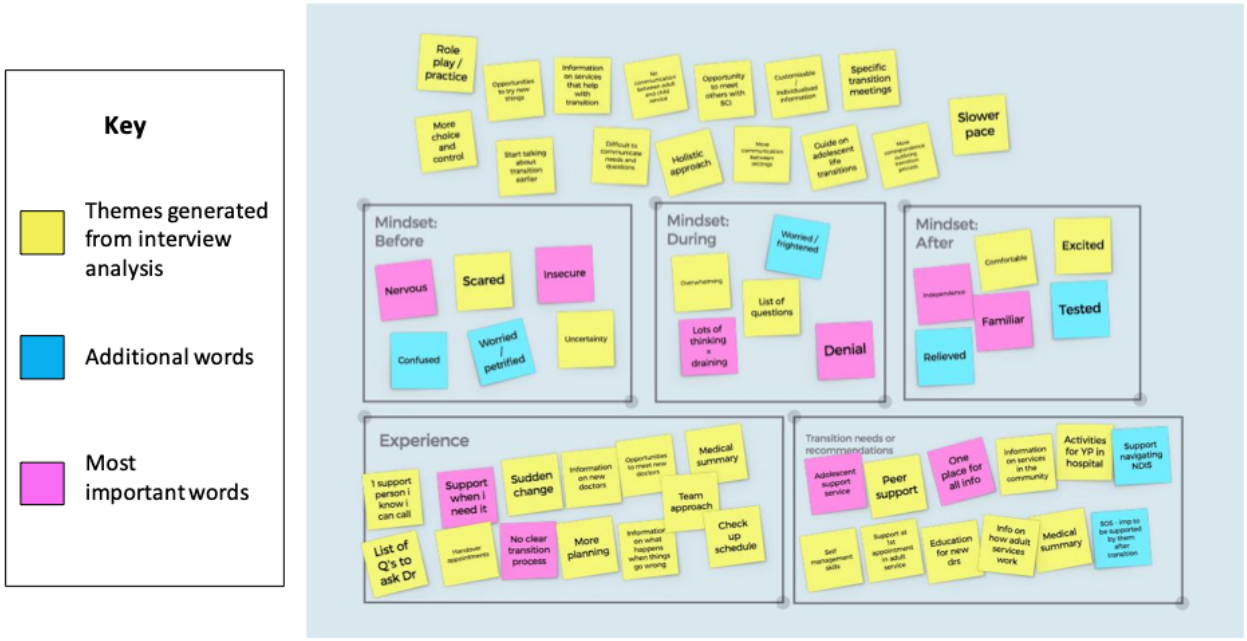
Each workshop began with a reminder of the aims of the study followed by a fun icebreaker activity (Multimedia Appendix 4) to help build rapport and ease any anxiety. Given that we were discussing issues relating to their personal experience transitioning between health care services, we asked that participants maintain confidentiality, especially if any sensitive issues were discussed. At the same time, each participant was encouraged to share as much or as little as they were comfortable with. This created an atmosphere that promoted open and honest communication and sharing. In the young persons' workshop, this information was reiterated at the start of the second activity (without the icebreaker) as some young people joined the workshop after the completion of the first activity.

Young Persons' Co-design Workshop

The young persons' workshop was 120 minutes long and consisted of 2 activities. The first activity required participants to review the thematic analysis of the semistructured interviews conducted by EAB and LMR as part of phase 1 of the research study. Participants were encouraged to reflect on some of the common ideas and concepts (codes) generated by the facilitators, confirm their authenticity, and compare and discuss analytical decisions as they grouped codes into themes. A web tool (Ideaflip) was used to create a whiteboard with Post-it notes to organize each idea or concept into predetermined boxes titled "Mindset: before," "Mindset: during," "Mindset: after," "Experience," and "Transition needs or recommendations" (Figure 1). These predetermined box labels aligned with the "Characteristics and Roles of Patients and Caregivers" domain of the Care Transitions Framework. Participants were also asked to add words or ideas that they believed had been omitted and highlight the most important word for them in each box. Participants' reflections on common ideas, concepts, and themes confirmed the thematic analysis conducted by EAB and LMR and have been published elsewhere (Bray et al, under review).

For the second activity, the researchers used the future workshop method [31,32] to facilitate discussion and generate ideas for the development of the HCT intervention. The future workshop method consists of 3 phases: a critique phase, a fantasy phase, and an implementation phase [31,32]. In the critique phase, the participants were asked to identify deficits or challenges related to the HCT experienced by young people with SCIs. The questions posed included the following: "What is the change you want to see? Or what did you need most to support your move but didn't have or receive?" In the fantasy phase, the participants were asked the following: "In a perfect world, how can this be achieved? Or what could we develop to support your move?" In the implementation phase, the participants transformed the "perfect world" ideas into a design for a practical and realizable HCT intervention ("How can we make this possible today?"). Participants had been provided with these 3 questions before the workshop in their information booklet. Using a Microsoft PowerPoint slide as a whiteboard, we discussed the participants' answers to the 3 posed questions and asked them to discuss their thought processes and ideas (Multimedia Appendix 5). The participants built on each other's ideas.

Figure 1. Screenshot of the first young peoples' co-design workshop activity.



Caregivers' Co-design Workshop

This 1-hour workshop mirrored activity 2 of the young persons' workshop (Multimedia Appendix 6). The questions posed to the caregivers included the following: (1) "What is the change you want to see? Or what did you need most to support your child's move but didn't have or receive?" (2) "In a perfect world how can this be achieved? Or what could we develop to support your child's move?" (3) "How can we make this possible today?"

Workshop Evaluation

At the end of the young persons' workshop, a short evaluation of the co-design workshop process took place. A web tool called Mentimeter (Mentimeter AB) was used to pose the following question: "Using 5 or more words, describe how you felt as a participant and what you thought of the workshop?"

Workshop Follow-up

Following the workshops, participants were provided with a link to a secure web-based Google document that included the questions posed in the workshops and a summary of the topics and ideas discussed. Participants were encouraged to cross-check information, add any information not covered during the workshops, and continue to develop the ideas generated.

Workshop Analysis

The audio recordings were transcribed and used in conjunction with the discussion prompters (Microsoft PowerPoint) to generate content themes for guiding the development of the HCT intervention. The researchers used a hybrid approach of deductive and inductive qualitative content analysis [26] of the transcripts, notes, and materials produced after each workshop. The reason this analysis method was chosen was that it allowed the researchers to organize and understand the data in a meaningful way and, through a manifest analysis, allowed the researchers to describe "*what* the informants actually say" by remaining close to the verbatim text [33]. The analysis was conducted in 3 stages: preparation, organization, and reporting [26]. In the first stage, the researchers immersed themselves in the data to obtain a sense of them as a whole. In the next stage, the researchers organized and condensed the data into meaning units through a process of open coding, categorization, and abstraction for the inductive approach and used the Care Transitions Framework to develop a categorization matrix for coding for the deductive approach (Multimedia Appendix 7). Finally, the contents of the categories and subcategories are described in detail as part of reporting the results.

Results

Participants

Young people joined and left the workshop at different times because of other commitments. In total, 2 participants, both female with tetraplegia aged 20 and 21 years, were present for the first activity of the young persons' workshop. A total of 3 participants, 2 (67%) female (n=1, 50% with tetraplegia and n=1, 50% with paraplegia) and 1 (33%) male with tetraplegia aged between 17 and 20 years were present for the second activity; 1 (33%) had attended the previous activity, and 2 (67%) new people joined for the second activity. In total, 4 participants, all mothers, attended the caregivers' workshop. A total of 75% (3/4) of the mothers had children with tetraplegia, and 25% (1/4) had a child with paraplegia. Only 1 young person and 1 caregiver had not yet transitioned. In total, 89% (8/9) of the original participants from phase 1 of the overall PAR study contributed to the co-design workshops.

Recommendations for the HCT Intervention

Overview

Data from the workshops were categorized in alignment with the Care Transitions Framework as summarized in Table 1.

Table 1. Workshop analysis content categories.

Care Transitions Framework domain and category	Subcategory
Intervention characteristics	
• What is the intervention designed to achieve?	• Coordinated handover between services: "For there to be more of a relationship built with the doctor before the transition occurs." • Greater independence: "I want to start moving that stuff to her, getting her to do things independently." • Peer connection: "Support group of the people going through the same thing as you."
• What are the features of the intervention?	• Information on the transition process and adult health care system: "Written information and summary on the transition process." • Medical summary and contact list: "Parallel lists of what was before and what was now. Like, pediatric versus adult." • One-stop shop for resources: "There should be pamphlets or a website with all the information." • Support to connect with others: "To be doing something together to help form relationships."
• Who is the intended target group?	—[a]

[a]No subcategory.

Category 1: What Is the Intervention Designed to Achieve?

Although caregivers wanted an intervention that supported their children in achieving greater independence, young people wanted the intervention to offer a space for peer connection. However, both caregivers and young people requested that the intervention support a coordinated handover between the pediatric and adult health care services and their multidisciplinary team members, including medical, nursing, rehabilitation, and allied health professionals.

1.1: Coordinated Handover Between Services: "For There to Be More of a Relationship Built With the Doctor Before the Transition Occurs"

Both young people and caregivers advocated for a smoother and more streamlined HCT. They believed that greater communication between the 2 health care settings (pediatric and adult) was essential for this to happen:

> *The adult service neurologist knew her paediatric neurologist and that is [a] huge help because they can chat between themselves, especially [because] we have the same neurologist for 15 years and he knew everything about Drew. It's so easy [for them] to communicate, rather than going through me.* [Morgan, caregiver]

In addition, participants wanted their pediatric health care team to organize the initial introductions of the family unit to the new adult health care team before transition:

> *I think just having an introduction as a family transitions across would be good.* [Jude, caregiver]

They felt this would reduce the need for young people and their caregivers to have to repeat their stories to different health care professionals:

> *Just a handover so people know where you are and you're not having to repeat yourself all the time.* [Kris, caregiver]

Furthermore, young people indicated a belief "that some doctors underestimate how big the transition actually is for young people" (Jamie, young person) and, as such, more support should be provided during the transition process. This would allow young people and their caregivers to build a relationship and rapport with their new health care team before moving and, in so doing, it would ease any anxiety around the transition:

> *For there to be more of a relationship built with the doctor before the transition occurs...so that the doctors are more aware of how the doctors in the children's hospital provide support and possibly adopt that, even though they're in the adult hospital.* [Jamie, young person]

1.2: Greater Independence: "I Want to Start Moving That Stuff to Her, Getting Her to Do Things Independently"

Caregivers acknowledged that, around the age of 16 years, young people legally could take charge of their own health care but that a lack of self-management skills at this age can result in unanticipated errors that can both be financially costly and have a detrimental impact on one's health:

> *The issue of being 16 [is that] Taylor was in charge, really on paper—he's in charge of his own health care, while he wanted us, as parents, to guide some of that. He wanted to make his own decisions and [has] every right to...Taylor is the voice of his own body. He's not always right, but it's 16 when he's making some of those choices and those choices are very expensive choices if you're buying a chair or you're paying for things...there's so many things that we've got wrong that have been a trial and error.* [Rory, caregiver]

Further to their lack of skills and knowledge, some caregivers identified that their child did not want to take charge of their own health care. However, caregivers saw the relinquishing of responsibility for health care decisions to the young person as an important milestone in the move to adult health care services:

> *She doesn't even want to entertain the thought of ordering products for herself or ringing up to get a new commode pusher. She wants me to do all that for her, I'm happy to, but I would like to see that, in the transition coming up in the next couple of years, I want to start moving that stuff to her, getting her to do things independently herself.* [Kris, caregiver]

Going forward in the young persons' move to adult health care services, caregivers saw that a requirement of the HCT intervention should be to build independence and the skills they need to manage their health care on their own:

> *So there's definitely something for me about...this kind of thing of letting young people become young people with that fierce independence, but to develop the skills that they actually see a chair as a piece of equipment that's necessary, that they see that they need that kind of life skills of being able to manage in situations.* [Rory, caregiver]

1.3: Peer Connection: "Support Group of the People Going Through the Same Thing as You"

All young people identified having an intervention that supported the opportunity to connect with others going through a similar experience as an important priority:

> *A support system where everyone going through the same experiences is able to kind of bond and start friendships.* [Jamie, young person]

Category 2: What Are the Features of the Intervention?

Young people and caregivers clearly outlined what they imagined an HCT intervention would need to include to achieve the desired outcomes outlined in the previous sections. These included information on the transition process and adult health care system, a medical summary including a contact list for health care professionals, accessible resources all located in one place, and a space to connect with others.

2.1: Information on the Transition Process and Adult Health Care System: "Written Information and Summary on the Transition Process"

Part of the struggle participants faced in their transition was a lack of knowledge of how and when the transition to adult health care would occur as well as how the adult health care system works:

> *It seems that some services change over or changed over when Jessie turned 16 and then others have stayed with paediatric services and so we've got this mix of some adult doctors ongoing with Jessie and some who are still in the paediatric system and it seems very messy for us. We're not sure whether we're in the adult system yet or in the children's system.* [Jude, caregiver]

> *I didn't have the knowledge. For example, in intensive care in paediatrics, we always had the same person and he was constant. Now, when we ended up in ICU for six weeks, every four days is a new doctor and that is so disturbing. Having 15 years [with] the same doctor and now, every four days, it was really hard to follow and it's not easy.* [Morgan, caregiver]

To address this knowledge gap, participants recommended having more information and education on the transition process and the adult health care system:

> *Written information and summary on the transition process.* [Drew, young person]

2.2: Medical Summary and Contact List: "Parallel Lists of What Was Before and What Was Now. Like, Paediatric Versus Adult."

Participants identified that a coordinated handover could be further supported with written documentation that is readily available to health care professionals, young people, and caregivers. A medical summary specifically created for the handover would mean that young people and their caregivers would not have to retell their stories:

> *A handover so people know where you are and you're not having to repeat yourself all the time. An updated information folio somewhere where you can access your information and where the doctors can access your information.* [Kris, caregiver]

Participants envisioned that the medical summary would include the young person's medical history, a schedule for annual appointments and scans, experiences they valued in the pediatric health care setting, and accomplishments:

> *The medical history of each person, what experiences from the paediatrics that they really enjoyed and would love to be integrated in the adult hospitals just to make it not as daunting when the transition occurs.* [Jamie, young person]

> *Getting a clear list of what are the ongoing check ins that Jessie needs for bladder, for bowel, for bones and how frequently we have to have those scans and tests done. I'm still trying to piece that together and it's incredibly messy and I keep thinking I'm going to miss him having an important scan or test done because I don't have a schedule that says "every two years, Jessie will need a bone scan. Every 12 months, he'll need a kidney test."* [Jude, caregiver]

Only 1 young person in the group described having a summary of their health care created for the handover process. However, despite it being given to the new health care team, Taylor felt that he had to repeat this information and supplement it. He also reported that he thought he probably had a copy of that summary but was not sure where it was now:

> *I think it had that, but I had to also repeat what could have been written down more than added to the summary.* [Taylor, young person]

In addition to the transition-specific medical summary, participants requested a contact list for their health care team that clearly displayed the pediatric health care professional and who would be taking over that role in the adult health care setting along with their contact details:

> *It would be nice to have the parallel lists of what was before and what was now. Like, paediatric versus*

> *adult...First [the] name of every condition, then name of every doctor or CNC [Clinical Nurse Consultant] and then equivalent or similar match in adult services and their contacts.* [Morgan, caregiver]

2.3: One-stop Shop of Resources: "There Should Be Pamphlets or a Website With All the Information"

What was evident from the discussions was that participants desired a "one-stop shop" or a repository where all HCT information was packaged in an easy-to-understand format and accessible:

> *There should be pamphlets or a website with all the information that people need to make it easily accessible in one place.* [Jamie, young person]
>
> *It's a package of education, health and everything else that's in our lives that can't be segregated.* [Rory, caregiver]

When asked about the information and resources they would like to be included in an HCT intervention, there were several requests (Table 2).

Young people also spoke about having the opportunity to share resources and collaborate, possibly using a web-based forum:

> *Just coming back to the opportunity to share the resources online, also possibly starting a forum within that in order to share—to help share those resources a bit more easily.* [Jamie, young person]

Finally, participants wanted information and resources to come from reputable sources such as physicians:

> *A list of resources provided by doctors possibly on the forum.* [Jamie, young person]

Table 2. Information and resources requested for inclusion in the health care transition intervention.

Information and resources requested	Participant quotes
Information on disability (general and SCI[a]-specific)	• "Things to do with disability and all stuff." [Ashley, young person]
Information on the difference between the pediatric and adult health care settings	• "How the adult system works and possibly differentiates from the children's system." [Jamie, young person] • "Someone to explain what's the difference between paediatric and adult." [Drew, young person]
Information on social activities	• "Information on sport. Like wheelchair sports and disabled sports." [Taylor, young person]
Alternate funding options	• "For people who don't get funding, like charities. Stuff like that to get equipment, wheelchairs, and stuff." [Taylor, young person]
Tips on building self-management skills	• "So managing all of these things is actually a skill which can be learnt either in some workshops or self-education...maybe you can request from a...Social Worker who might come and show some tips to your child and some mind mapping or Excel spreadsheets." [Morgan, caregiver]
Education and employment support	• "It's not just the transition of his care and things. It's that transition to what you do beyond school and how do you do that when you have a spinal cord injury?" [Jude, caregiver]

[a]SCI: spinal cord injury.

2.4: Support to Connect With Others: "To Be Doing Something Together to Help Form Relationships"

When asked to expand on how they would like to connect with others going through a similar experience, young people recommended having both one-on-one and group support options. They recommended that the one-on-one support be a formal program tailored to the young person's individual needs, matching them with someone who is either of a similar age or injury or who may have experience of transition, depending on what the young person desires:

> *Maybe the same age, but also similar injuries. So if it's someone who is able to walk, then someone who—I guess [a] support worker who may be able to walk.* [Taylor, young person]

The group support would be more informal and would involve monthly catch-ups of a more social nature (bowling, trivia, and games on the web):

> *Do an activity during that, instead of talking. To be doing something together to help form relationships.* [Jamie young person]

Category 3: Who Is the Intended Target Group?

The ultimate beneficiaries of the intervention are young people with SCIs. However, participants noted that they saw the intervention "as a package that everybody takes a role [in]" (Rory, caregiver). As such, the HCT intervention needed to support young people, their caregivers, and health care professionals.

Workshop Evaluation

A total of 75% (3/4) of the young people completed the evaluation at the end of the co-design workshop answering the following question: "Using 5 or more words, describe how you feel as a participant today and what you thought of the workshop?" The feedback was positive, and young people reported that they felt valued and listened to and that the

workshops were fun and informative (Multimedia Appendix 8).

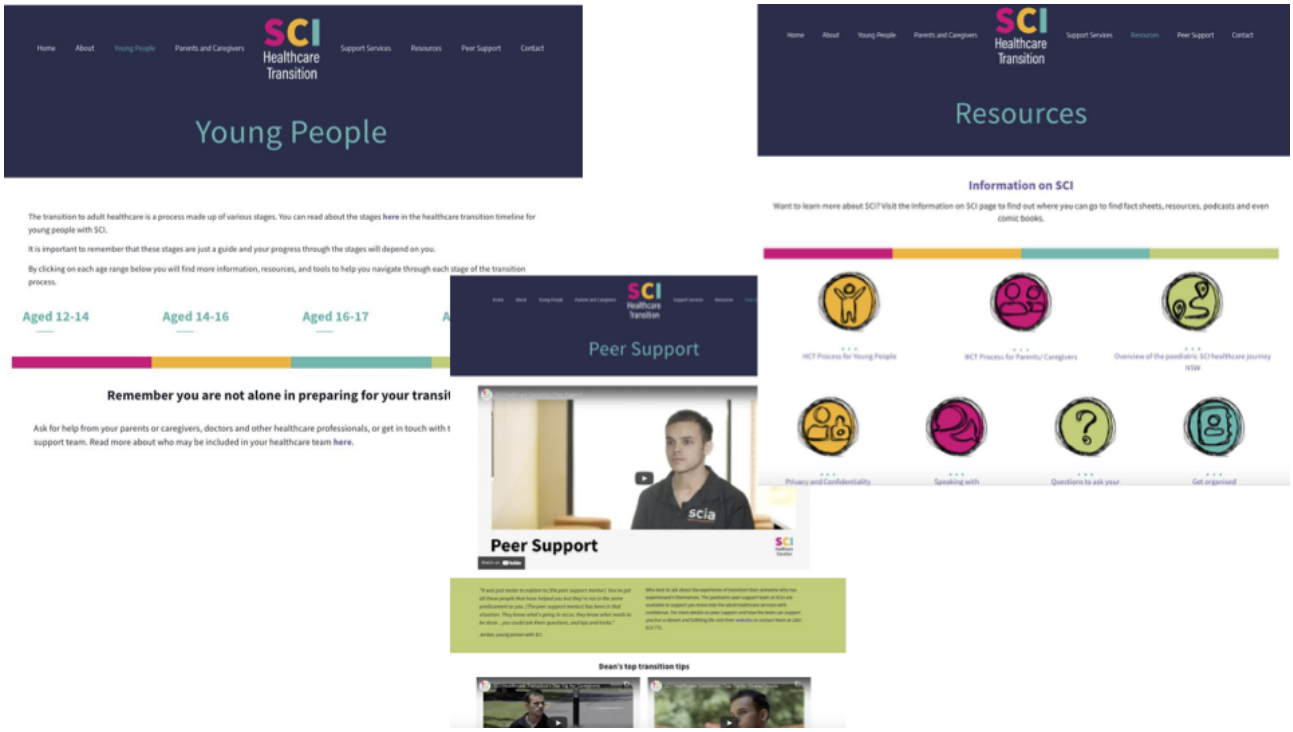
The Prototype HCT Intervention

On the basis of the recommendations from participants in the workshops, it was determined that a website would be an appropriate platform on which to deliver the suite of information needs. Content categories from the workshop were used to guide the development of the prototype HCT intervention, the SCI Healthcare Transition website [34] (Figure 2). The development of the website occurred over a period of 4 months and was supported by a web designer, motion graphics designer and video editor.

The website was designed with the aim of supporting young people with SCIs to achieve greater health care independence, support the smooth and coordinated handover from children's to adult health care services, and offer young people with SCIs a way to access peer support. It aimed to achieve this by providing young people with SCIs and caregivers with a step-by-step guide to HCT categorized by age that included tools (eg, a goal-planning worksheet), tips (eg, a PDF on tips for talking to health care professionals), and resources (eg, information on SCIs) to help prepare for the move. It also provided a directory on where to access further support from health care professionals and peer mentors. The website used a combination of eye-catching colors and graphics along with videos, interactive quizzes, and PDFs that are both downloadable and can be filled in, with the intention of appealing to young people (Figure 3).

Figure 2. Screenshot of the SCI Healthcare Transition website home page. SCI: spinal cord injury.

Figure 3. A total of 3 screenshots of the SCI Healthcare Transition website. SCI: spinal cord injury.



Focus Group Evaluation of the Prototype HCT Intervention Development (Phase 2b)

Methods

Overview

Similar to the process outlined for phase 2a, 2 focus groups were run on the web (owing to the COVID-19 pandemic) via videoconference (Zoom): one with young people with SCIs and caregivers and the other with the study's reference group of SCI health care professionals. The proposed HCT intervention was presented to the 2 groups, each of whom provided constructive feedback on the overall content and the layout and structure, allowing for further refinement of the intervention. The focus group was facilitated by 2 of the study's researchers, one taking on the role of lead facilitator (EAB) and the second acting as cofacilitator and note-taker (LMR). The focus groups were run as a single group (all participants together all the time). Similar to the workshops, discussion prompters (Microsoft PowerPoint) were used to guide the discussion and feedback. The focus groups were recorded and transcribed with the permission of the participants.

Co-design Process Evaluation

At the end of the focus group with young persons with SCIs and caregivers, a short evaluation of the co-design process took place. The researcher reflected on the co-design process with participants and gathered evaluative feedback.

Focus Group Analysis

As in phase 2a, the focus group transcriptions were used in conjunction with the discussion prompters (Microsoft PowerPoint) to develop content themes for guiding the refinement of the prototype HCT intervention.

Results

Participants

Of the 9 participants contacted from phase 1 of the overall PAR study, 4 (44%) provided feedback to the focus groups. In total, 3 participants attended the young people with SCIs and caregiver focus group: 1 (33%) young person (female, aged 21 years with tetraplegia) and 2 (67%) caregivers (both mothers of children with tetraplegia). A caregiver (mother of a child with paraplegia) was unable to attend the focus group but submitted written responses to the questions.

A total of 3 participants were invited, and all attended the SCI health care professionals' focus group. The health care workers were from 3 different pediatric SCI service providers, and each had different professions (clinical nurse consultant, occupational therapist, and physiotherapist).

Recommendations

Overview

Data from the focus groups were categorized in alignment with the Care Transitions Framework as summarized in Table 3.

Table 3. Focus group analysis content categories.

Care Transition Framework domain and category	Subcategory
Intervention characteristics	
• Does the website achieve what it was designed to achieve?	—[a]
• Website features	• Successful features: "That was...really awesome" • Recommendations for improvement: "One thing I thought would have been really useful is..."
Process of implementation	
• Website implementation	—

[a]No subcategory.

Category 4: Does the Website Achieve What It Was Designed to Achieve?

The aim of the website, as identified in the workshop analysis, was to support young people with SCIs to achieve greater health care independence, support the smooth and coordinated handover from children's to adult health care services, and offer young people with SCIs a way to access peer support. Participants were asked to keep these aims in mind when answering the first question: "Do the website's features achieve these outcomes? If not, why not?" All participants responded positively, and no recommendations for changes with regard to the overall outcomes were made:

> *I think in terms of helping to guide and support the health care transition from kids to adults, it [the website] did a pretty comprehensive job.* [Jude, caregiver]

Category 5: Website Features

Participants spoke about the features that they valued, however, had clear ideas on how the websites content could be further developed and refined.

5.1: Successful Features: "That Was...Really Awesome"

On the whole, participants spoke very highly of the website:

> *Your website is absolutely amazing.* [Morgan, caregiver]

Young people and caregivers valued the wealth of information and noted that the fact sheets, resources, and links to podcasts and organizations were "handy to have in one place" (Morgan, caregiver) and "will encourage patients and their families to develop a necessary knowledge base" (Morgan, caregiver). Health care professionals affirmed this sentiment:

> *I suppose it's kind of that link that I've always struggled with about trying to give more ownership to the young person of actually getting them to fill out bits themselves and then that kind of reveals any gaps in knowledge.* [Avery, health care professional]

All participants liked the simplicity of the quizzes and checklists, and their interest was piqued by the motivational videos at the start and end of the intervention:

> *I love the About page how it's got your video and the Peer Support one's got Dean...that was...really awesome...because straight away you're like, "I want to click on that." "I want to see what that's about" and it's just really kind of accessible and, teenagers, they kind of want to be fed information.* [Robin, health care professional]

More specifically, participants reported that the medical summary template and health care goal-planning worksheet were comprehensive and useful for everyone, even for those who had already transitioned. Furthermore, the transition checklist was "a nice way of reminding young people about their responsibilities and their self-awareness" (Morgan, caregiver).

Not only did the participants approve of the content, but they also appreciated the website's design scheme and simplicity. This was important for young people and caregivers as the transition between health care services is both stressful and overwhelming, and they often did not have the time or patience to navigate different websites to find information:

> *I just want to say how nice and colourful it is and that's also important for people. Just to make it more inviting and happier because, as you know, it might be quite overwhelming when you are going through all of that.* [Morgan, caregiver]

5.2: Recommendations for Improvement: "One Thing I Thought Would Have Been Really Useful Is..."

The participants had valuable recommendations for the further improvement of the website. They asked for more information on the National Disability Insurance Scheme (disability support funding), more information on the general life transitions of adolescence (study, employment, and living independently), and more education on SCIs (bladder and bowel management). In addition to the medical summary template already provided, young people and caregivers also wanted a schedule of health appointments template to record all necessary regular and ongoing health appointments; for example, bladder scans and bone density tests. Young people and caregivers could then get their health care team to check that they had not missed any important health checks:

> *One thing I thought would have been really useful is...once you've got your appointments calendar lined up, to just make sure that your rehab [rehabilitation] doctor looks over it to make sure you haven't*

> *forgotten to book in a bone scan or whatever else might be needed because it's really hard to remember all the bits and pieces that you need to check up on and I worry that there's something that we're missing because no-one's putting it together as a whole.* [Jude, caregiver]

Participants also requested to hear or read more success stories and expand the website in the future to include a forum or mobile app for further peer connection opportunities:

> *Just add testimonials from participants with the same situation as Drew or with SCI. So it's a good thing because you can actually build a community here. You can gain support. It's hard. It's hard to find support these days. So I think it's a good thing. And also for caregivers like me.* [Morgan, caregiver]

> *It would be awesome if, in a further extension of this, that they [young people] could somehow be connected, either on the website or in a social media group or an app [application] that they could opt into.* [Reese, health care professional]

Category 6: Website Implementation

With regard to the website implementation, the researchers wanted to gain an understanding of the usability of the website as well as how best to inform young people and caregivers about it.

Although the participants acknowledged that the website needed to be functional and practical for young people of varying abilities and those that used different technologies (eg, eye gaze tracking software), they believed the website was easy to use and generally accessible to young people with multiple levels of physical abilities because of its simplicity:

> *All of the pages were very simple. They weren't too overcrowded with information. So it was really easy and quick to flick through. Jessie also had a look. He was able to navigate around it quite easily and simply using his usual equipment on the computer.* [Jude, caregiver]

Participants appreciated the use of various formats of delivery, in particular the animation on the home page of the website, as they thought information delivered in this way was easier to digest than if a health care professional was speaking to them. In addition, participants reiterated that the simplicity of the website lent itself to being easy to use, and information was easy to comprehend for those that had limited time or a short attention span:

> *If anything is too complex, too busy, it just takes too much energy and it's too exhausting to try and navigate through it and sift through what you need to know.* [Morgan, caregiver]

Participants believed that there was a place for all members of the SCI community to be involved in the implementation of the website, from health care professionals in the hospital setting to those in the community as well as community-based SCI support organizations. Participants believed it to be the role of all stakeholders to promote the use of the website by linking to it from their own websites and advertising it in their newsletters and social media pages. Positively, the health care professionals who worked in these roles also envisaged themselves drawing on the website in their education sessions related to transition preparation:

> *I would actually...Show them the basics of it and go, "Go and have a look...This is all part of your preparation.*" [Avery, health care professional]

> *I would use this, absolutely, as a resource to actually talk our patients through the process of transition and it doesn't sound scary when it's coming from you guys and the way that you've presented that information.* [Avery, health care professional]

Co-design Process Evaluation

Reflecting on the co-design process and their involvement, young people with SCIs and caregivers appreciated being given the opportunity to participate in the codevelopment of an intervention to support the HCT of young people with SCIs and caregivers. It made them feel respected, valued, and heard and provided participants with a sense of achievement:

> *It made us feel really respected that someone took the time to ask us what was useful for us to have in this. So often people preach at you and tell you what they think you should know and it was nice as part of this process for you to pause and ask, what would be useful for us? What did we want to see in here?* [Jude, caregiver]

> *It feels engaging. Engaging and helpful.* [Drew, young person]

> *I have a sense of achievement, it's nice to have something you did to help other people.* [Morgan, caregiver]

Discussion

Principal Findings

Young people with SCIs and caregivers currently encounter obstacles in their HCT and report needing additional advice, information, and support to prepare for the move. This study has described the co-design and development of an HCT intervention to support the transition of young people with SCIs and caregivers from the pediatric to the adult health care setting. Information and advice to prepare for transition was purposefully presented in a manner that introduced young people to the transition process early and prompted them to learn more about their SCI and to start taking more responsibility for their own health care. Short videos informed young people about pediatric peer support services and offered "top transition tips" from a young adult who had experienced an HCT. The other tabs on the website provided links to support services, informational resources, and PDFs to support self-management skill development.

Although the purpose of this intervention was to provide support for young people with SCIs and their caregivers during their transition from pediatric to adult health care services, it must be acknowledged that this transition occurs during a broader

transition process—the transition to adulthood. The transition to adulthood and its implications have been previously discussed [9,35] and, as such, this was not the focus of this study. However, in the feedback focus groups, participants reported the need for more information on the general transitions, including study, employment, and independent living. Consequently, information on where to access support on topics such as sexuality, education, and employment was added to the website, although its focus remained on supporting the move from pediatric to adult health care services.

The participatory co-design approach used in this study supported the active engagement of young people with SCIs and caregivers in the design process and resulted in the development of an intervention that addressed the current gaps in the HCT process as identified by end users. These findings support the observations of others [36,37]. For example, a study from Ireland by Coyne et al [37] reported on the co-design of a website to support the transition of young people with long-term illnesses to adult health care services. Their study highlighted that a participatory co-design approach yielded a reliable, functional, and acceptable intervention to support young people in their transition to adult health care [37]. Beaudry et al [36] similarly described a participatory co-design approach in the development of a chatbot that aimed to promote the attainment of self-care skills during the transition to adult care. They also reported that the resulting intervention was feasible for supporting engagement during HCT. Furthermore, the involvement of health care professionals in the feedback focus groups in our study ensured that we gained a broader scope for the design of the intervention, ensuring that it not only fulfilled the needs of young people with SCIs but also complemented current services.

Including end users in disability research brings knowledge and experience that may not be held by the researchers themselves and that can add to the diversity of skills and knowledge required for more appropriately designed research [38]. Furthermore, PAR and co-design principles foster empowerment as people with disabilities gain control over their lives and make decisions on matters that affect them [14]. Neither Coyne et al [37] nor Beaudry et al [36] evaluated young peoples' experiences of being involved in the co-design process; however, our study did. Evaluation data from young people on their involvement in the co-design process highlighted that their inclusion empowered them, gave them a voice, and provided them with an opportunity to contribute to an intervention that would make a difference in their lives and the lives of others. This provides further evidence of the importance of giving young people with disabilities the opportunity to authentically and meaningfully participate in the research and codevelopment of interventions that affect their lives.

Strengths and Limitations

Evaluation of the co-design process indicated that the participants valued the opportunity to be part of the development of a solution and appreciated being given a voice. However, the iterative and cyclical nature of the co-design process did present some challenges. Recruitment for the study and maintenance of engagement was a challenge across the different phases of the study. Of the 9 young people and caregivers who participated in the interviews during phase 1 of the study, 8 (89%) returned to participate in the co-design workshops, and 4 (44%) participated in the focus groups in phase 2. Reasons for the dropout included family stressors, relocation to another country, and nonresponse to phone or email. Owing to a paucity of participant numbers, we did not achieve a representative research sample, with no LGBTQ+ and no cultural and linguistically diverse representation (including Aboriginal and Torres Strait Islander input). A further limitation of our study relates to the inability to compare our findings with other similar studies because of a paucity of written literature on this topic [16].

Owing to the COVID-19 pandemic, activities initially intended to be held in person were moved to the web. This modification had its advantages as it eliminated geographical and mobility barriers to participation and fostered inclusive research practices. However, because of the additional pressures of COVID-19 and homeschooling on families, it was difficult to find times that were suitable to all, and it required a substantial amount of preparatory work on the part of the principal researcher (EAB) to organize workshops and focus groups.

The next phase of the PAR study is to assess the acceptability and feasibility of the HCT intervention. We plan to roll out the website with the same participants and conduct short evaluation telephone interviews based on the 8 focus areas in the framework by Bowen et al [39].

Conclusions

Engaging young people with SCIs and caregivers in the co-design of an HCT intervention has produced, in a relatively short time frame, a great depth of insight into the transition needs of young people with SCIs and their priorities for support. The result has been the collaborative development of an intervention that young people with SCIs, caregivers, and health care professionals believe will support the transition from child to adult health care services and equip young people with SCIs with practical and helpful tools to take charge of their health care. This is the first study to co-design and develop an HCT intervention in partnership with young people with SCIs and caregivers. Although the study sample was small, it has shown that it is possible to meaningfully engage young people with SCIs and caregivers in the co-design of an HCT intervention that leads to enhanced end-user acceptability.

Acknowledgments

The authors thank all participants for their valuable feedback, creativity, and time. The SpineCare Foundation, Sydney, New South Wales, Australia, funded this project (grant number: not applicable). The funders had no involvement in the review or approval of the manuscript for publication.

Authors' Contributions

All authors were responsible for the study conception and design. EAB and LMR acquired the data, completed the data analysis and verification, and drafted the manuscript. All authors were responsible for reviewing and finalizing the manuscript, provided approval for the final version of the report, and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Project phase and objective alignment with the Care Transitions Framework domains. It displays how the phases of the project and study objectives align with the Care Transitions Framework domains.
[PDF File (Adobe PDF File), 191 KB - formative_v6i7e38616_app1.pdf]

Multimedia Appendix 2
Young person co-design workshop participant booklet (activity booklet sent to young people in preparation for the co-design workshop).
[PDF File (Adobe PDF File), 203 KB - formative_v6i7e38616_app2.pdf]

Multimedia Appendix 3
Caregiver co-design workshop participant booklet (activity booklet sent to young people in preparation for the co-design workshop).
[PDF File (Adobe PDF File), 175 KB - formative_v6i7e38616_app3.pdf]

Multimedia Appendix 4
Screenshots of the icebreaker used across both co-design workshops. This image presents 2 screenshots of the randomized question spin wheel used as an icebreaker in the co-design workshops.
[PDF File (Adobe PDF File), 100 KB - formative_v6i7e38616_app4.pdf]

Multimedia Appendix 5
Screenshots of the 3-stage process from activity 2 of the young person workshop. This figure shows young people's responses to the 3 questions posed in the co-design activity.
[PDF File (Adobe PDF File), 380 KB - formative_v6i7e38616_app5.pdf]

Multimedia Appendix 6
Screenshots of the 3-stage process of the caregiver workshop. This image shows caregivers' responses to the 3 questions posed in the co-design activity.
[PDF File (Adobe PDF File), 468 KB - formative_v6i7e38616_app6.pdf]

Multimedia Appendix 7
Examples of meaning units, condensed meaning units, codes, subcategories, and categories (an example of the how the data were condensed into meaning units, coded, and categorized).
[PDF File (Adobe PDF File), 100 KB - formative_v6i7e38616_app7.pdf]

Multimedia Appendix 8
Young person workshop evaluation (this image displays young people's responses to the workshop evaluation question).
[PDF File (Adobe PDF File), 645 KB - formative_v6i7e38616_app8.pdf]

References

1. Mulcahey MJ, Vogel LC, Sheikh M, Arango-Lasprilla JC, Augutis M, Garner E, et al. Recommendations for the National Institute for Neurologic Disorders and Stroke spinal cord injury common data elements for children and youth with SCI. Spinal Cord 2017 Apr;55(4):331-340. [doi: 10.1038/sc.2016.139] [Medline: 27845358]
2. Schmidt A, Ilango SM, McManus MA, Rogers KK, White PH. Outcomes of pediatric to adult health care transition interventions: an updated systematic review. J Pediatr Nurs 2020;51:92-107. [doi: 10.1016/j.pedn.2020.01.002] [Medline: 31981969]
3. Key Principles for Transition of Young People from Paediatric to Adult Health Care. Agency for Clinical Innovation and Trapeze, The Sydney Children's Hospitals Network. 2014. URL: https://aci.health.nsw.gov.au/__data/assets/pdf_file/0011/251696/Key_Principles_for_Transition.pdf [accessed 2022-05-30]

4. Blum RW, Garell D, Hodgman CH, Jorissen TW, Okinow NA, Orr DP, et al. Transition from child-centered to adult health-care systems for adolescents with chronic conditions. A position paper of the Society for Adolescent Medicine. J Adolesc Health 1993 Nov;14(7):570-576. [doi: 10.1016/1054-139x(93)90143-d] [Medline: 8312295]
5. Stewart D. Transition to adult services for young people with disabilities: current evidence to guide future research. Dev Med Child Neurol 2009 Oct;51 Suppl 4:169-173 [FREE Full text] [doi: 10.1111/j.1469-8749.2009.03419.x] [Medline: 19740226]
6. Gray WN, Schaefer MR, Resmini-Rawlinson A, Wagoner ST. Barriers to transition from pediatric to adult care: a systematic review. J Pediatr Psychol 2018 Jun 01;43(5):488-502. [doi: 10.1093/jpepsy/jsx142] [Medline: 29190360]
7. White PH, Cooley WC, Transitions Clinical Report Authoring Group, American Academy of Pediatrics, American Academy of Family Physicians, American College of Physicians. Supporting the health care transition from adolescence to adulthood in the medical home. Pediatrics 2018 Nov;142(5):e20183610. [doi: 10.1542/peds.2018-2587] [Medline: 30348754]
8. Acuña Mora M, Saarijärvi M, Moons P, Sparud-Lundin C, Bratt EL, Goossens E. The scope of research on transfer and transition in young persons with chronic conditions. J Adolesc Health 2019 Nov;65(5):581-589. [doi: 10.1016/j.jadohealth.2019.07.014] [Medline: 31540780]
9. Porto A, Anderson L, Kalinich T, Deane KC, Vogel LC, Zebracki K. Understanding transition for youth with spinal cord injury: youth and caregiver perceptions. J Spinal Cord Med 2020 Jul;43(4):505-511 [FREE Full text] [doi: 10.1080/10790268.2019.1574437] [Medline: 30758272]
10. Baker I, de Paula A, Serratore L, Hanna M, Diviney K, Clark N, et al. Towards independence: the New South Wales (Australia) experience of transition to adulthood of young people with spinal cord injury. Top Spinal Cord Inj Rehabil 2010 Jul;16(1):55-65. [doi: 10.1310/sci1601-55]
11. Steen M, Manschot M, De Koning N. Benefits of co-design in service design projects. Int J Des 2011;5(2):53-60.
12. Thabrew H, Fleming T, Hetrick S, Merry S. Co-design of eHealth interventions with children and young people. Front Psychiatry 2018 Oct 18;9:481 [FREE Full text] [doi: 10.3389/fpsyt.2018.00481] [Medline: 30405450]
13. 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]
14. Balcazar F, Keys C, Kaplan MA, Suarez-Balcazar Y. Participatory action research and people with disabilities: principles and challenges. Can J Rehabil 1998;12:105-112.
15. Kelly PJ. Practical suggestions for community interventions using participatory action research. Public Health Nurs 2005;22(1):65-73. [doi: 10.1111/j.0737-1209.2005.22110.x] [Medline: 15670327]
16. Bray EA, Everett B, George A, Salamonson Y, Ramjan LM. Co-designed healthcare transition interventions for adolescents and young adults with chronic conditions: a scoping review. Disabil Rehabil 2021 Oct 01:1-22. [doi: 10.1080/09638288.2021.1979667] [Medline: 34595986]
17. Evidence and utilisation of spinal cord injury services in NSW: Evidence Series. Agency for Clinical Innovation and Government of New South Wales. 2020 Aug. URL: https://aci.health.nsw.gov.au/__data/assets/pdf_file/0008/595673/ACI-Evidence-and-utilisation-of-SCI-NSW.pdf [accessed 2022-05-30]
18. Bray EA, George A, Everett B, Salamonson Y, Ramjan L. Protocol for developing a healthcare transition intervention for young people with spinal cord injuries using a participatory action research approach. BMJ Open 2021 Jul 29;11(7):e053212 [FREE Full text] [doi: 10.1136/bmjopen-2021-053212] [Medline: 34326059]
19. Rojas SL, Ashok M, Morss Dy S, Wines RC, Teixeira-Poit S. Contextual Frameworks for Research on the Implementation of Complex System Interventions. Rockville, MD, USA: Agency for Healthcare Research and Quality (US); Mar 2014.
20. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
21. New PW, Baxter D, Farry A, Noonan VK. Estimating the incidence and prevalence of traumatic spinal cord injury in Australia. Arch Phys Med Rehabil 2015 Jan;96(1):76-83. [doi: 10.1016/j.apmr.2014.08.013] [Medline: 25218255]
22. New PW, Farry A, Baxter D, Noonan VK. Prevalence of non-traumatic spinal cord injury in Victoria, Australia. Spinal Cord 2013 Feb;51(2):99-102. [doi: 10.1038/sc.2012.61] [Medline: 22665222]
23. Vogel LC, Betz RR, Mulcahey MJ, Zebracki K. Spinal cord injuries and disorders in children and adolescents. In: Kirshblum S, Lin VW, editors. Spinal Cord Medicine. 3rd edition. New York, NY, USA: Demos Medical; 2017:926-954.
24. Choi EK, Bae E, Jang M. Transition programs for adolescents and young adults with spina bifida: a mixed-methods systematic review. J Adv Nurs 2021 Feb;77(2):608-621. [doi: 10.1111/jan.14651] [Medline: 33222278]
25. Holmbeck GN, Kritikos TK, Stern A, Ridosh M, Friedman CV. The transition to adult health care in youth with spina bifida: theory, measurement, and interventions. J Nurs Scholarsh 2021 Mar;53(2):198-207. [doi: 10.1111/jnu.12626] [Medline: 33482054]
26. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
27. Erlingsson C, Brysiewicz P. A hands-on guide to doing content analysis. Afr J Emerg Med 2017 Sep;7(3):93-99 [FREE Full text] [doi: 10.1016/j.afjem.2017.08.001] [Medline: 30456117]

28. Ponterotto JG. Brief note on the origins, evolution, and meaning of the qualitative research concept thick description. Qual Rep 2006 Jan 16;11(3):538-549. [doi: 10.46743/2160-3715/2006.1666]
29. Shenton AK. Strategies for ensuring trustworthiness in qualitative research projects. Educ Inf 2004 Jul 19;22(2):63-75. [doi: 10.3233/efi-2004-22201]
30. Olmos-Vega FM, Stalmeijer RE, Varpio L, Kahlke R. A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. Med Teach 2022 Apr 07:1-11. [doi: 10.1080/0142159X.2022.2057287] [Medline: 35389310]
31. Alminde S, Warming H. Future workshops as a means to democratic, inclusive and empowering research with children, young people and others. Qual Res 2019 Jul 27;20(4):432-448. [doi: 10.1177/1468794119863165]
32. Müllert N, Jungk R. Future Workshops: How to Create Desirable Futures. London, UK: Institute for Social Inventions; 1987.
33. Bengtsson M. How to plan and perform a qualitative study using content analysis. NursingPlus Open 2016;2:8-14. [doi: 10.1016/j.npls.2016.01.001]
34. Bray EA. SCI Healthcare Transition. 2022. URL: https://scihealthcaretransition.com/ [accessed 2022-05-30]
35. Zebracki K, Anderson C, Chlan K, Vogel L. Outcomes of adults with pediatric-onset spinal cord injury: longitudinal findings and implications on transition to adulthood. Top Spinal Cord Inj Rehabil 2010 Jul;16(1):17-25. [doi: 10.1310/sci1601-17]
36. Beaudry J, Consigli A, Clark C, Robinson KJ. Getting ready for adult healthcare: designing a chatbot to coach adolescents with special health needs through the transitions of care. J Pediatr Nurs 2019;49:85-91. [doi: 10.1016/j.pedn.2019.09.004] [Medline: 31644960]
37. Coyne I, Prizeman G, Sheehan A, Malone H, While AE. An e-health intervention to support the transition of young people with long-term illnesses to adult healthcare services: design and early use. Patient Educ Couns 2016 Sep;99(9):1496-1504. [doi: 10.1016/j.pec.2016.06.005] [Medline: 27372524]
38. Joss N, Cooklin A, Oldenburg B. A scoping review of end user involvement in disability research. Disabil Health J 2016 Apr;9(2):189-196. [doi: 10.1016/j.dhjo.2015.10.001] [Medline: 26596694]
39. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]

Abbreviations

HCT: health care transition
NSW: New South Wales
PAR: participatory action research
SCI: spinal cord injury

Edited by A Mavragani; submitted 09.04.22; peer-reviewed by M Mostafa, C Kiekens; comments to author 04.05.22; revised version received 17.05.22; accepted 21.06.22; published 28.07.22.

Please cite as:
Bray EA, Everett B, George A, Salamonson Y, Ramjan LM
Developing a Health Care Transition Intervention With Young People With Spinal Cord Injuries: Co-design Approach
JMIR Form Res 2022;6(7):e38616
URL: https://formative.jmir.org/2022/7/e38616
doi:10.2196/38616
PMID:35900814

Original Paper

Effects of a Gamified, Behavior Change Technique–Based Mobile App on Increasing Physical Activity and Reducing Anxiety in Adults With Autism Spectrum Disorder: Feasibility Randomized Controlled Trial

Daehyoung Lee[1], PhD; Georgia C Frey[2], PhD; Donetta J Cothran[2], PhD; Jaroslaw Harezlak[3], PhD; Patrick C Shih[4], PhD

[1]Department of Applied Human Sciences, University of Minnesota Duluth, Duluth, MN, United States

[2]Department of Kinesiology, Indiana University, Bloomington, IN, United States

[3]Department of Epidemiology and Biostatistics, Indiana University, Bloomington, IN, United States

[4]Department of Informatics, Indiana University, Bloomington, IN, United States

Corresponding Author:
Daehyoung Lee, PhD
Department of Applied Human Sciences
University of Minnesota Duluth
123 Sports and Health Center
1216 Ordean Court
Duluth, MN, 55812
United States
Phone: 1 2187267816
Email: lee03284@d.umn.edu

Abstract

Background: Physical activity (PA) has an impact on physical and mental health in neurotypical populations, and addressing these variables may improve the prevalent burden of anxiety in adults with autism spectrum disorder (ASD). Gamified mobile apps using behavior change techniques present a promising way of increasing PA and reducing sedentary time, thus reducing anxiety in adults with ASD.

Objective: This study aimed to compare the effectiveness of a gamified and behavior change technique–based mobile app, PuzzleWalk, versus a commercially available app, Google Fit, on increasing PA and reducing sedentary time as an adjunct anxiety treatment for this population.

Methods: A total of 24 adults with ASD were assigned to either the PuzzleWalk or Google Fit group for 5 weeks using a covariate-adaptive randomization design. PA and anxiety were assessed over 7 days at 3 different data collection periods (ie, baseline, intervention start, and intervention end) using triaxial accelerometers and the Beck Anxiety Inventory. Group differences in outcome variables were assessed using repeated-measures analysis of covariance, adjusting for age, sex, and BMI.

Results: The findings indicated that the PuzzleWalk group spent a significantly longer amount of time on app use compared with the Google Fit group ($F_{2,38}$=5.07; P=.01; partial η^2=0.21), whereas anxiety was unfavorably associated with increases in light PA and decreases in sedentary time after intervention (all P<.05).

Conclusions: Further research is needed to clarify the determinants of physical and mental health and their interrelationship in adults with ASD to identify the factors that facilitate the use and adoption of mobile health technologies in these individuals. Despite these mixed results, the small changes in PA or anxiety may be clinically significant for adults with ASD.

Trial Registration: ClinicalTrials.gov NCT05466617; https://clinicaltrials.gov/show/NCT05466617

(JMIR Form Res 2022;6(7):e35701) doi:10.2196/35701

KEYWORDS

gamification; behavior change techniques; physical activity; sedentary behavior; anxiety; autism; mobile app; mental health; mHealth; mobile phone

Introduction

Regular physical activity (PA) helps to reduce anxiety in the neurotypical population [1,2] and presents a potentially effective adjunct treatment for anxiety in people with autism spectrum disorder (ASD). Anxiety is one of the most common and debilitating mental health issues among adults with ASD [3,4]. A recent systematic review and meta-analysis study revealed that the lifetime prevalence of anxiety was >40% in a large sample of adults with ASD (n=26,070) included in the study [4]. The negative impact of chronic anxiety on those with ASD has been well documented, and evidence suggests that the presence of comorbid psychiatric disorders significantly interferes with active daily living and further increases the risk of clinical morbidity in these individuals [5,6].

Despite the prevalence and negative impact of anxiety on the everyday lives of those with ASD [7], there are few effective treatment options for this symptom. Most people with ASD rely on medications, which carry a high dependency risk [8,9], or cognitive behavioral therapy, which has demonstrated mixed results in alleviating this symptom [10,11]. In addition, antianxiety medication use is associated with several side effects, such as metabolic syndrome and weight gain [12,13], whereas cognitive behavioral therapy is both costly and labor intensive [14,15]. As a result, efforts have been made to identify adjunct treatments to help adults with ASD better manage anxiety [16,17]. The adult population with ASD is increasing [18,19], and there is a critical need for accessible and cost-effective adjunct anxiety treatments to alleviate this problem.

Although the level of benefits may vary from person to person, participation in regular PA is proven to help alleviate anxiety symptoms by improving self-efficacy [20], strengthening the sympathetic nervous system [21], and increasing neurogenesis in the human brain [22]. Research on PA interventions for adults with ASD is markedly lacking; however, there is support for the positive impact of PA or exercise on mental health in this population. Hillier et al [23] found that salivary cortisol and self-reported anxiety were reduced immediately after a low-intensity exercise session 1 day per week in a small sample of young adults with ASD. These effects were not maintained over the course of the 8-week intervention; however, the findings suggest that even low-intensity exercise can have a positive, albeit short-term, effect on anxiety in these individuals [23]. It is of interest to determine whether participation in regular PA can induce similar results in adults with ASD in the long term.

Although it is well understood that PA is a leading health indicator for both physical and mental health [24-26], there is little information on PA participation in adults with ASD. To date, only a few studies have addressed PA levels in adults with ASD, and the results have varied greatly. Eaves and Ho [27] reported that adults with ASD only engaged in moderate to vigorous PA (MVPA) once per week and spent approximately 13 hours per day sitting. In contrast, Frey et al [28] and Garcia-Pastor et al [29] found that adults with ASD met the guidelines of 150 minutes of weekly MVPA but were also highly sedentary. Lalonde et al [30] also observed that adults with ASD had daily step counts similar to neurotypical adults. It is difficult to draw clear conclusions about these findings because of variability in the assessment methods and the functional ability of the samples. Most of the previous study findings were based on participants who attended segregated schools or centers and required extensive support, which was not representative of adults with ASD who were more intellectually able and autonomous.

To date, a limited number of studies have specifically attempted to increase PA in adults with ASD by using objective measures. Lalonde et al [30] used a goal-setting–based and reinforcement-based treatment to increase pedometer-measured walking steps in 5 young adults with ASD who attended a special education program. All participants were able to increase the number of walking steps to sufficient levels to achieve health benefits (ie, ≥10,000 steps). Although these are encouraging findings, the highly structured and prompted nature of the intervention makes it inappropriate to generalize the findings to adults with ASD who are more intellectually able and autonomous. There has also been an effort to address physical fitness in individuals with ASD using a video modeling–incorporated mobile app. Although this novel intervention approach is promising, as the app was effective in increasing the heart rate and energy expenditure of participants with ASD, the findings were limited to children with ASD who required substantial assistance and supervision for proper execution [31]. It is important to identify PA and sedentary time interventions that are more suitable for independent adults with ASD who are self-determined for health behaviors.

Technology-based PA interventions have proven to be a promising method of increasing PA in the neurotypical population [32-34] and may be well suited for adults with ASD. Gamified mobile technology is a particularly promising tool that can increase PA and reduce sedentary time while also meeting the unique needs and interests of these individuals [35-37]. Lee et al [37] found that adults with ASD used technological devices for >6 hours per day, primarily for surfing the internet, using social media, and web-based gaming [37]. Emerging evidence suggests that adults with ASD are attracted to technology use because the human-technology interface creates consistency and predictability, as well as because of a lower social burden, compared with traditional face-to-face interaction [38,39]. Furthermore, individuals with ASD typically have distinctive strengths in visuospatial functioning, which is common to technology-based games, as well as a preference for learning and interaction through visual information [40-42]. Consequently, in the past few decades, health practitioners have actively used mobile technologies to offer a cost-effective and low-barrier platform for visual learning to identify and improve health outcomes in diverse clinical populations, including those with ASD [37,38,43].

Gamified behavioral interventions using smartphone apps have the advantage of providing personalization, feelings of amusement, and desire for continuation [44-46] and have rapidly expanded their technological potential to monitor and improve daily PA participation in adults with obesity and sedentary workers [35,47]. Nevertheless, the success of gamified mobile apps in promoting PA and reducing sedentary behavior is

questionable as most of the existing health or fitness apps in the commercial market are not sustainable stand-alone interventions and lack scientific evidence and health behavior theory in the app development process [48,49]. The overall low quality of evidence regarding the effectiveness of long-term behavior change makes it currently difficult to apply commercial PA-promoting apps to adults with ASD [50]. The purpose of this study was to (1) examine the effects of the competitive gamification and behavior change theory–based mobile app PuzzleWalk [37] on increasing PA and reducing sedentary time and anxiety in adults with ASD and (2) compare PuzzleWalk to a commercially available platform, Google Fit. It is hypothesized that (1) the use of PuzzleWalk will lead to higher levels of light PA and MVPA and lower levels of sedentary time and anxiety in adults with ASD than the use of Google Fit and (2) the increased PA or decreased sedentary time from both apps will be associated with reduced levels of anxiety in adults with ASD.

Methods

Participants

A total of 29 adults aged ≥18 years and diagnosed with ASD were recruited through state and regional agencies that serve people with ASD across the United States and online autism support groups on social media such as Facebook and Reddit. Evidence of ASD diagnosed by a qualified medical professional such as a pediatrician or clinical psychologist (ie, when and where) was required for study participation and obtained via self-report. In addition, eligible study participants met the following inclusion criteria: (1) self-reported medical diagnosis of anxiety or self-identification of experiencing anxiety symptoms for the past 3 or more months, (2) access to a supported device (smartphones with Android 4.4 and higher or iOS 9.0 and higher operating system), (3) cognitive ability to understand the purpose of the study, and (4) no prior experience using the PA mobile apps used in the study. Individuals with low cognitive function, co-occurring intellectual disabilities, or mobility impairments were excluded from this study. We hereby use the term *intellectually able* adults with ASD to refer to those who can make their own decisions regarding health behaviors without much assistance. A participant with a self-reported mild learning disorder that did not significantly interfere with active daily living and study participation was included in the study. A formal screening interview was conducted with each participant through a phone or face-to-face video call to verify participant eligibility and identify potential barriers to study participation. All participants provided written or digital consent before data collection.

Ethics Approval

The Institutional Review Board of Indiana University approved this study (protocol number 1807483245).

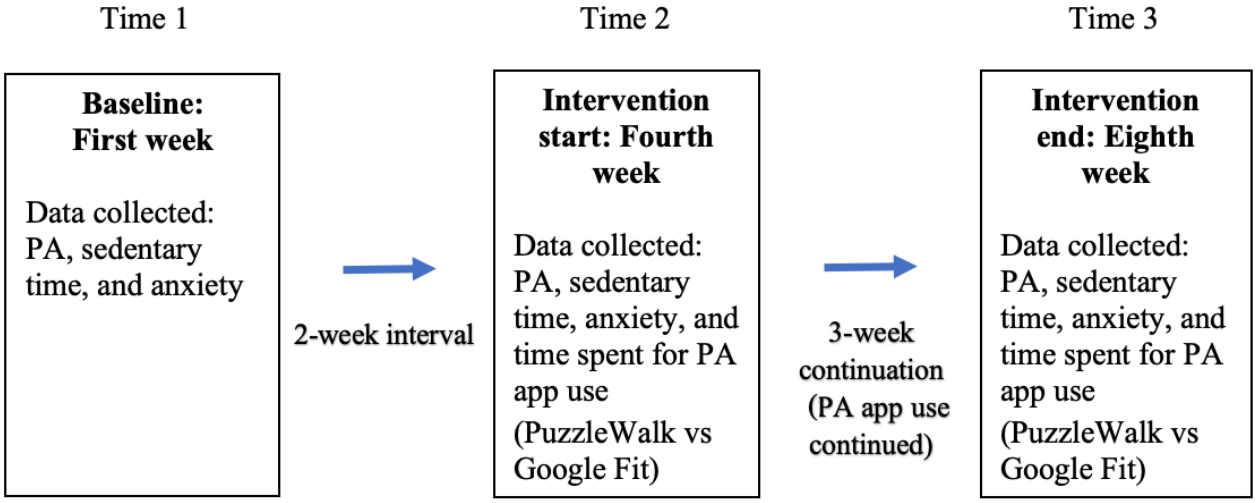
Procedure

On verification of the participant's eligibility through the screening interview, participants completed a web-based demographic survey with an emphasis on BMI, waist circumference, medication use, and autism symptoms. Self-reported height and weight information were collected to calculate each participant's BMI (ie, kg/m^2 or lbs/in^2 × 703) [51]. Participants were also asked to provide self-measured waist circumference, which can be an indicator of obesity-related disease risk [52]. The Autism Spectrum Quotient 10-item (AQ-10) questionnaire was included in the survey to examine the severity of autism symptoms (eg, social interaction deficits, sensory issues, and other autism-specific characteristics) [53]. The AQ-10 was designed for health care professionals to administer an informal ASD assessment of individuals with typical cognitive function [54]. The maximum possible score is 10, and individuals with a score of ≥6 are advised to consider referral for formal ASD assessment. The AQ-10 is one of the few autism screening tools available for adults [18]. After the completion of the demographic survey, study materials, including an accelerometer with an elastic belt, a USB cable charger, and study instruction sheets (ie, how to wear and charge the accelerometer and how to install and use the daily anxiety assessment [DAA] app for anxiety assessment) were either mailed or handed to remote and local participants, respectively.

A rubric was used to assess participant comfort and knowledge of study procedures before the start of the baseline and intervention periods, and case-by-case decisions were made by the research team when individuals were ready to start data collection. Before the start of the intervention period, participants were assigned to either the PuzzleWalk or Google Fit group according to age, sex, and BMI using a covariate-adaptive randomization process. Specifically, a minimization technique was applied to the randomization process by distributing the participants into 2 groups based on the aforementioned variables, which were identified before the start of data collection [55]. Covariate randomization aimed to minimize the imbalance in baseline characteristics across the 2 groups included in the study [56]. All participants received visualized step-by-step instructions (eg, search and download on Google Play or App Store, user registration, goal setting, and PA behavior tracking) on the assigned PA app (PuzzleWalk or Google Fit) and used it from the beginning of the intervention start (fourth week) until the end of the intervention (eighth week).

Participation required an approximate 2-month commitment: the first week for baseline and the fourth to eighth weeks for intervention. A 2-week interval was implemented between baseline and intervention onset to reduce the novelty effect in response to the accelerometer and PA app use [57] (see Figure 1 for the study timeline). Both the PuzzleWalk and Google Fit groups received reminders regarding the use of the PA app during the first week of the intervention period and autonomously continued to use the app until the intervention ended. Data collection occurred in the fall and early spring to avoid the impact of inclement weather on activity patterns. Participants who successfully completed the 2-month study received a US $100 e-gift card as a token of appreciation.

Figure 1. Data collection time points. PA: physical activity.



Mobile Apps

Google Fit

Google Fit (Google LLC) is a PA-tracking platform developed by Google for Android and Apple iOS. The app was first launched to the public in 2014, and since its release, it has been one of the most popular health/fitness apps in the United States, with >2.6 million monthly active users (Statista 2022 [58]). The app uses a sensor built into a smartphone device to automatically track PA, including steps and active minutes. It also allows users to journal and record a variety of forms of PA (eg, cycling, weightlifting, and yoga) by manually setting the activity tracking mode. Google Fit uses a heart point–based reward system as a gamification strategy to provide users with individualized exercise tips incorporated with PA recommendations outlined by the American Heart Association. Google Fit users can earn one heart point for each minute of MVPA, with possible virtual rewards (ie, celebrative animation on the interface in addition to a green circle morphed into the user's profile image) when they reach a certain number of PA milestones (eg, 30 minutes of moderate PA a day or 150 minutes of MVPA per week). The number of heart points received based on active minutes is the app's primary gamification strategy. To the best of our knowledge, there are no peer-reviewed, data-based publications on the app's functional reliability and behavior change effectiveness. There are a few industrial reports and studies on Google Fit, although these studies focused on the basic functionality of the app and compatibility with wearable trackers [59,60]. We chose Google Fit as a comparison platform as it is free and easy to use, and, most importantly, it has been extensively used by neurotypical adults. As such, the comparison of Google Fit in this study will provide valuable insights into the usability/feasibility of a commercially available health app for promoting PA in adults with ASD.

PuzzleWalk

A gamified behavior change app, PuzzleWalk, available for both Android and Apple iOS, was developed to increase PA and reduce sedentary behavior in adults with ASD following a participatory, user-centered development process, including a needs analysis, literature review, and prototype design [37]. PuzzleWalk incorporates behavior change techniques (BCTs), a theory-based method of promoting healthy behavior change by leveraging psychological determinants, such as autonomy, perceived competence, and intrinsic and extrinsic motivation [61]. The example techniques included in PuzzleWalk are a comprehensive, visualized user guide, self-monitoring of target performance, contingent rewards, and goal setting [37].

It is a *spot the difference* puzzle game comprising 660 major city images around the world (see Figure 2). This format was chosen because it is easy to understand the purpose of the game, and it can quickly capture the user's interests without a complex comprehension process. Moreover, this visual image–based game facilitates visual interaction, which is a unique strength of individuals with ASD [40]. The most unique design element of PuzzleWalk is the conversion algorithm between steps and game-solving time. Specifically, the user's accumulated steps are directly converted to game-solving time to motivate PA participation. A review of the literature indicated that only a few available PA apps use this gamified token economy strategy for PA promotion. Pokémon Go uses a similar gamification strategy as the app links walking activities to the Pokémon character–hunting game supported by location tracking and augmented reality technologies; however, there is no direct conversion algorithm between PA (steps) and game time/opportunity. PuzzleWalk also uses a gamified leaderboard that ranks active users based on their steps and puzzle scores, with tangible rewards (ie, US $10 e-gift cards) provided to the top 3 score leaders at the end of each month. This gamified leaderboard leveraged a BCT of prompt rewards contingent on efforts toward a target behavior [62,63].

Figure 2. PuzzleWalk user interfaces.

Mobile App Use Assessment

PA app use for the past 1 week was assessed through a self-report survey at 3 different data collection time points (ie, end of the fourth week, start of the eighth week, and end of the eighth week) (see Table 1). Participants were asked to report the frequency (days of app use during the past 7 days) and duration (hours and minutes usually spent for app use). The time spent using the PA app was calculated using a minutes per day format.

Table 1. Durations of data collection.

Instrument used	Time 1: baseline (first week)	Time 2: intervention start (fourth week)	Time 3: intervention end (eighth week)	Total
Accelerometry (days)	7	7	7	21
Beck Anxiety Inventory for the past 7-day anxiety assessment	Days 1 and 7	Days 1 and 7	Days 1 and 7	6 times
Daily anxiety assessment (days)	7	7	7	21
Survey for physical activity app use	N/A[a]	Day 7	Days 1 and 7	3 times

[a]N/A: not applicable.

Anxiety Assessment

The Beck Anxiety Inventory (BAI) was used to assess participants' prolonged state of anxiety. The BAI is a self-report scale comprising 21 items that measure the severity of anxiety symptoms during the past week, with high internal consistency (Cronbach α=.92) and test-retest reliability (r=0.75) [64]. Scores range from 0 to 63, with scores of ≥26 indicating potentially concerning levels of anxiety (ie, 0-7=minimal; 8-15=mild; 16-25=moderate, and ≥26=concerningly severe). The participants were asked to report the extent to which they had been bothered by each of the 21 symptoms in the past week. The participants completed the BAI at the start and end of each data collection period (a total of 6 times).

In addition to the BAI, time-specific and type of anxiety trigger questions were asked daily during each data collection period to better identify the contexts of potential anxiety triggers such as environmental, psychological, or sensory factors [65]. At the start of the study, participants downloaded the DAA mobile app

developed by the research team. The participants received a prompt at 8 PM each day during the data collection period to answer specific questions. This time was deemed appropriate for assessing perceived emotional changes over the course of the day based on pilot work and considering typical work/school schedules and bedtimes. The questions included "When did you feel anxious today?" and "What caused your anxiety today?" with the possible choices of (1) change or disruption to routine, (2) sensory oversensitivity or overstimulation, (3) confusion and worries about social and communication situations, (4) specific fears or phobias, (5) too many demands or expectations, (6) being prevented from preferred behaviors or interests, and (7) not applicable. As the traditional retrospective assessment approach can be less reliable [66], the DAA adopted the ecologic momentary assessment method to improve response reliability [67]. Low compliance can be an issue with this type of daily report; therefore, we implemented routine strategies to address this issue, including establishing an efficient data/compliance-tracking system and providing regular reminders to complete the task and monetary incentives [68]. Reminders were sent to the participants via email or SMS text messages based on their preferences recorded in the demographic survey. Only participants with valid survey compliance (ie, participation for 6 times in the BAI and participation in at least four DAAs in each data collection period) and monitor wear (ie, at least 3 valid days in each data collection period) received a monetary incentive.

PA and Sedentary Time Assessment

Daily walking steps, PA intensity, and sedentary time were measured using GT3X+ and ActiGraph triaxial accelerometers (ActiGraph). ActiGraph accelerometers have been extensively used to measure PA with moderate to high reliability in both laboratory and free-living conditions [69-71]. All participants were asked to wear an accelerometer on their right hip during waking hours for 7 consecutive days, including at least 2 weekdays and 1 weekend day over the 3 different data collection periods (baseline, start of the fourth-week intervention, and end of the eighth-week intervention). Accelerometers were programmed to calculate data in 60-second epochs [72]. Daily walking steps were measured using the ActiGraph pedometer function, and sedentary time and activity intensity were identified using the following activity counts per minute cutoffs: <100=sedentary, 100-2019=light, 2020-5999=moderate, >5999=vigorous, and ≥2020=MVPA [73]. These cutoff points were proposed by Troiano et al [73] and have been validated in large data sets of free-living adults [74-76]. The minimum wear time for a valid day is ≥10 hours per day for wake time [77]. Although there is no scientific consensus on the minimum number of valid days, 4 days have been widely accepted in previous studies to reliably estimate habitual PA [78,79]. However, this study required at least 3 valid days of monitor wear in each data collection period in an effort to minimize sample loss [78].

In addition, an intraclass correlation (ICC) analysis was performed to validate the use of at least 3 valid days of monitor wear by examining the relationship between 3 valid days and ≥4 valid days on all PA and sedentary time outcomes at baseline. ICC values of ≥0.75 are generally regarded as good or acceptable reliability [80]. The ICC results demonstrated that the reliability of 3 valid days for all PA and sedentary time variables was excellent (all ICC >0.90); therefore, participants with ≥3 valid days of monitor wear were included in the analyses. Overall, 2 participants' accelerometry data at the intervention start (fourth week) and 1 participant's accelerometry data at the intervention end (eighth week) did not satisfy the minimum number of 3 valid days; thus, their baseline accelerometry data were imputed according to the baseline observation carried forward method [81]. The criterion for nonwear time was 90 minutes of consecutive 0 counts [82], and the accelerometers were set to collect data at sampling rate of 30 Hz.

Data Analysis

Independent *t* tests and chi-square tests were performed for continuous and ordinal variables, respectively, to compare the baseline differences between the 2 groups. Data are presented as mean (SD) or frequency (percentage), according to the variable type. According to the scoring guidelines, the BAI scores for all 21 items (ie, not at all=0, mild=1, moderate=2, and severe=3) were summed to yield a total score of anxiety severity [64]. The BAI scores collected twice during each data collection period were averaged to represent each time point (ie, baseline, intervention start, and intervention end). Data on common anxiety triggers and time-specific occurrences were screened to include only participants with at least four responses to the DAA at each data collection time point. All participants met the compliance criteria, and 134 counts (baseline, N=24), 124 counts (intervention start, 23/24, 96%), and 107 counts (intervention end, 20/24, 83%) were included in the frequency (percentage) analyses.

Objectively measured PA data were first systematically cleaned and verified by examining the ranges and missing values according to the established validation criteria. The collected raw accelerometry data were then converted into activity counts using the ActiLife 6 Data Analysis Software (ActiGraph). With the aforementioned activity intensity cutoffs, the PA data were processed and extracted into an editable spreadsheet. Manual data screening was sequentially performed to verify the minimum required hours and valid days of monitor wear for data analysis; invalid data were eliminated.

An intention-to-treat analysis was conducted to maximize external validity, and the baseline observation carried forward method was used to impute missing data after randomization [81,83,84]. According to the guidelines on missing data in clinical trials by the European Medicines Agency (2010), the baseline observation carried forward method can be appropriate in randomized trial design studies in which researchers reasonably assume that the outcomes of a participant would return to their baseline levels in the long term after dropout [85]. In light of the general span of a mobile health intervention's effectiveness in promoting PA (eg, up to 3 months) [86], this method can also help maintain external validity and minimize sample loss.

The dependent variables were sedentary time (minutes per day), light PA (minutes per day), MVPA (minutes per day), steps per day, total activity counts (vector magnitude), average BAI

anxiety score, and PA app use (minutes per day). To assess PA and anxiety changes over the 3 data collection periods; all measures collected at baseline, intervention start, and intervention end were compared between time points and groups using a repeated-measures analysis of covariance (ANCOVA). Repeated-measures ANCOVA models were adjusted for baseline characteristics, including age, sex, and BMI. The Mauchly test of sphericity was used for each outcome variable to examine the equality of variances of within-group differences across the 3 different data collection time points. In general, if the *P* value was <.05, the assumption of sphericity was violated. Owing to the violation of the sphericity assumption, the Greenhouse-Geisser correction was applied to MVPA, steps, and total activity count variables to interpret the results of the within-group effects.

The effect size (partial η^2) was calculated and defined as >0.02=small, >0.13=medium, or >0.26=large [87]. Owing to the violation of the normality assumption, Spearman rank correlation analyses were performed to determine the baseline correlations between the outcome variables and the impact of increased PA or decreased sedentary time on anxiety change. The degree of change in the outcome variables following PA app use was calculated by subtracting the baseline value from the average of intervention outcomes (eg, MVPA change = [MVPA at intervention start + intervention end]/2 – MVPA at baseline). Data analyses were performed using SPSS 26.0, and significance was declared at *P*<.05 (2-tailed).

Results

A total of 29 adults with ASD initially volunteered to participate in this study. Approximately 17% (5/29) of participants were eliminated before the start of data collection as they did not meet the eligibility criteria or lost study materials. The remaining 83% (24/29) of participants met the eligibility criteria and were enrolled in the study. Of the 24 participants, 3 (13%) from the PuzzleWalk group and 1 (4%) participant from the Google Fit group dropped out during either the intervention start or intervention end time point because of personal obligation (n=1, 4%), invalid monitor wear compliance (n=1, 4%), and restrictions on outdoor activities because of the COVID-19 pandemic (n=2, 8%). The retention rate was 83.3%. On the basis of the intention-to-treat standard, the baseline observation carried forward method was applied to these 4 cases; thus, no participants were lost after data collection was started. Overall, there were no statistically significant baseline differences between the 2 groups. The participant characteristics are presented in Table 2.

The average valid monitor wear for baseline, start of the intervention, and end of the intervention were for 5.8 (82.9%; SD 1.6) days, 5.7 (81.4%; SD 1.6) days, and 5.6 (80.0%; SD 1.4) days, respectively. Of the 24 participants, 13 (54%) at baseline, 10 (46%) at the start of the intervention, and 7 (35%) at the end of the intervention = wore the monitor for the full 7 days. On average, participants wore the monitor for 14.4 (SD 1.7) hours per day, 14.3 (SD 2.1) hours per day, and 14.0 (SD 2.0) hours per day during each data collection period. Regarding anxiety occurrence, adults with ASD experienced anxiety more frequently during the late afternoon—between 3 PM and 7 PM. Overall, participants felt relatively less anxiety during the week of intervention start; however, this positive change was slightly diminished during the last week of intervention (Figure 3).

Figures 4-7 show the descriptive statistics (mean and percentage change) for all outcome measures, including sedentary time, light PA, MVPA, steps, total activity counts, BAI score, and time spent on app use across the 3 data collection periods between the PuzzleWalk and Google Fit groups. The only baseline difference was in daily steps (PuzzleWalk mean 5157.3, SD 2987.2 steps per day, vs Google Fit mean 3094.0, SD 1506.1 steps per day; *P*=.04). There were no significant changes in any of the PA or sedentary time variables over time in either group. The app use time was significantly different between the 2 groups at intervention start (*P*=.046) and intervention end (*P*=.045). The PuzzleWalk group showed a significantly decreased time spent on app use at the start of the final intervention week (*P*=.04); however, the time increased at the end of the intervention (*P*=.049).

Repeated-measures ANCOVAs were performed with age, sex, and BMI as covariates to test if there was a time×group interaction effect between the 2 groups (Tables 3 and 4). There was a significant within-group change over time in sedentary time (*P*=.003) and MVPA (*P*=.04). PA app use was the only variable that resulted in statistically significant pairwise and overall between-group and time×group interaction differences. Specifically, PuzzleWalk participants showed a significant decrease in time spent on PA app use from the start of the intervention to the start of the eighth week (mean 203.5, SD 62.6 minutes per day, to mean 82.9, 38.2 minutes per day; *P*=.008); however, time significantly increased from the start of the eighth week to the end of the intervention (mean 82.9, SD 38.2 minutes per day, to 162.9, 48.3 minutes per day; *P*=.01). In addition, pairwise group differences were found to be statistically significant between PuzzleWalk and Google Fit groups at the start of the intervention (mean 203.5, SD 62.6 minutes per day, vs 12.9, 57.7 minutes per day; *P*=.04) and at the end of the intervention (mean 162.9, SD 48.3 minutes per day, vs 3.3, SD 44.5 minutes per day; *P*=.03) periods. When it comes to overall time×group interactions, there were significant overall between-group (*P*=.04) and time×group interaction (*P*=.01) effects on the time spent on PA app use, indicating that the overall increase in PA app use time was considerably higher in the PuzzleWalk group than in the Google Fit group.

Sedentary time was significantly negatively associated with MVPA (*P*=.03), steps (*P*<.01), and total activity counts (*P*=.021), whereas MVPA was positively associated with steps (*P*<.01) and total activity counts (*P*<.01) at baseline, and these relationships remained significant after the intervention. Anxiety level was not significantly associated with any PA variables or sedentary time at baseline but was changed to have a significant negative association with sedentary time (*P*=.02) and positive associations with light PA (*P*=.045), steps (*P*=.03), and total activity counts (*P*=.045) after the intervention (Table 5).

Table 2. Baseline characteristics of study participants (N=24).

Characteristics	PuzzleWalk (n=12)	Google Fit (n=12)	Total	*P* value
Age (years), mean (SD)	27.1 (7.5)	31.9 (11.3)	29.5 (9.7)	.23
Age at diagnosis (years), n (%)				
Early childhood (birth to age 5)	3 (25)	1 (8)	4 (17)	.74
Later childhood (age 6-11)	2 (17)	3 (25)	5 (21)	.74
Adolescence (age 12-17)	1 (8)	1 (8)	2 (8)	.74
Adulthood (age ≥18)	6 (50)	7 (58)	13 (54)	.74
Female, n (%)	9 (75)	6 (50)	15 (63)	.40
Autism symptoms (AQ-10 score[a,b]), mean (SD)	7.9 (1.6)	8.6 (1.6)	8.3 (1.6)	.33
BMI, mean (SD)	32.2 (7.9)	30.3 (8.9)	31.3 (8.3)	.59
Waist circumference, mean (SD)	38.1 (5.5)	39.3 (5.7)	38.7 (5.5)	.63
Education, n (%)				
High school or General Education Diploma	0 (0)	4 (33)	4 (17)	.16
Some college	5 (42)	4 (33)	9 (38)	.16
College degree	6 (50)	3 (25)	9 (38)	.16
Postgraduate	1 (8)	1 (8)	2 (8)	.16
Employment status, n (%)				
Unemployed	1 (8)	4 (33)	5 (21)	.44
Nonpaid work (eg, volunteer)	1 (8)	2 (17)	3 (13)	.44
Keeping house or homemaker	1 (8)	0 (0)	1 (4)	.44
Student	4 (33)	3 (25)	7 (29)	.44
Paid employment	5 (42)	3 (25)	8 (33)	.44

[a]AQ-10: Autism Spectrum Quotient 10-item.

[b]With a possible maximum score of 10, a higher score on the AQ-10 indicates the presence of more autism symptoms.

Figure 3. Time-specific occurrence of anxiety and perceived anxiety triggers from baseline to intervention end. Multiple answers were allowed for both questions. At least four responses to the Daily Anxiety Assessment were required for each 7-day data collection period. The loss of 4 participants during the intervention period was accounted for in the percentage calculation.

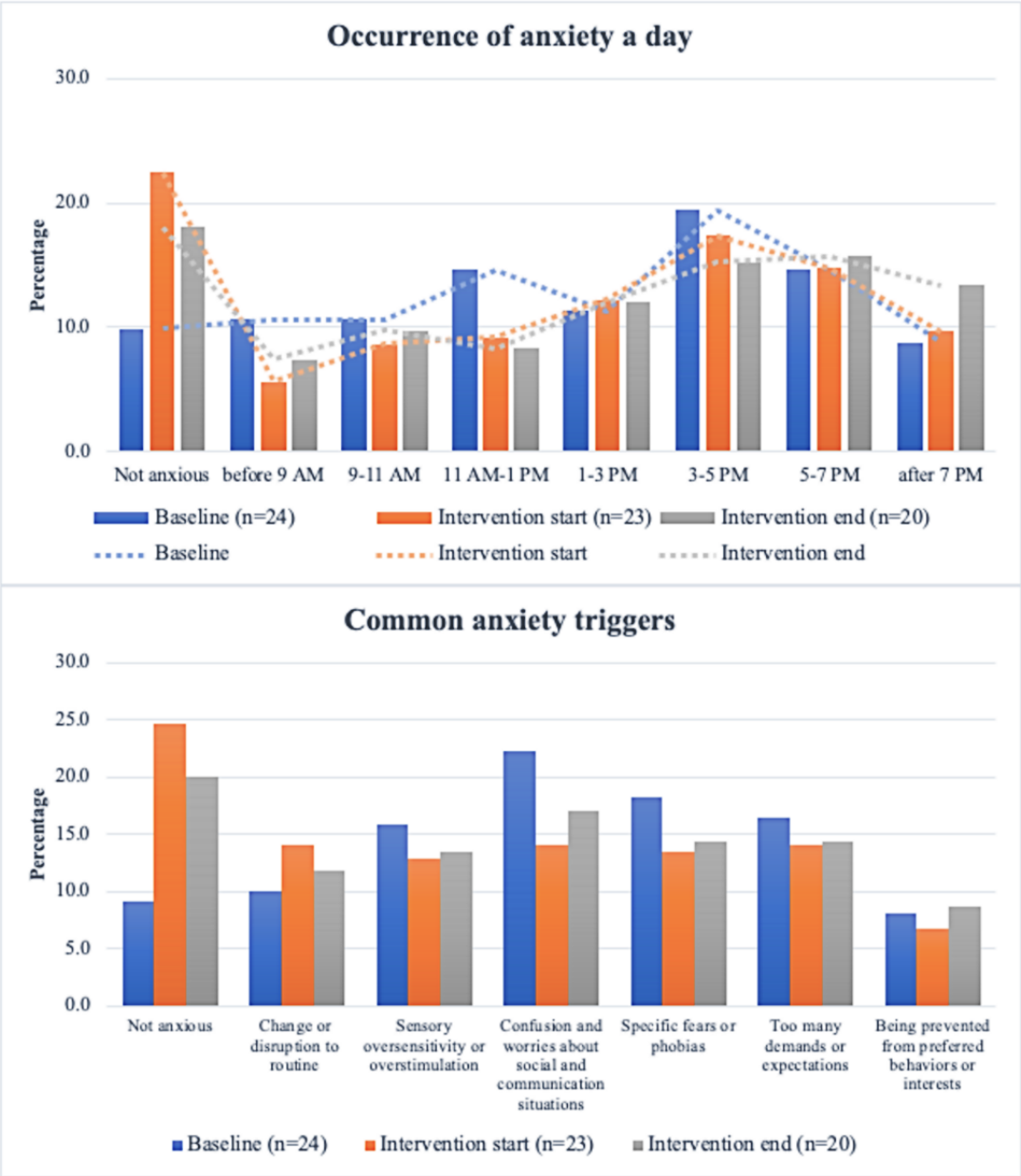


Figure 4. Between- and within-group comparisons of sedentary time and light physical activity. Data are presented as mean (percentage change from baseline).

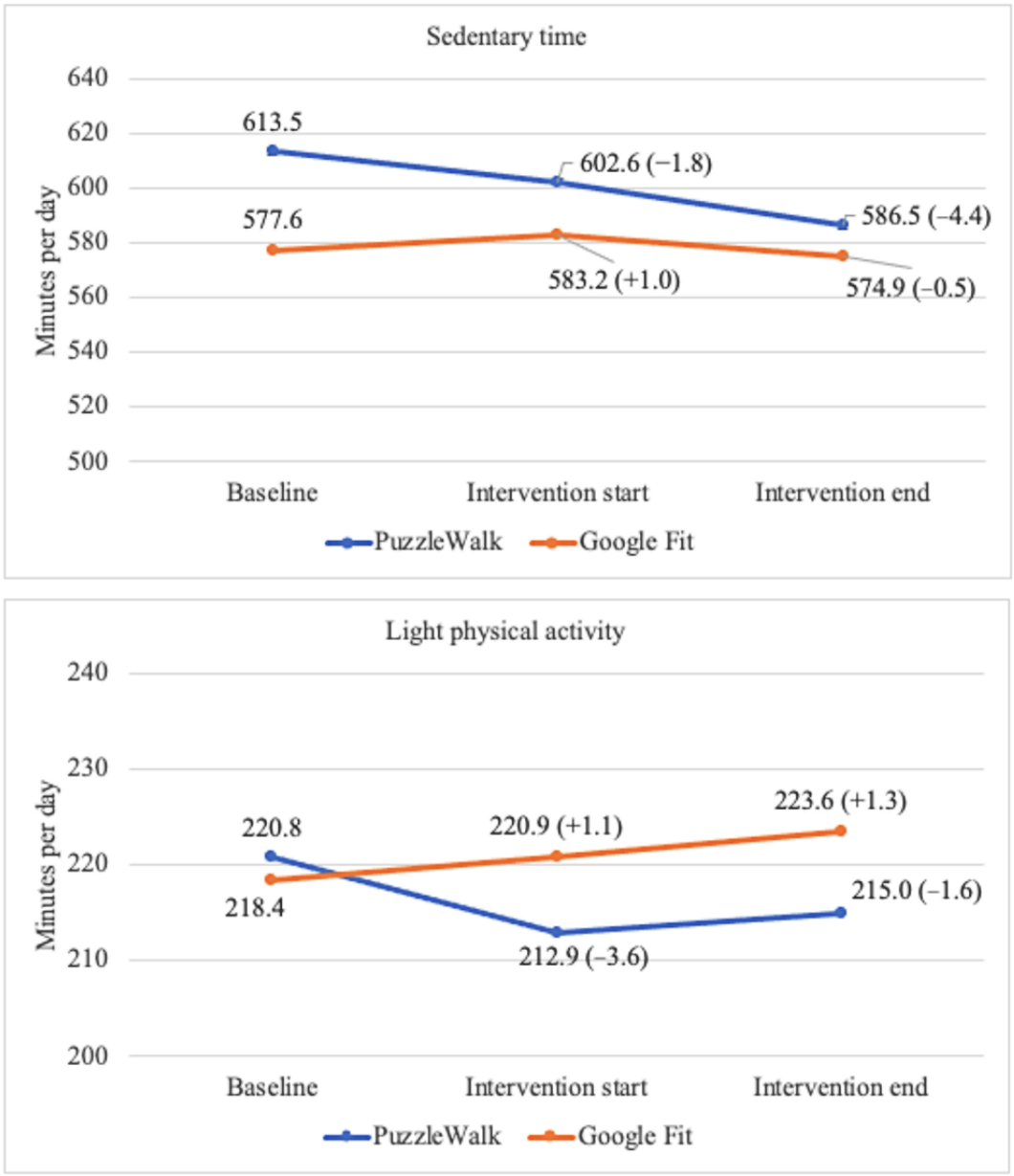


Figure 5. Between- and within-group comparisons of moderate to vigorous physical activity and steps. Data are presented as mean (percentage change from baseline). *P<.05, between-group difference.

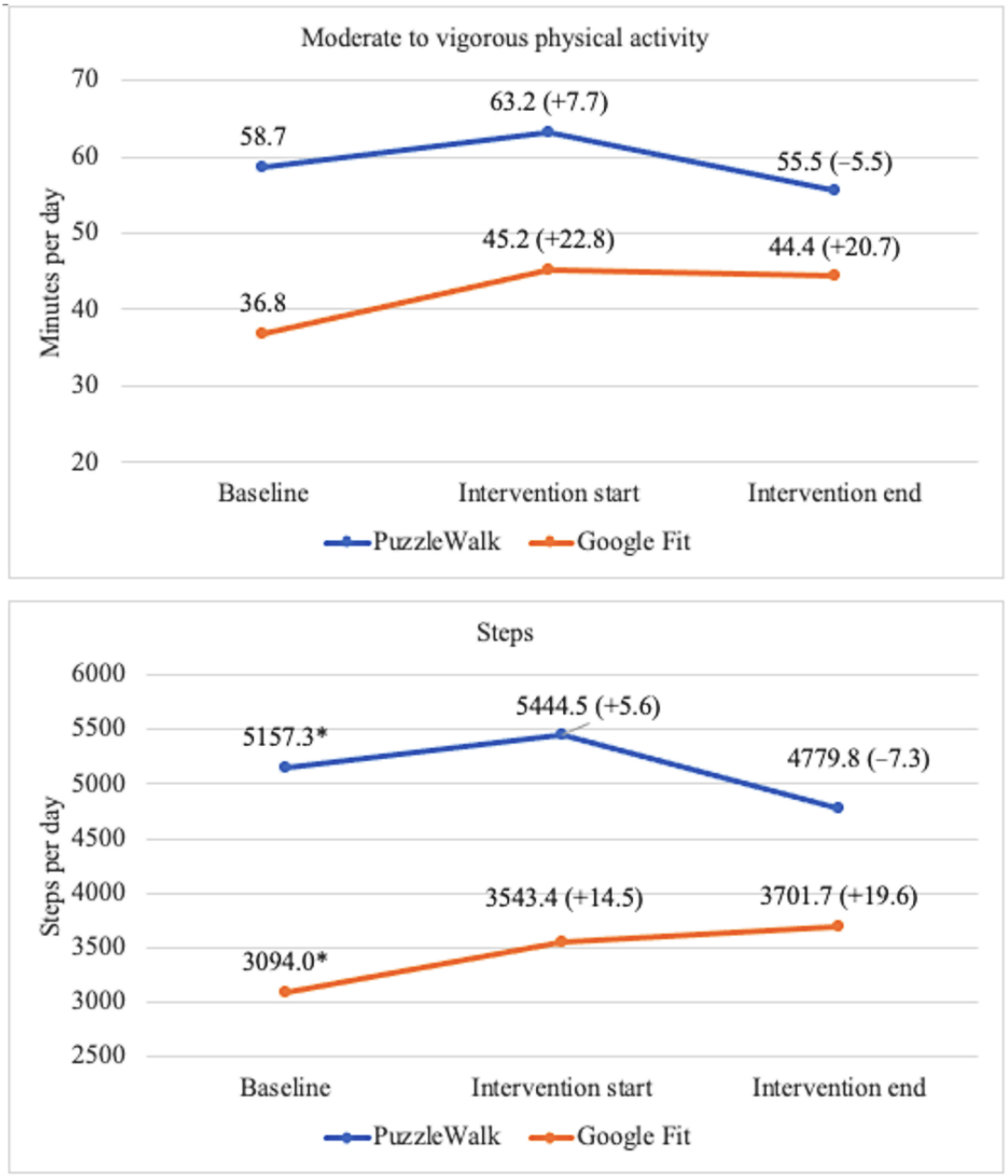


Figure 6. Between- and within-group comparisons of total activity counts and anxiety level. Data are presented as mean (percentage change from baseline).

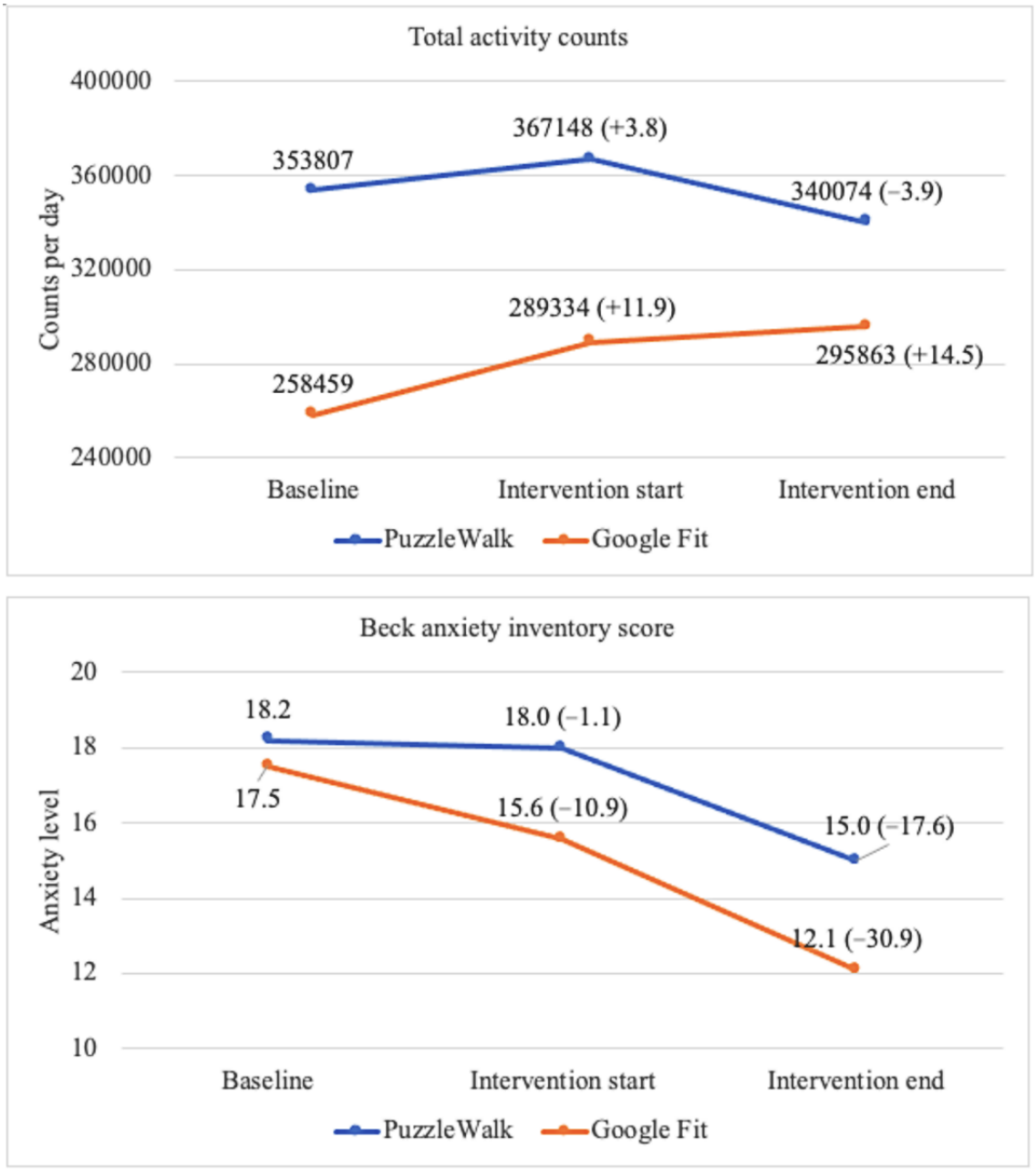


Figure 7. Between- and within-group comparisons of app usage time. Data are presented as mean (percentage change from baseline). *P<.05, between-group difference; **P<.05, within-group difference.

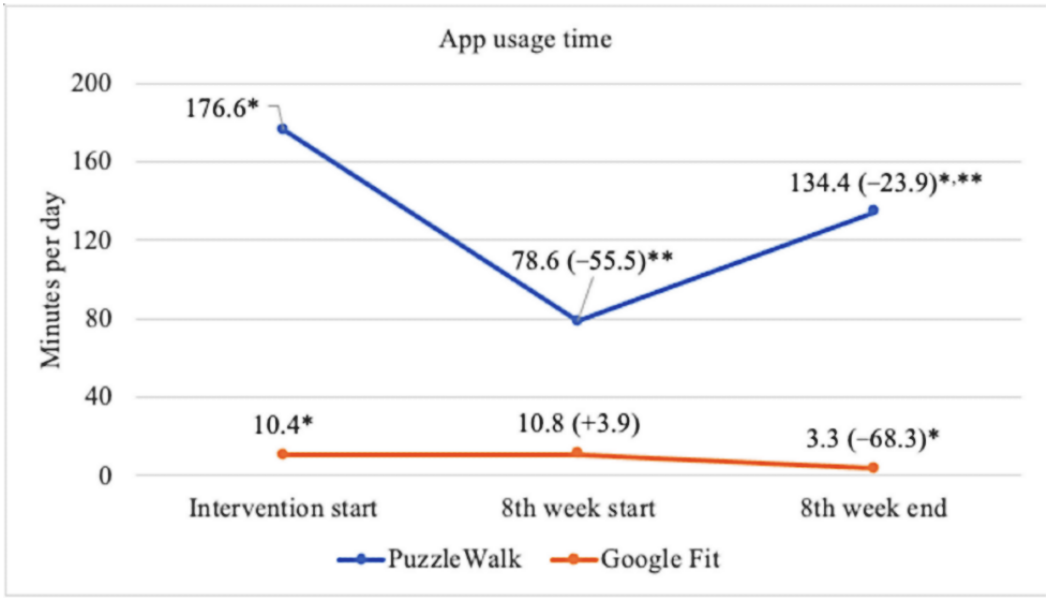


Table 3. Summary of changes in PA[a] and anxiety from baseline to intervention end[b].

Variable	Baseline, mean (SE)	Intervention start, mean (SE)	Intervention end, mean (SE)	Time, F test (η_p^2; df=2,38)	Between-group, F test (η_p^2; df=2,38)	Time×group, F test (η_p^2; df=2,38)
Sedentary time (minutes per day)						
PuzzleWalk	644.0 (31.0)	635.8 (36.2)	627.1 (43.9)	6.83 (0.26)[c]	1.96 (0.09)	0.13 (0.01)
Google Fit	567.5 (28.6)	572.9 (33.3)	557.3 (40.5)	6.83 (0.26)[c]	1.96 (0.09)	0.13 (0.01)
Light PA (minutes per day)						
PuzzleWalk	226.2 (17.5)	220.5 (20.7)	220.7 (19.5)	0.23 (0.01)	0.02 (0.00)	0.24 (0.01)
Google Fit	215.6 (16.1)	217.3 (19.0)	224.4 (18.0)	0.23 (0.01)	0.02 (0.00)	0.24 (0.01)
Moderate to vigorous PA (minutes per day)						
PuzzleWalk	55.9 (9.2)	61.3 (11.2)	50.8 (9.1)	3.47 (0.15)[c]	1.04 (0.05)	1.52 (0.07)
Google Fit	36.9 (8.5)	45.7 (10.3)	46.3 (8.4)	3.47 (0.15)[c]	1.04 (0.05)	1.52 (0.07)
Steps (steps per day)						
PuzzleWalk	4815.9 (787.1)	5083.0 (892.8)	4298.6 (779.2)	1.80 (0.09)	1.28 (0.06)	1.71 (0.08)
Google Fit	3125.7 (725.1)	3609.1 (822.5)	3836.3 (717.9)	1.80 (0.09)	1.28 (0.06)	1.71 (0.08)
Total activity count (counts per day)						
PuzzleWalk	343,351 (42,869)	363,626 (50,927)	321,134 (47,377)	3.42 (0.15)	0.96 (0.05)	1.52 (0.07)
Google Fit	256,596 (39,328)	288,175 (46,917)	304,998 (43,647)	3.42 (0.15)	0.96 (0.05)	1.52 (0.07)
Anxiety level						
PuzzleWalk	17.3 (2.1)	16.2 (2.2)	14.5 (2.9)	2.78 (0.13)	0.02 (0.00)	0.32 (0.02)
Google Fit	17.6 (2.0)	16.5 (2.0)	12.9 (2.6)	2.78 (0.13)	0.02 (0.00)	0.32 (0.02)

[a]PA: physical activity.

[b]Estimated means (SE) for repeated-measures analysis of covariance when adjusted for age, sex, and BMI. [c]P<.05, post hoc comparison within groups, between groups, or overall difference.

Table 4. Summary of changes in PA[a] app use from the intervention start to the intervention end[b].

Variable	Intervention start, mean (SE)	Eighth-week start, mean (SE)	Eighth-week end, mean (SE)	Time, *F* test (η_p^2; *df*=2,38)	Between-group, *F* test (η_p^2; *df*=2,38)	Time×group, *F* test (η_p^2; *df*=2,38)
PA app use (minutes per day)						
PuzzleWalk	203.5 (62.6)[c]	82.9 (38.7)[c]	162.9 (48.3)[c]	0.80 (0.04)	4.37 (0.19)[c]	5.07 (0.21)[c]
Google Fit	12.9 (57.7)[c]	13.2 (35.7)	3.3 (44.5)[c]	0.80 (0.04)	4.37 (0.19)[c]	5.07 (0.21)[c]

[a]PA: physical activity.

[b]Estimated means (SE) for repeated-measures analysis of covariance when adjusted for age, sex, and BMI. [c]*P*<.05, post hoc comparison within groups, between groups, or overall difference.

Table 5. Correlations between PA[a] and anxiety at baseline and their changes over time[b].

Time point and assessments	Sedentary time (minutes per day)	Light PA (minutes per day)	MVPA[c] (minutes per day)	Steps (steps per day)	TAC[d] (counts per day)	Anxiety level
Baseline						
Sedentary time (minutes per day)	1.00	−0.18	−0.46[e]	−0.53[e]	−0.49[e]	0.10
Light PA (minutes per day)	−0.18	1.00	0.20	0.17	0.33	0.11
MVPA (minutes per day)	−0.46[e]	0.20	1.00	0.93[f]	0.97[f]	−0.21
Steps (steps per day)	−0.53[f]	0.17	0.93[f]	1.00	0.92[f]	−0.21
TAC (counts per day)	−0.49[e]	0.33	0.97[f]	0.92[f]	1.00	−0.22
Anxiety level	0.10	−0.11	−0.21	−0.21	−0.22	1.00
Change						
Sedentary time (minutes per day)	1.00	−0.27	−0.59[f]	−0.66[f]	−0.54	−0.49[e]
Light PA (minutes per day)	−0.27	1.00	0.55[f]	0.50[e]	0.68[f]	0.41[e]
MVPA (minutes per day)	−0.59[f]	0.55[f]	1.00	0.91[f]	0.90[f]	0.40
Steps (steps per day)	−0.66[f]	0.50[e]	0.91[f]	1.00	0.87[f]	0.45[e]
TAC (counts per day)	−0.54[f]	0.68[f]	0.90[f]	0.87[f]	1.00	0.41[e]
Anxiety level	−0.49[e]	0.41[e]	0.40	0.45[e]	0.41[e]	1.00

[a]PA: physical activity.

[b]Data are presented as Spearman correlation coefficients, r_s.

[c]MVPA: moderate to vigorous physical activity.

[d]TAC: total activity count.

[e]*P*<.05.

[f]*P*<.01.

Discussion

Principal Findings

The findings from this feasibility study indicate that gamified PA-promoting mobile apps can be an effective tool for decreasing sedentary time and increasing MVPA in intellectually able adults with ASD. Notably, the PuzzleWalk mobile app, developed using BCTs and autism-specific design elements, was comparable with the commercially popular Google Fit in inducing changes in PA and sedentary time. However, the positive improvements in PA and sedentary time did not significantly reduce anxiety levels, although overall anxiety severity for participants with ASD was positively changed from moderate to mild at the end of the intervention.

It appears that intellectually able adults with ASD can benefit similarly from both commercially available and best practice–guided PA mobile apps. Although there were no significant differences in PA and sedentary time between the 2 groups, there were varying trends in the data according to groups, which are worthy of mention. In the PuzzleWalk group, there was a downward trend in sedentary time and a short-term upward trend in MVPA, step count, and total activity count. Conversely, there was a slight intervention-induced increase in light PA and a clear upward trend in steps and total activity counts in the Google Fit group. These observations, in

conjunction with the significantly greater time spent on app use in the PuzzleWalk group compared with the Google Fit group, suggest that the specific research-based design elements of PuzzleWalk, including BCTs and competitive gamification, may pose an advantage over commercially popular apps that do not incorporate these elements. As shown in previous gamified PA interventions for neurotypical samples, adding competition elements to the intervention may be more effective in inducing PA or sedentary behavior changes than collaboration when there are no social connections among users [88,89]. It is of interest to see whether these trends are enhanced in a larger sample of adults with ASD.

Increasing evidence suggests that mobile app interventions can be effective in promoting PA mostly in the short term (eg, up to 3 months) [86]. In a recent study, the short- and long-term effects of a mobile phone–based PA app, together with brief in-person counseling, were evaluated for PA behavior change among inactive middle-aged women (n=210). The findings of this randomized clinical trial indicated that the intervention group's MVPA and total steps were significantly increased during the first 3-month intervention period, although the intervention effect was not maintained in the following 6 months [88]. Similar results of regression fallacy have been reported in earlier randomized controlled trials that examined the efficacy of smartphone sensor–based interventions on PA promotion in healthy adults, which found a substantial PA decline during long-term follow-up [83,89]. Given that adherence and engagement in mobile app interventions generally decline before 3 months [33,86], the attenuation of intervention effects observed in this study (ie, regression of PA outcomes at the intervention end) may correspond with the overall decline in time spent on PA app use during the intervention period. As such, it is critical for mobile app interventions to maximize participant engagement with the app and exposure to the intervention through additional health behavior management strategies such as motivational short message services and telephone coaching [86,90].

This study aimed to identify an anxiety treatment adjunct to, and not in replacement of, conventional therapies. The overall anxiety level decreased from moderate to mild, with decreases in sedentary time and increases in PA over time, in both groups. Although this change was not statistically significant, it may be of clinical or personal value to adults with ASD. A recent review and meta-analysis reported that PA may have a protective effect against anxiety symptoms and disorders; however, the authors cautioned against drawing firm conclusions because of the limited and heterogenic available research [1]. Kim et al [91] recently reported that the optimal range of PA for decreasing anxiety symptoms was 600 to 9000 metabolic equivalents (METs) minutes per week and 1200 to 3000 METs minutes per week for neurotypical men and women, respectively [91]. Individuals with ASD in this study had mean PA levels of 569.9 (SD 45.6) METs minutes per week at the intervention end, which did not meet this criterion [92]; therefore, the intervention-induced changes in PA may have been insufficient to further affect anxiety levels. In addition, many adults with ASD have difficulty articulating their emotions and feelings, and this could have affected the ability of participants in this study to report the full extent of their anxiety symptoms [93]. Despite the lack of robust outcomes regarding intervention effectiveness, additional research on the long-term effects of PA interventions as a simple and economical way of alleviating the health burden of anxiety in adults with ASD is encouraged.

Interestingly, small intervention-induced changes in PA, sedentary time, and anxiety were unfavorably associated. Increases in light PA, steps, and total activity counts and decreases in sedentary time were associated with increased anxiety in adults with ASD. This is contrary to the preponderance of research and a general understanding of the relationships among PA, sedentary time, and anxiety [24,94]. Previous studies have shown that anxiety symptoms are associated with low PA levels and increased sedentary time [95]. Furthermore, Lee et al [96] found that adults with ASD reporting higher sedentary time had increased odds of developing physical and mental health conditions, including anxiety, associated with cardiovascular risk. The underlying mechanisms that explain the relationship between PA, sedentary time, and anxiety are understudied [1]; however, a potential explanation for the findings in this study is that PA-related social situations and environmental changes may have negatively contributed to anxiety symptoms in the sample [7,97]. Confusion and worries about social and communication situations were common anxiety triggers in adults with ASD in this study, as has been previously reported among those with ASD who are often self-aware of their incompetence and difficulty in social situations [7,65,97,98]. As the mobile apps used in this study prompted outdoor walking, which is difficult to perform in social isolation, it is reasonable to conclude that the prospect of entering the public to walk could have elevated anxiety in study participants. Specific types of anxiety were not assessed in this study; however, Hollocks et al [4] found that social phobia was a highly prevalent type of anxiety disorder in adults with ASD. Therefore, future efforts to study PA and sedentary time interventions for this population must consider the social expectations of the prescribed activity.

The major strengths of this study include (1) the use of a covariate-adaptive randomized controlled trial design to control the influence of covariates in the results [99], (2) objective assessment of PA and sedentary time, and (3) relatively good adherence to the study protocol and low attrition. Limitations include a small sample size; the use of self-report for the assessment of anxiety symptoms and app use time; and potential underestimation of PA, as accelerometry cannot accurately detect bicycling, swimming, and other upper-body movements [100]. Regardless, this feasibility, proof-of-concept study demonstrated that intellectually able adults with ASD could favor a gamified PA-promoting mobile app, and this use can induce small improvements in PA, sedentary time, and anxiety, which are worthy of further investigation.

Conclusions

This study demonstrates that a gamified BCT-based mobile app, PuzzleWalk, may be able to decrease the level of sedentary time and create a short-term impact on increasing MVPA, daily steps, and total activity counts among adults with ASD. However, the findings also suggest that anxiety in adults with ASD was

unfavorably related to increased light PA, steps, and total activity counts and decreased sedentary time after the intervention. Further longitudinal research is warranted to evaluate the long-term efficacy of PuzzleWalk in improving physical and mental health and to elucidate the underlying mechanisms that explain the roles of PA and sedentary time in changing anxiety symptoms in adults with ASD. Given the promising usability of a gamified app as a supplementary behavior change tactic, it is recommended that the design elements of mobile health interventions be user centered to meet the unique needs and leverage the strengths of the target population (eg, visual support for users with autism). Furthermore, as supported by many previous findings, mobile health interventions should focus on increasing sustainability and long-term behavior change that can continuously promote regular PA participation. Finally, there is a need for more experimental studies conducted in real-world settings to verify the evidence for gamification and other BCTs in underserved populations [62].

Acknowledgments

This study was supported by the Indiana University Collaborative Research Grants Program and ACSM Foundation Doctoral Student Research Grant from the American College of Sports Medicine Foundation (19-01155).

Authors' Contributions

DL and GCF were involved in the conceptualization, data collection, formal analysis, project administration, and writing of the original draft. DJC, JH, and PCS were engaged in the conceptualization, formal analysis, and reviewing and editing of the manuscript. All the authors have read and accepted the manuscript.

Conflicts of Interest

None declared.

Editorial Notice

This randomized study was only retrospectively registered. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials because the risk of bias appears low and the study was considered formative, guiding the development of the application. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 6163 KB - formative_v6i7e35701_app1.pdf]

References

1. McDowell CP, Dishman RK, Gordon BR, Herring MP. Physical activity and anxiety: a systematic review and meta-analysis of prospective cohort studies. Am J Prev Med 2019 Oct;57(4):545-556. [doi: 10.1016/j.amepre.2019.05.012] [Medline: 31542132]
2. Asmundson GJ, Fetzner MG, Deboer LB, Powers MB, Otto MW, Smits JA. Let's get physical: a contemporary review of the anxiolytic effects of exercise for anxiety and its disorders. Depress Anxiety 2013 Apr;30(4):362-373. [doi: 10.1002/da.22043] [Medline: 23300122]
3. Croen LA, Zerbo O, Qian Y, Massolo ML, Rich S, Sidney S, et al. The health status of adults on the autism spectrum. Autism 2015 Oct;19(7):814-823. [doi: 10.1177/1362361315577517] [Medline: 25911091]
4. Hollocks MJ, Lerh JW, Magiati I, Meiser-Stedman R, Brugha TS. Anxiety and depression in adults with autism spectrum disorder: a systematic review and meta-analysis. Psychol Med 2019 Mar;49(4):559-572. [doi: 10.1017/S0033291718002283] [Medline: 30178724]
5. Fortuna RJ, Robinson L, Smith TH, Meccarello J, Bullen B, Nobis K, et al. Health conditions and functional status in adults with autism: a cross-sectional evaluation. J Gen Intern Med 2016 Jan;31(1):77-84 [FREE Full text] [doi: 10.1007/s11606-015-3509-x] [Medline: 26361965]
6. Billstedt E, Gillberg IC, Gillberg C. Autism after adolescence: population-based 13- to 22-year follow-up study of 120 individuals with autism diagnosed in childhood. J Autism Dev Disord 2005 Jun;35(3):351-360. [doi: 10.1007/s10803-005-3302-5] [Medline: 16119476]
7. Trembath D, Germano C, Johanson G, Dissanayake C. The experience of anxiety in young adults with autism spectrum disorders. Focus Autism Other Dev Disabl 2012 Sep 27;27(4):213-224. [doi: 10.1177/1088357612454916]
8. Buck TR, Viskochil J, Farley M, Coon H, McMahon WM, Morgan J, et al. Psychiatric comorbidity and medication use in adults with autism spectrum disorder. J Autism Dev Disord 2014 Dec;44(12):3063-3071 [FREE Full text] [doi: 10.1007/s10803-014-2170-2] [Medline: 24958436]

9. Jobski K, Höfer J, Hoffmann F, Bachmann C. Use of psychotropic drugs in patients with autism spectrum disorders: a systematic review. Acta Psychiatr Scand 2017 Jan;135(1):8-28. [doi: 10.1111/acps.12644] [Medline: 27624381]
10. Gaus VL. Cognitive-Behavioral Therapy for Adult Asperger Syndrome. New York, NY, USA: The Guilford Press; 2007.
11. Lang R, Regester A, Lauderdale S, Ashbaugh K, Haring A. Treatment of anxiety in autism spectrum disorders using cognitive behaviour therapy: a systematic review. Dev Neurorehabil 2010 Feb;13(1):53-63. [doi: 10.3109/17518420903236288] [Medline: 20067346]
12. Farach FJ, Pruitt LD, Jun JJ, Jerud AB, Zoellner LA, Roy-Byrne PP. Pharmacological treatment of anxiety disorders: current treatments and future directions. J Anxiety Disord 2012 Dec;26(8):833-843 [FREE Full text] [doi: 10.1016/j.janxdis.2012.07.009] [Medline: 23023162]
13. Beyazyüz M, Albayrak Y, Eğilmez OB, Albayrak N, Beyazyüz E. Relationship between SSRIs and metabolic syndrome abnormalities in patients with generalized anxiety disorder: a prospective study. Psychiatry Investig 2013 Jun;10(2):148-154 [FREE Full text] [doi: 10.4306/pi.2013.10.2.148] [Medline: 23798963]
14. Kobak KA, Wolitzky-Taylor K, Craske MG, Rose RD. Therapist training on cognitive behavior therapy for anxiety disorders using Internet-based technologies. Cognit Ther Res 2017 Apr;41(2):252-265 [FREE Full text] [doi: 10.1007/s10608-016-9819-4] [Medline: 28435174]
15. Ross EL, Vijan S, Miller EM, Valenstein M, Zivin K. The cost-effectiveness of cognitive behavioral therapy versus second-generation antidepressants for initial treatment of major depressive disorder in the United States: a decision analytic model. Ann Intern Med 2019 Dec 03;171(11):785-795 [FREE Full text] [doi: 10.7326/M18-1480] [Medline: 31658472]
16. Ekman E, Hiltunen AJ. Modified CBT using visualization for Autism Spectrum Disorder (ASD), anxiety and avoidance behavior--a quasi-experimental open pilot study. Scand J Psychol 2015 Dec;56(6):641-648 [FREE Full text] [doi: 10.1111/sjop.12255] [Medline: 26565732]
17. Parr JR, Brice S, Welsh P, Ingham B, Le Couteur A, Evans G, et al. Treating anxiety in autistic adults: study protocol for the Personalised Anxiety Treatment-Autism (PAT-A©) pilot randomised controlled feasibility trial. Trials 2020 Mar 14;21(1):265 [FREE Full text] [doi: 10.1186/s13063-020-4161-2] [Medline: 32171316]
18. Dietz PM, Rose CE, McArthur D, Maenner M. National and state estimates of adults with autism spectrum disorder. J Autism Dev Disord 2020 Dec;50(12):4258-4266 [FREE Full text] [doi: 10.1007/s10803-020-04494-4] [Medline: 32390121]
19. Nightingale S. Autism spectrum disorders. Nat Rev Drug Discov 2012 Oct;11(10):745-746. [doi: 10.1038/nrd3771] [Medline: 23000684]
20. Katula JA, Blissmer BJ, McAuley E. Exercise intensity and self-efficacy effects on anxiety reduction in healthy, older adults. J Behav Med 1999 Jun;22(3):233-247. [doi: 10.1023/a:1018768423349] [Medline: 10422616]
21. Åstrand PO, Rodahl K, Dahl HA, Strømme SB. Textbook of Work Physiology: Physiological Bases of Exercise. 4th edition. Champaign, IL, USA: Human Kinetics; 2003.
22. Duman RS. Neurotrophic factors and regulation of mood: role of exercise, diet and metabolism. Neurobiol Aging 2005 Dec;26 Suppl 1:88-93. [doi: 10.1016/j.neurobiolaging.2005.08.018] [Medline: 16226350]
23. Hillier A, Murphy D, Ferrara C. A pilot study: short-term reduction in salivary cortisol following low level physical exercise and relaxation among adolescents and young adults on the autism spectrum. Stress Health 2011 Feb 17;27(5):395-402. [doi: 10.1002/smi.1391]
24. Biddle S. Physical activity and mental health: evidence is growing. World Psychiatry 2016 Jun;15(2):176-177 [FREE Full text] [doi: 10.1002/wps.20331] [Medline: 27265709]
25. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity--a systematic review of longitudinal studies. BMC Public Health 2013 Sep 08;13:813 [FREE Full text] [doi: 10.1186/1471-2458-13-813] [Medline: 24010994]
26. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. CMAJ 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
27. Eaves LC, Ho HH. Young adult outcome of autism spectrum disorders. J Autism Dev Disord 2008 Apr;38(4):739-747. [doi: 10.1007/s10803-007-0441-x] [Medline: 17764027]
28. Frey GC, Lee D, Cothran DJ, Harezlak J, Shih PC. Concordance between accelerometer-derived and self-reported physical activity and sedentary time in adults with autism: 623. Med Sci Sports Exerc 2021 Aug 1;53(8S):208-209.
29. Garcia-Pastor T, Salinero JJ, Theirs CI, Ruiz-Vicente D. Obesity status and physical activity level in children and adults with autism spectrum disorders: a pilot study. J Autism Dev Disord 2019 Jan;49(1):165-172. [doi: 10.1007/s10803-018-3692-9] [Medline: 30043355]
30. LaLonde KB, MacNeill BR, Eversole LW, Ragotzy SP, Poling A. Increasing physical activity in young adults with autism spectrum disorders. Res Autism Spectr Disord 2014 Dec 01;8(12):1679-1684. [doi: 10.1016/j.rasd.2014.09.001]
31. Bittner MD, Rigby BR, Silliman-French L, Nichols DL, Dillon SR. Use of technology to facilitate physical activity in children with autism spectrum disorders: a pilot study. Physiol Behav 2017 Aug 01;177:242-246. [doi: 10.1016/j.physbeh.2017.05.012] [Medline: 28502837]
32. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007 Apr 27;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]

33. Jee H. Review of researches on smartphone applications for physical activity promotion in healthy adults. J Exerc Rehabil 2017 Feb;13(1):3-11 [FREE Full text] [doi: 10.12965/jer.1732928.464] [Medline: 28349027]
34. Wang JB, Cataldo JK, Ayala GX, Natarajan L, Cadmus-Bertram LA, White MM, et al. Mobile and wearable device features that matter in promoting physical activity. J Mob Technol Med 2016 Jul;5(2):2-11 [FREE Full text] [doi: 10.7309/jmtm.5.2.2] [Medline: 27493694]
35. Gremaud AL, Carr LJ, Simmering JE, Evans NJ, Cremer JF, Segre AM, et al. Gamifying accelerometer use increases physical activity levels of sedentary office workers. J Am Heart Assoc 2018 Jul 02;7(13):e007735 [FREE Full text] [doi: 10.1161/JAHA.117.007735] [Medline: 29967221]
36. Camargo MC, Barros RM, Brancher JD, Barros VT, Santana M. Designing gamified interventions for autism spectrum disorder: a systematic review. In: Proceedings of the 1st Joint International Conference on Entertainment Computing and Serious Games. 2019 Presented at: ICEC-JCSG '19; November 11-15, 2019; Arequipa, Peru p. 341-352. [doi: 10.1007/978-3-030-34644-7_28]
37. Lee D, Frey G, Cheng A, Shih PC. Puzzle walk: a gamified mobile app to increase physical activity in adults with autism spectrum disorder. In: Proceedings of the 10th International Conference on Virtual Worlds and Games for Serious Applications. 2018 Presented at: VS-Games '18; September 5-7, 2018; Würzburg, Germany p. 1-4. [doi: 10.1109/vs-games.2018.8493439]
38. Kientz JA, Goodwin MS, Hayes GR, Abowd GD. Interactive technologies for autism. In: Baecker R, Sixsmith A, Helal S, Hayes G, editors. Synthesis Lectures on Assistive, Rehabilitative, and Health-Preserving Technologies. Williston, VT, USA: Morgan and Claypool Publishers; Nov 2013:1-177.
39. Goodwin MS. Enhancing and accelerating the pace of autism research and treatment: the promise of developing innovative technology. Focus Autism Other Dev Disabl 2008 Mar 26;23(2):125-128. [doi: 10.1177/1088357608316678]
40. Savner JL, Myles BS. Making Visual Supports Work in the Home and Community: Strategies for Individuals with Autism and Asperger Syndrome. Shawnee, KS, USA: AAPC Publishing; Apr 20, 2000.
41. Quill KA. Instructional considerations for young children with autism: the rationale for visually cued instruction. J Autism Dev Disord 1997 Dec;27(6):697-714. [doi: 10.1023/a:1025806900162] [Medline: 9455729]
42. Cardon TA. Technology and the Treatment of Children with Autism Spectrum Disorder. Cham, Switzerland: Springer; 2016.
43. Bassette L, Kulwicki J, Dieringer ST, Zoder-Martell KA, Heneisen R. The use of a multicomponent behavioral intervention to promote physical activity in adolescents with autism spectrum disorders across inclusive community settings. Behav Anal Pract 2018 Dec 31;11(4):358-369 [FREE Full text] [doi: 10.1007/s40617-018-00285-7] [Medline: 30538909]
44. Alahäivälä T, Oinas-Kukkonen H. Understanding persuasion contexts in health gamification: a systematic analysis of gamified health behavior change support systems literature. Int J Med Inform 2016 Dec;96:62-70. [doi: 10.1016/j.ijmedinf.2016.02.006] [Medline: 26944611]
45. Lister C, West JH, Cannon B, Sax T, Brodegard D. Just a fad? Gamification in health and fitness apps. JMIR Serious Games 2014 Aug 04;2(2):e9 [FREE Full text] [doi: 10.2196/games.3413] [Medline: 25654660]
46. Ferrara J. Games for persuasion: argumentation, procedurality, and the lie of gamification. Games Cult 2013 Aug 27;8(4):289-304. [doi: 10.1177/1555412013496891]
47. Patel MS, Benjamin EJ, Volpp KG, Fox CS, Small DS, Massaro JM, et al. Effect of a game-based intervention designed to enhance social incentives to increase physical activity among families: the BE FIT randomized clinical trial. JAMA Intern Med 2017 Nov 01;177(11):1586-1593. [doi: 10.1001/jamainternmed.2017.3458] [Medline: 28973115]
48. Conroy DE, Yang CH, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. Am J Prev Med 2014 Jun;46(6):649-652. [doi: 10.1016/j.amepre.2014.01.010] [Medline: 24842742]
49. Yang CH, Maher JP, Conroy DE. Implementation of behavior change techniques in mobile applications for physical activity. Am J Prev Med 2015 Apr;48(4):452-455. [doi: 10.1016/j.amepre.2014.10.010] [Medline: 25576494]
50. Byambasuren O, Sanders S, Beller E, Glasziou P. Prescribable mHealth apps identified from an overview of systematic reviews. NPJ Digit Med 2018 May 9;1:12. [doi: 10.1038/s41746-018-0021-9] [Medline: 31304297]
51. National Institutes of Health. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults--the evidence report. National Institutes of Health. Obes Res 1998 Sep;6 Suppl 2:51S-209S. [Medline: 9813653]
52. Klein S, Allison DB, Heymsfield SB, Kelley DE, Leibel RL, Nonas C, et al. Waist circumference and cardiometabolic risk: a consensus statement from shaping America's health: Association for Weight Management and Obesity Prevention; NAASO, the Obesity Society; the American Society for Nutrition; and the American Diabetes Association. Obesity (Silver Spring) 2007 May;15(5):1061-1067 [FREE Full text] [doi: 10.1038/oby.2007.632] [Medline: 17495180]
53. Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E. The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. J Autism Dev Disord 2001 Feb;31(1):5-17. [doi: 10.1023/a:1005653411471] [Medline: 11439754]
54. Woodbury-Smith MR, Robinson J, Wheelwright S, Baron-Cohen S. Screening adults for Asperger Syndrome using the AQ: a preliminary study of its diagnostic validity in clinical practice. J Autism Dev Disord 2005 Jun;35(3):331-335. [doi: 10.1007/s10803-005-3300-7] [Medline: 16119474]

55. Taves DR. Minimization: a new method of assigning patients to treatment and control groups. Clin Pharmacol Ther 1974 May;15(5):443-453. [doi: 10.1002/cpt1974155443] [Medline: 4597226]
56. Lin Y, Zhu M, Su Z. The pursuit of balance: an overview of covariate-adaptive randomization techniques in clinical trials. Contemp Clin Trials 2015 Nov;45(Pt A):21-25. [doi: 10.1016/j.cct.2015.07.011] [Medline: 26244705]
57. Cajita MI, Kline CE, Burke LE, Bigini EG, Imes CC. Feasible but not yet efficacious: a scoping review of wearable activity monitors in interventions targeting physical activity, sedentary behavior, and sleep. Curr Epidemiol Rep 2020 Mar;7(1):25-38 [FREE Full text] [doi: 10.1007/s40471-020-00229-2] [Medline: 33365227]
58. Ceci L. Most popular health and fitness apps in the United States as of May 2018, by monthly active users (in millions). Statista. 2022 Jan 18. URL: https://www.statista.com/statistics/650748/health-fitness-app-usage-usa/ [accessed 2022-06-20]
59. Henriksen A, Haugen Mikalsen M, Woldaregay AZ, Muzny M, Hartvigsen G, Hopstock LA, et al. Using fitness trackers and smartwatches to measure physical activity in research: analysis of consumer wrist-worn wearables. J Med Internet Res 2018 Mar 22;20(3):e110 [FREE Full text] [doi: 10.2196/jmir.9157] [Medline: 29567635]
60. Menaspà P. Effortless activity tracking with Google Fit. Br J Sports Med 2015 Dec;49(24):1598. [doi: 10.1136/bjsports-2015-094925] [Medline: 26359347]
61. 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]
62. Lee D, Frey GC, Min A, Kim B, Cothran DJ, Bellini S, et al. Usability inquiry of a gamified behavior change app for increasing physical activity and reducing sedentary behavior in adults with and without autism spectrum disorder. Health Informatics J 2020 Dec;26(4):2992-3008 [FREE Full text] [doi: 10.1177/1460458220952909] [Medline: 32951500]
63. Kim B, Lee D, Min A, Paik S, Frey G, Bellini S, et al. PuzzleWalk: a theory-driven iterative design inquiry of a mobile game for promoting physical activity in adults with autism spectrum disorder. PLoS One 2020 Sep 10;15(9):e0237966 [FREE Full text] [doi: 10.1371/journal.pone.0237966] [Medline: 32911501]
64. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988 Dec;56(6):893-897. [doi: 10.1037//0022-006x.56.6.893] [Medline: 3204199]
65. Ozsivadjian A, Knott F, Magiati I. Parent and child perspectives on the nature of anxiety in children and young people with autism spectrum disorders: a focus group study. Autism 2012 Mar;16(2):107-121. [doi: 10.1177/1362361311431703] [Medline: 22297200]
66. Moskowitz DS, Young SN. Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. J Psychiatry Neurosci 2006 Jan;31(1):13-20 [FREE Full text] [Medline: 16496031]
67. Liao Y, Skelton K, Dunton G, Bruening M. A systematic review of methods and procedures used in ecological momentary assessments of diet and physical activity research in youth: an adapted STROBE checklist for reporting EMA studies (CREMAS). J Med Internet Res 2016 Jun 21;18(6):e151 [FREE Full text] [doi: 10.2196/jmir.4954] [Medline: 27328833]
68. Zweben A, Fucito LM, O'Malley SS. Effective strategies for maintaining research participation in clinical trials. Drug Inf J 2009 Jul;43(4):10.1177/009286150904300411 [FREE Full text] [doi: 10.1177/009286150904300411] [Medline: 24311825]
69. O'Neil ME, Fragala-Pinkham MA, Forman JL, Trost SG. Measuring reliability and validity of the ActiGraph GT3X accelerometer for children with cerebral palsy: a feasibility study. J Pediatr Rehabil Med 2014;7(3):233-240. [doi: 10.3233/PRM-140292] [Medline: 25260506]
70. Jarrett H, Fitzgerald L, Routen AC. Interinstrument reliability of the ActiGraph GT3X+ ambulatory activity monitor during free-living conditions in adults. J Phys Act Health 2015 Mar;12(3):382-387. [doi: 10.1123/jpah.2013-0070] [Medline: 24828685]
71. Aadland E, Ylvisåker E. Reliability of the Actigraph GT3X+ accelerometer in adults under free-living conditions. PLoS One 2015 Aug 14;10(8):e0134606 [FREE Full text] [doi: 10.1371/journal.pone.0134606] [Medline: 26274586]
72. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. Med Sci Sports Exerc 1998 May;30(5):777-781. [doi: 10.1097/00005768-199805000-00021] [Medline: 9588623]
73. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
74. Evenson KR, Wen F, Metzger JS, Herring AH. Physical activity and sedentary behavior patterns using accelerometry from a national sample of United States adults. Int J Behav Nutr Phys Act 2015 Feb 15;12:20 [FREE Full text] [doi: 10.1186/s12966-015-0183-7] [Medline: 25889192]
75. Matthews CE, Chen KY, Freedson PS, Buchowski MS, Beech BM, Pate RR, et al. Amount of time spent in sedentary behaviors in the United States, 2003-2004. Am J Epidemiol 2008 Apr 01;167(7):875-881 [FREE Full text] [doi: 10.1093/aje/kwm390] [Medline: 18303006]
76. Metzger JS, Catellier DJ, Evenson KR, Treuth MS, Rosamond WD, Siega-Riz AM. Patterns of objectively measured physical activity in the United States. Med Sci Sports Exerc 2008 Apr;40(4):630-638. [doi: 10.1249/MSS.0b013e3181620ebc] [Medline: 18317384]
77. Tudor-Locke C, Camhi SM, Troiano RP. A catalog of rules, variables, and definitions applied to accelerometer data in the National Health and Nutrition Examination Survey, 2003-2006. Prev Chronic Dis 2012;9:E113 [FREE Full text] [doi: 10.5888/pcd9.110332] [Medline: 22698174]

78. Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nyström C, Mora-Gonzalez J, Löf M, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. Sports Med 2017 Sep;47(9):1821-1845 [FREE Full text] [doi: 10.1007/s40279-017-0716-0] [Medline: 28303543]
79. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S531-S543. [doi: 10.1249/01.mss.0000185657.86065.98] [Medline: 16294116]
80. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
81. Liu-Seifert H, Zhang S, D'Souza D, Skljarevski V. A closer look at the baseline-observation-carried-forward (BOCF). Patient Prefer Adherence 2010 Feb 04;4:11-16 [FREE Full text] [doi: 10.2147/ppa.s8135] [Medline: 20165594]
82. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011 Feb;43(2):357-364 [FREE Full text] [doi: 10.1249/MSS.0b013e3181ed61a3] [Medline: 20581716]
83. Duncan M, Vandelanotte C, Kolt GS, Rosenkranz RR, Caperchione CM, George ES, et al. Effectiveness of a Web- and mobile phone-based intervention to promote physical activity and healthy eating in middle-aged males: randomized controlled trial of the ManUp study. J Med Internet Res 2014 Jun 12;16(6):e136 [FREE Full text] [doi: 10.2196/jmir.3107] [Medline: 24927299]
84. Driehuis F, Barte JC, Ter Bogt NC, Beltman FW, Smit AJ, van der Meer K, et al. Maintenance of lifestyle changes: 3-year results of the Groningen Overweight and Lifestyle study. Patient Educ Couns 2012 Aug;88(2):249-255 [FREE Full text] [doi: 10.1016/j.pec.2012.03.017] [Medline: 22560253]
85. Committee for Medicinal Products for Human Use (CHMP). Guideline on missing data in confirmatory clinical trials. European Medicines Agency. 2010 Jul 2. URL: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-missing-data-confirmatory-clinical-trials_en.pdf [accessed 2019-09-20]
86. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can smartphone apps increase physical activity? Systematic review and meta-analysis. J Med Internet Res 2019 Mar 19;21(3):e12053 [FREE Full text] [doi: 10.2196/12053] [Medline: 30888321]
87. Bakeman R. Recommended effect size statistics for repeated measures designs. Behav Res Methods 2005 Aug;37(3):379-384. [doi: 10.3758/bf03192707] [Medline: 16405133]
88. Fukuoka Y, Haskell W, Lin F, Vittinghoff E. Short- and long-term effects of a mobile phone app in conjunction with brief in-person counseling on physical activity among physically inactive women: the mPED randomized clinical trial. JAMA Netw Open 2019 May 03;2(5):e194281. [doi: 10.1001/jamanetworkopen.2019.4281] [Medline: 31125101]
89. Shin DW, Yun JM, Shin JH, Kwon H, Min HY, Joh HK, et al. Enhancing physical activity and reducing obesity through smartcare and financial incentives: a pilot randomized trial. Obesity (Silver Spring) 2017 Feb;25(2):302-310. [doi: 10.1002/oby.21731] [Medline: 28063226]
90. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
91. Kim SY, Jeon SW, Lee MY, Shin DW, Lim WJ, Shin YC, et al. The association between physical activity and anxiety symptoms for general adult populations: an analysis of the dose-response relationship. Psychiatry Investig 2020 Jan;17(1):29-36 [FREE Full text] [doi: 10.30773/pi.2019.0078] [Medline: 31856560]
92. Brooks AG, Gunn SM, Withers RT, Gore CJ, Plummer JL. Predicting walking METs and energy expenditure from speed or accelerometry. Med Sci Sports Exerc 2005 Jul;37(7):1216-1223. [doi: 10.1249/01.mss.0000170074.19649.0e] [Medline: 16015141]
93. Poquérusse J, Pastore L, Dellantonio S, Esposito G. Alexithymia and autism spectrum disorder: a complex relationship. Front Psychol 2018 Jul 17;9:1196 [FREE Full text] [doi: 10.3389/fpsyg.2018.01196] [Medline: 30065681]
94. Vancampfort D, Stubbs B, Herring MP, Hallgren M, Koyanagi A. Sedentary behavior and anxiety: association and influential factors among 42,469 community-dwelling adults in six low- and middle-income countries. Gen Hosp Psychiatry 2018;50:26-32. [doi: 10.1016/j.genhosppsych.2017.09.006] [Medline: 28987919]
95. Edwards MK, Loprinzi PD. Experimentally increasing sedentary behavior results in increased anxiety in an active young adult population. J Affect Disord 2016 Nov 01;204:166-173. [doi: 10.1016/j.jad.2016.06.045] [Medline: 27351099]
96. Lee D, Kennedy J, Cothran D, Shih P, Dickinson S, Golzarri-Arroyo L, et al. Association between sedentary behavior and cardiovascular disease risk in adults with autism spectrum disorder. In: International Society of Behavioral Nutrition and Physical Activity (ISBNPA) 2019 Annual Meeting. 2019 Presented at: ISBNPA '19; June 4-7, 2019; Prague, Czech Republic p. 180.
97. Robertson A, Stanfield A, Watt J, Barry F, Day M, Cormack M, et al. The experience and impact of anxiety in autistic adults: a thematic analysis. Res Autism Spectr Disord 2018 Feb;46:8-18. [doi: 10.1016/j.rasd.2017.11.006]
98. Maddox BB, White SW. Comorbid social anxiety disorder in adults with autism spectrum disorder. J Autism Dev Disord 2015 Dec;45(12):3949-3960. [doi: 10.1007/s10803-015-2531-5] [Medline: 26243138]

99. Kang M, Ragan BG, Park JH. Issues in outcomes research: an overview of randomization techniques for clinical trials. J Athl Train 2008;43(2):215-221 [FREE Full text] [doi: 10.4085/1062-6050-43.2.215] [Medline: 18345348]
100. Arvidsson D, Fridolfsson J, Börjesson M. Measurement of physical activity in clinical practice using accelerometers. J Intern Med 2019 Aug;286(2):137-153 [FREE Full text] [doi: 10.1111/joim.12908] [Medline: 30993807]

Abbreviations

ANCOVA: analysis of covariance
AQ-10: Autism Spectrum Quotient 10-item
ASD: autism spectrum disorder
BAI: Beck Anxiety Inventory
BCT: behavior change technique
DAA: daily anxiety assessment
ICC: intraclass correlation
MET: metabolic equivalent
MVPA: moderate to vigorous physical activity
PA: physical activity

Edited by G Eysenbach; submitted 14.12.21; peer-reviewed by A Teles, O Korn; comments to author 29.04.22; revised version received 23.05.22; accepted 29.05.22; published 28.07.22.

Please cite as:
Lee D, Frey GC, Cothran DJ, Harezlak J, Shih PC
Effects of a Gamified, Behavior Change Technique–Based Mobile App on Increasing Physical Activity and Reducing Anxiety in Adults With Autism Spectrum Disorder: Feasibility Randomized Controlled Trial
JMIR Form Res 2022;6(7):e35701
URL: https://formative.jmir.org/2022/7/e35701
doi:10.2196/35701
PMID:35900808

Original Paper

Using the PMAQ-AB Mobile App and Management System to Evaluate the Quality of Primary Health Care in Brazil: Qualitative Case Study

Osvaldo de Goes Bay Júnior[1*], PhD; Cícera Renata Diniz Vieira Silva[2*], PhD; Cláudia Santos Martiniano[3*], PhD; Lygia Maria de Figueiredo Melo[4*], PhD; Marize Barros de Souza[4], PhD; Monique da Silva Lopes[5], MSc; Ardigleusa Alves Coelho[3], PhD; Paulo de Medeiros Rocha[6*], PhD; Themis Xavier de Albuquerque Pinheiro[6], PhD; Nadja de Sá Pinto Dantas Rocha[6], PhD; Severina Alice da Costa Uchôa[6*], PhD

[1]Postgraduate Program in Collective Health, Federal University of Rio Grande do Norte, Natal, Brazil
[2]Federal University of Campina Grande, Cajazeiras, Brazil
[3]Department of Nursing, State University of Paraíba, Campina Grande, Brazil
[4]School of Health, Federal University of Rio Grande do Norte, Natal, Brazil
[5]Parnamirim Municipal Health Department, Parnamirim, Brazil
[6]Department of Pediatrics, Federal University of Rio Grande do Norte, Natal, Brazil
[*]these authors contributed equally

Corresponding Author:
Osvaldo de Goes Bay Júnior, PhD
Postgraduate Program in Collective Health
Federal University of Rio Grande do Norte
Av. Sen. Salgado Filho.
Lagoa Nova
Natal, 59064-630
Brazil
Phone: 55 84991343560
Email: osvaldobay_jr@hotmail.com

Abstract

Background: The application of cell phones, similar portable devices (ie, tablets), apps, the internet, and GPS in evaluation have established new ways of collecting, storing, retrieving, transmitting, and processing data or information. However, evidence is incipient as to which technological resources remain at the center of assessment practice and the factors that promote their use by the assessment community.

Objective: This study aimed to analyze the relationship between the use of the National Program for Improving Primary Healthcare Access and Quality's (PMAQ-AB; Programa Nacional de Melhoria do Acesso e da Qualidade da Atenção Básica) mobile app and management system and the external evaluation quality of Brazil's PMAQ-AB.

Methods: We conducted a qualitative case study during the external evaluation of Brazil's PMAQ-AB. Data collection consisted of interviews, focus groups, and document analysis. A total of 7 members from the Department of Primary Care of the Ministry of Health and 47 researchers from various higher education and research institutions across the country participated in the study. Data were categorized using the ATLAS.ti software program, according to the quality standards of the Joint Committee on Standards for Educational Evaluation, following the content analysis approach by Bardin.

Results: The results related to feasibility, thematic scope, field activity management, standardized data collection, data consistency, and transparency. They demonstrated improvements and opportunities for advancements in evaluation mediated by the use of information technology (IT), favored the emergence of new practices and remodeling of existing ones, and took into account the multiple components required by the complex assessment of access and quality in primary health care. Difficulties in technology operation, inoperative systems, and lack of investment in equipment and human resources posed challenges to increasing the effectiveness of IT in evaluation.

Conclusions: The use of technology-based tools—the app and the management system—during the external evaluation offered evaluators a greater opportunity for stakeholder engagement. This also allowed the insertion of different organizational, operational,

and methodological components that are capable of triggering influences and confluences. In addition, this allowed connections in collaborative and synergistic networks to increase the quality and allow the development of a more consistent and efficient evaluation process with greater possibility of incorporating the results into public health policies.

(JMIR Form Res 2022;6(7):e35996) doi:10.2196/35996

KEYWORDS

information technology; information technology management; program evaluation; health evaluation; meta-evaluation; primary health care

Introduction

There has been a remarkable increase in investments in, and access to, information technology (IT) globally [1]. The application of cell phones, similar portable devices (ie, tablets), apps, information management systems, the internet, and GPS has expanded in evaluations [2], helping to overcome challenges related to time, resources, and limited data quality [2-4]. Technological tools currently available for evaluators are mainly used for data collection, management, storage, processing, and retrieval [5-7], as well as for improving coverage, accuracy, efficiency, and efficacy of evaluations, adding value to the information that is produced, which supports management decisions [3,4,7].

A growing demand from the national government in Brazil for improved data on primary health care (PHC) has boosted the development of new technology-based tools—an app and a management system [8]—which were then applied in the external evaluation of Brazil's National Program for Improving Primary Healthcare Access and Quality (PMAQ-AB; Programa Nacional de Melhoria do Acesso e da Qualidade da Atenção Básica) in its three cycles: 2011-2012, 2013-2014, and 2015-2019. The evaluation consisted of interviews with PHC teams and observations of infrastructure and functioning. Thus, a total of 42,975 PHC teams from 5570 Brazilian municipalities were evaluated in the last evaluation cycle [9,10].

Therefore, the external evaluation of the PMAQ-AB used IT through an app (ie, a PMAQ-AB instrument) that was developed for this purpose and was accessible via tablets in order to overcome the challenges of improving the availability, quality, and understanding of data related to PHC access and quality in Brazil. It enabled online and offline use and data transference to a cloud management platform. Furthermore, the External Evaluation Management System (EEMS) [11] was implemented, and it enabled monitoring and management of field data collection, guaranteeing shorter decision-making time [8,12]. The use of components and technological resources brought knowledge updates to the evaluation team to achieve the objectives proposed by the evaluation. Textbox 1 describes some features of the PMAQ-AB external evaluation app and the EEMS.

Textbox 1. Features of the PMAQ-AB app and the External Evaluation Management System.

National Program for Improving Primary Healthcare Access and Quality (PMAQ-AB) mobile app for external evaluation:

- Control of access to the app: access to the app after the interviewer has previously registered
- Control of feasibility of the evaluation app: filling in the identification details of the team of evaluators and the team to be evaluated
- Control of the evaluation app: it is mandatory to have answers to all questions
- Management of the conducted evaluations: this allows sending the finalized evaluations, viewing the obtained metadata, and viewing the location obtained by the GPS
- Finalization of evaluations: the questionnaire is finalized and locked for editing and is then ready to be sent
- Submission of evaluations: all evaluation data are transmitted to a cloud management system, requiring internet access at the time of submission
- Field diary: communication channel between supervisors and their interviewers

External Evaluation Management System:

- Monitoring panel: allows the external assessment to be viewed as it happens
- Registration of the fieldwork team, which would allow access to the external evaluation app
- Filling in the field diary if the user is a supervisor
- Verification of the consistency of the data collected by the evaluators
- The evaluations can be accompanied by the general coordinator while conducting the external evaluation
- Conflict regarding questionnaires: managing the resubmission of assessments; the supervisor must determine which is the correct submission among those sent

The use of software can support external evaluations, but its applicability brings new ethical and methodological challenges that experts need to face so that the use of tools, platforms, and digital approaches reach their full potential [2,13]. The growing

incorporation of IT into work processes, driven by the accelerated evolution and variety of technological innovations, requires evaluators to use specific knowledge, means, techniques, and equipment. Since technology can influence the relationships, norms, practices, and aims of evaluation, it cannot be considered a neutral and random organizational phenomenon [14].

Advances in the understanding of which technological elements remain at the core of evaluation practice and the factors that promote their use by the evaluation community are needed [15]. Technologies are continuously changing; therefore, seeking evidence to elucidate the impacts and the reasons for their application may yield relevant contributions to the evaluation field [16]. Studies indicate the need for a systematic follow-up of technological trends and their influence on the roles and responsibilities of evaluators [2,15,17]. Moreover, only a few empirical studies have focused on the interaction between evaluators and technologies, mainly from the perspective of how this interaction facilitates or hampers quality standards needed for evaluations, which attribute value or merit to promote successful health care policies or programs. Therefore, evaluating the theoretical and methodological basis of the evaluation (ie, a meta-evaluation) is needed to understand the extent of success, which can be guided by quality standards [18-22].

It is also important to emphasize that the success or failure of IT implementation mainly depends on the relationships that are established during its use in practice [23]. Considering the importance of adopting technological innovations for evaluations with broad scopes and extensive territorial coverage—including remote areas, such as the external evaluation of the PMAQ-AB—this study aimed to analyze the relationship between the use of the PMAQ-AB mobile app and management system and the external evaluation quality of Brazil's PMAQ-AB.

Methods

Study Design

We conducted a summative meta-evaluation [21,22] after the assessment process was completed. To do so, we carried out a qualitative case study [24] from the perspective that its use would enable analyzing complex social phenomena in depth and in the context of the real world.

The theory that underlies this analysis is sociomateriality [25], which recognizes the importance of relationships and interactions between the social and the material; it emerges as a theoretical approach that can contribute to giving visibility to the understanding of IT in evaluative practices.

A meta-evaluation was performed in the context of the coordination of the PMAQ-AB external evaluation. Higher education and research institutions and the Department of Primary Care of the Ministry of Health (Departamento de Atenção Básica [DAB]–Ministério da Saúde [MS]) conducted the evaluations from the three cycles. Data collection occurred between July 2018 and December 2019 at the DAB-MS, Brasilia, and the main campuses of higher education and research institutions in Pelotas, Belo Horizonte, Rio de Janeiro, Salvador, Teresina, and Aracaju, Brazil.

The EEMS was intentionally chosen due to its innovative technological system and for being an important step in the large, complex, and innovative payment program (PMAQ-AB) that takes into account the performance of PHC teams [8-10].

Study Sample

We initially conducted a document study. We analyzed documents with public access that discussed external evaluation of the PMAQ-AB or those that were available at the DAB-MS for the purpose of training the fieldwork team. In this sense, the documents were used as communicative devices to elucidate the observed event. Thus, four documents regarding the third evaluative cycle that were published between 2017 and 2019 were included, and documents regarding the two previous cycles were excluded because their content was repetitive, as shown in Table 1.

Table 1. Documents used for data collection.

Document no.	Document title	Document type
D1	PMAQ-AB[a] Application: User Manual; Laboratory of Technological Innovation in Health, Federal University of Rio Grande do Norte	Manual on the use of the external evaluation data collection app made available and used for training the fieldwork team
D2	Field Management System: User Manual; Laboratory of Technological Innovation in Health, Federal University of Rio Grande do Norte	Manual on the External Evaluation Management System made available and used for training the fieldwork team
D3	Manual for PMAQ-AB Fieldwork: 3rd Cycle [12]	Aims to present the PMAQ-AB
D4	Methodological Note for the Certification of Primary Healthcare Teams [10]	Aims to present the applied methodology to certify teams that joined the third PMAQ-AB cycle

[a]National Program for Improving Primary Healthcare Access and Quality (Programa Nacional de Melhoria do Acesso e da Qualidade da Atenção Básica).

We used a purposive sample of 54 participants: 7 from the DAB-MS (1 coordinator and 6 technicians) and 47 from higher education and research institutions (6 coordinators and 41 researchers). Participants were part of the PMAQ-AB external evaluation team (third cycle).

Instrument Construction

The Item Matrix for Evaluating the External Evaluation of Primary Health Care [26] was used to elaborate our data collection instrument. Dimensions, subdimensions, items, and questions that were used to guide data collection were extracted (Table 2).

Table 2. Dimensions, subdimensions, items, and guiding questions used to collect data on the use of information technology in the evaluation of access and quality of primary health care, Brazil, 2020.

Dimensions and subdimensions	Items	Guiding questions
Stakeholder engagement		
• Stakeholder identification (utility) • Evaluator credibility (utility)	• Stakeholder identification • Degree of stakeholder involvement and interaction, and participation mechanisms	• Comment on the degree of involvement and interaction and the participation mechanisms needed to identify the needs of interested parties and those affected by the claimant (Department of Primary Care) in the external assessment.
Evaluation design		
• Practical procedures (feasibility) • Evaluation impact (utility)	• Mechanisms for following the PMAQ-AB[a] external evaluation • Viability and feasibility of operational and methodological procedures	• In your opinion, did the use of the interview as a data collection technique provide credibility to the data collected in the external evaluation? • In your opinion, were there any strategies for monitoring the evaluation by the interested parties during the external evaluation?
Evidence systematization and analysis		
• Information scope and selection (utility) • Valid information (accuracy) • Systematic information (accuracy)	• Credibility of collected data through data collection technique • Reaching useful, valid results through data collection procedures • Use of tablets and software programs to increase credibility, trustworthiness, agility, and security in the process of administering questionnaires and in data storage and treatment	• Comment on the feasibility and viability of the operating procedures adopted during the external assessment to gather information. • Did the information technology tools (tablets and software) used during the external evaluation enable credibility, reliability, agility, and security in the process of applying the external evaluation questionnaires? • Were the data collection procedures adequate to achieve useful and valid results?

[a]PMAQ-AB: National Program for Improving Primary Healthcare Access and Quality (Programa Nacional de Melhoria do Acesso e da Qualidade da Atenção Básica).

Data Collection Procedures

We conducted seven semistructured interviews: one with the coordinator of the General Commission for Monitoring and Evaluation of Primary Healthcare from the DAB-MS and six with the coordinators of higher education and research institutions. We conducted focus groups, following the method by Kitzinger [27], with one moderator and one rapporteur. Seven focus groups were conducted: one group included 7 members from the DAB-MS, and each of six groups included at least 6 researchers from the PMAQ-AB external evaluation.

Interviews and focus groups were audio recorded for approximately 1 hour. Notes were taken by the moderator in the focus group; however, few notes summarizing answers were taken. All material was transcribed, and transcripts were read and compared with the original recording immediately following the focus group. Each interview and focus group was identified using "I" or "FG," respectively, followed by the sequential data collection number (eg, 1, 2,..., n). I-DAB and FG-DAB identified the evaluation members from the DAB-MS. The textual fragments extracted from each document (D) were identified by the reading sequence of the documents (ie, D1, D2, D3, and D4), as shown in Table 1.

Data Analysis

Content analysis, according to Bardin [28], was performed using the ATLAS.ti software program (version 8.4.24; Informer Technologies, Inc). First, interviews, focus groups, and documents were transcribed and imported into the software. A peer-review strategy was used, including a group of 6 researchers (OdGBJ, LMdFM, CSM, MBdS, NdSPDR, and TXdAP), to select citations (ie, context units) and link them to specific codes (ie, recording units). At this stage, comments were given to facilitate comprehension, and initial systematization of ideas was performed to interpret the collected information. Second, reports were developed and submitted to be assessed by 2 researchers (PdMR and SAdCU) in order to validate the linkage between citations and codes.

Following data codification, we selected and extracted citations, codes, and code groups representing dimensions, subdimensions, and items a priori for this study. The analyzed material was organized into two thematic categories: (1) technological and organizational infrastructure adding value to the utility of external evaluation and (2) use of IT and paths for feasibility and accuracy of external evaluation.

Ethics Approval

This study was approved by the research ethics committee of the Onofre Lopes University Hospital, Federal University of Rio Grande do Norte (Certificate of Presentation of Ethical Appreciation: 84537418.1.0000.5292) and followed the resolutions of the Brazilian National Commission for Ethics in Research. All participants signed the Free and Informed Consent Term.

Results

Credibility of the External Evaluator and Organization of the Fieldwork Team

The higher education and research institutions constitute the main institutional support for research and the training of researchers. They establish research and technology centers in the country with the capacity to meet the operational needs of external evaluation. The technological resources used in the fieldwork result from the cooperation established with the higher education and research institutions. This demonstrates that the DAB-MS was cautious about the evaluators' expertise in terms of evaluative research, production, and IT application (ie, assessment focus), as set out below:

> *The higher education and research institutions have an organizational structure to conduct data collection that defines a common profile of professional attributions to operate during the external evaluation phase.* [D3]

> *The higher education and research institutions play a prominent role in the external evaluation of the PMAQ-AB, as they are prepared to develop research processes from conception to consolidation of data analysis and results, despite logistical challenges.* [I2]

The operationalization of the external evaluation using IT required complex logistics that encouraged clear and objective communication between team members, delimiting the scope of work for each one. The documents analyzed in this study emphasize that the organizational structure and the definition of roles for the fieldwork team are fundamental for developing the evaluation and obtaining good results from the field activity.

> *The team that performs the evaluation is composed of a field coordinator, supervisors, and interviewers. Understanding the roles of team members, workflows and responsibilities point to progress in the work process and in the qualification of the obtained data. The application of the external assessment instruments in loco with the team of professionals and users was carried out by the interviewers with the help of tablets.* [D3]

The analyzed documents emphasize that the data obtained through interviews with professionals and users were generally used to certify the PHC teams and to improve public health policies; however, some participants considered the importance of external evaluation results to be an important database that could guide teaching and stimulate research on PHC in Brazil. This was particularly evident in the interview with the DAB-MS.

> *The results of the external evaluation served to certify the teams and guide the improvement of public health policies.* [D4]

> *It is one of the main databases to understand, study, investigate, and analyze what is happening in primary health care in Brazil.* [I-DAB]

Regarding data collection at the national level, the result of which will imply payment for performance, the documents analyzed in this study emphasized the importance of maintaining transparency and efficiency throughout the external evaluation process of the PMAQ-AB. Considering the issues inherent to the dynamics and complexity encountered in the field, a partnership with the higher education and research institutions was proposed due to their ability to promote the organization of actions developed by the work team.

PMAQ-AB Mobile App: Comprehensive Evaluation With a Wide Thematic Scope

An electronic modality was chosen, with data collected and sent through a specific app developed for use on a tablet. It is stated in one of the analyzed documents that the app can improve the performance of evaluators, facilitate data collection, and optimize transmission.

> *The External Evaluation of the PMAQ app aims to be a fundamental tool for data collection across the country. It was developed to simplify data collection, facilitate its use, and allow a better experience by the end user.* [D1]

Interview and focus group participants highlighted that the external evaluation of the PMAQ-AB app, which was available for mobile devices (ie, tablets), made data collection across the country feasible, boosted field activity, reached a large number of respondents, and broadened coverage and evaluation scope. According to the documents analyzed, 42,975 PHC health teams from 5570 Brazilian municipalities were evaluated, in loco and simultaneously. The external evaluation instrument is composed of 903 questions.

> *The move from a paper tool to a tablet has supported advanced data collection as it is a comprehensive territory with broad coverage of teams to be assessed.* [FG1]

> *Technology helped with data collection and processing to cover as many variables as possible.* [I3]

PMAQ-AB Mobile App: Validation Structure and Obtaining Valid Results

The PMAQ-AB app allowed automatic verification of data consistency. Validation rules were established within the evaluation instrument to avoid entering incorrect information, with specific criteria for filling in the field, valid records, and expected input values to guarantee the integrity of the entered data, as illustrated in the following quote:

> *In the external evaluation tool, validation is related to fill-in of blanks (answers to patterns). Validation criteria are a) expected value typed in the tool; b) size of answer (number of characters); c) lack of information when the question is "non-applicable."* [D3]

Within this validation framework, an additional strategy was to create a duplicate questionnaire notification system, in which the app accused the interviewer of the existence of a conflicting evaluation in the database. Duplication of questionnaires leads to data inconsistency, as illustrated in the following quote:

> *Interviewers will be notified of duplicated questionnaires during submission. An alert informing the existence of an equivalent module filled for that team will appear in the external evaluation tool.* [D3]

GPS was used to obtain the coordinates of the units that would participate in the external assessment to improve field activity. For some participants, this strategy could improve field activity by monitoring the location of interviewers during data collection, as noted below:

> *GPS was a technological advance used to improve field activity, identifying whether data are being collected at the defined location.* [FG5]

The Development of the EEMS as a Resource for Managing the Evaluation Practice

The implementation of the EEMS appears in the analyzed documents as being strategic for online monitoring of data collection. The follow-up took place through reports, visualization of graphs, and the status of completed assessments. Access to the system is public, allowing monitoring by managers and local workers. The system has restricted access for higher education and research institutions, with reports prepared for the management of teams of interviewers.

An external evaluation panel was available at the EEMS. It is a public access system that enables follow-up of field activity, allowing for the external evaluation to be viewed as data that are collected in the field and uploaded to the national database. These reports can be downloaded in a spreadsheet format.

> *The system for public follow-up is divided into two sections: Informative Sections (with necessary information regarding conduction and follow-up of field activity) and Field Activity Follow-up Panel (allowing visualization of the external evaluation in parallel to its progress).* [D2]

The EEMS allowed transparency and monitoring of data collection for coordinators and supervisors. In this proposal, the documents emphasize that with restricted access, it was possible to obtain an overview of the field, generate statistics and fieldwork diaries, and download the evaluations carried out.

> *EEMS was developed to allow follow-up and management of field data collection and development of reports containing information collected in real time, facilitating management of higher education and research institutions.* [D3]

The EEMS allowed supervisors and coordinators, even though they were not in the territory, to have daily control of the interviewers' activities during data collection. It was possible to identify and resolve inconsistencies, pending issues, and errors with the information recorded and sent after the application of the questionnaires, thereby increasing the accuracy of the data at each cycle. This was particularly evident in the focus groups.

> *The management system allowed daily control of collected data and follow-up of daily demand, reducing the number of inconsistencies after data collection; in many situations, we already had the answers to the problem.* [FG4]

> *After sending data, they could be tracked using the platform. We could check time and inconsistencies, even though you were not in the field. The platform could resolve issues or observe what could be wrong out in the field.* [FG2]

> *EEMS could establish alerts in case the questionnaires were filled by the interviewer at an incoherent time or time interval estimated for its completion.* [D3]

Data consistency was verified by the evaluators according to the guidelines defined in the protocol for the analysis of the consistency and validation of collected data endorsed by the administration and teaching and research institutions. The parameters required by the app and the EEMS were defined a priori by the DAB-MS.

Use of IT: Skills Development, Inclusion of New Actors, and Possible Planning and Execution Problems

The use of IT has brought new training and learning opportunities for evaluators, often developing unusual skills. Considering the technological tools used to carry out the external evaluation of the PMAQ-AB, the documents highlight that training and simulation were fundamental for standardizing field activity throughout the country.

> *Presentation of registration tools and daily monitoring of field activity with EEMS demonstrations for all access profiles. Presentation of materials and tools for field activity. Simulation of use of electronic equipment and EEMS.* [D3]

Despite the existence of training and education in essential content and basic precepts for understanding and acting in the field, in some focus groups, questions were raised about the lack of knowledge of some functionalities offered by the EEMS, which were known during the field activity. Monitoring the

development and implementation of technology is essential for participants to minimize problems related to its applicability in practice.

> *Technological problems could be fixed if we were following the evolution and implementation of technology. When we joined the database and the validator, we were unable to identify the location of the inconsistencies in the management system. Little by little, we got to know their functions.* [FG6]

When the tablets were chosen to carry out the evaluation, the need became clear for investments in infrastructure (ie, hardware) for large-scale acquisition, due to the territorial extension of Brazil and the number of evaluations to be carried out simultaneously. Budgetary restrictions often do not allow for the acquisition of equipment with guaranteed use, making it difficult to carry out the evaluation, as one interviewee commented:

> *When we think about IT, tablets have a lifespan that cannot be ignored in the field and, as always, we work at the operational limit of tablets due to a lack of resources. We always have to consider the costs to ensure the hardware works and allows us to carry out the assessment.* [I5]

The app used by the external evaluation was developed from the principle of facilitating and simplifying the process of collecting and sending data; however, some flaws were found in the app due to installation of its updates during the field activity. To address this and other issues, an IT team was included in the external assessment, as exemplified in the following quote:

> *The number of updates caused the app to malfunction. The IT support team at our institution resolved the issues.* [FG6]

Discussion

Principal Findings

According to the results, evaluators interacted with people and with important technological components during the external evaluation of the PMAQ-AB, triggering influences, confluences, and collaborative and synergistic connections to increase evaluation quality. Technological resources made it possible to carry out the external evaluation of the PMAQ-AB across the country, supported operationalization of an evaluation with a broad thematic scope, provided resources for managing field activity, minimized errors, accelerated and standardized data collection, ensured information comprehensiveness with minimum inconsistencies, and allowed useful and valid results for team certification, health policies, and research.

Mobile technology with a specific app contributed to data collection following the objectives and expectations of all involved in the evaluation. Available resources were sufficient to provide rigorous application of a new external evaluation questionnaire, automatic verification of data consistency, and simplification of the work process in the field. Technology can also be used to guarantee quality and transparency during the evaluative process.

The EEMS collected the data from the finalized evaluation, enabled the coordination of the teams, and monitored the application of the instrument throughout the national territory. We highlight the ability to provide an overview of evaluations to the public, a daily registry of field activity, and management of resending the applied questionnaires among available resources. Difficulties could be identified, modified, and adjusted during data collection almost in real time. According to the results, the EEMS increased evaluation efficiency, reduced errors, and supported evaluators with a structure to manage the evaluative practice.

The collaboration between management and higher education and research institutions conferred credibility to incorporate IT into the evaluation. We highlighted the involvement of higher education and research institutions as important social actors for the external evaluation of the PMAQ-AB; their expertise added value to findings because they were evaluators who were external to the DAB-MS. Participation of higher education and research institutions during development and improvement of the EEMS was also emphasized, providing rules and resources for controlling field activity.

Presentation of materials and tools for data collection, attitudes and behaviors from human actors, and simulation of technology in the external evaluation of the PMAQ-AB are inherent to the learning process and were reflected in the application of the tool for data collection, standardized field activity, and improved communication.

Despite benefits, interaction with the mobile app and the EEMS during field activity presented challenges in the external evaluation of the PMAQ-AB. Factors such as technical aptitude, difficulties in technology operation, occasional system failure, and lack of investments in equipment and human resources were reported as potential barriers to increased efficiency of IT during evaluation.

Integration of Findings With the Current Literature

Enhanced efficiency and efficacy of data collection and accuracy are among the potential benefits of technology for evaluation practice [6,7,15,29,30], thus overcoming challenges related to timing, resources, and restrictions concerning data quality [2-4]. Technology also allows adjustments during evaluation, improves interventions and results, ensures feasible evaluations [6], ensures broad coverage during data collection (ie, inclusion of vulnerable groups and those who are difficult to reach) [2], and brings new voices and social participation (ie, diverse information) [6]. Technology can facilitate integration of different data sources to create a more comprehensive evaluation system [3], store large amounts of data [31], facilitate access and rapid exchange of information [1,4,32], and encourage evaluators to share public data [6].

Functionalities within the PMAQ-AB app allowed better integration of information, minimum inconsistencies, and secure data collection. Evaluators and app developers must guarantee adequate storage (ie, server security) [3], mainly due to the possibility of storing large amounts of data.

Consistencies are automatically verified with the advancement of digitalized evaluation, and data presenting discrepancies can

be identified during collection, contributing to their integrity [30]. Some authors consider app development relevant because the questionnaire can be applied in specific time intervals, and unanswered questions are signaled to the interviewer [3]. This scenario of greater control due to technology can create opportunities to improve the accountability of those involved with the evaluation [29].

Establishing a protocol for data management before initiating data collection increases efficiency, reduces errors, and supports evaluators with a structure adjusted to deal with challenges. This protocol must be shared with the entire team during evaluation planning [3]. Management and follow-up are conducted online by a leading researcher or leader of the evaluation team when interviewers conduct several evaluations. They are responsible for verifying data and identifying errors during data collection. Real-time feedback can also be provided in case of inconsistencies [5,6,30].

Regarding challenges for IT, the literature points out efficiency during the evaluation, the lack of technical ability [3], hardware malfunction [33], low institutional capacity due to insufficient budget to include technology into their operations, and difficulties regarding technological implementation [34]. Findings of what works and does not work should also be presented because this can help individuals and organizations to understand information, share learned lessons, make decisions, and continue progress [30].

Implications for Practice Emerging From This Work

The inclusion of new technologies into organizations influences new practices or reshapes existing practices [35,36]. The success of implementing information technology depends on interactions during the process and the final product. The use of different tools and processes aims to enhance and guarantee the quality of the evaluators' practice, and some of these efforts reflect existing structures [37].

The independence of the evaluator in the development and application of technological resources in conducting an evaluation also implies credibility for the collected data [38]. Implantation of systems for data collection management ensures credibility, use of the evaluation [39], and development of more participative and democratic evaluative approaches [40]. Technology can play a central role in evaluation management.

Although evaluators prefer the inclusion of technology due to the above-mentioned advantages, evaluative practices may require unusual demands (ie, knowledge and abilities). The development of evaluators' competencies is fundamental for facing potential challenges related to technology [41]. Therefore, specific training regarding the system (ie, how to solve common problems or obtain additional technical support) must be included in the initial evaluative planning. Development of data security and interpersonal abilities is also relevant, both at individual evaluator and organizational levels [40].

Limitations of This Study

Collaboration between the DAB-MS and higher education and research institutions was successful during the meta-evaluation of the PMAQ-AB external evaluation. However, the inclusion of stakeholders (eg, municipal managers, health professionals, and users from PHC) would enhance the inclusiveness of the evaluation and broaden results. This inclusion was not feasible in this study, mainly because a meta-evaluation of the PMAQ-AB was not expected. An additional limitation of this study was the exclusion of software developers, hindering the understanding of technological points relevant to its applicability. This research fills a gap in empirical studies on the use of IT conceived from its sociomateriality and its effects on the quality of evaluative and meta-evaluative practices.

Conclusions

This study demonstrated that incorporating technology-based tools—the PMAQ-AB mobile app and the EEMS—during an external evaluation for data collection conferred utility, feasibility, and accuracy to the external evaluation of the PMAQ-AB. Implantation of technological resources was important to enhance the health evaluative system, with a reduction in operational and administrative costs; homogenization of data collection, guaranteeing data security and integrity; and optimized time and information flow. Cost-effective and clever use of technology offered evaluators a chance to include stakeholders and operational and methodological components that contributed to institutionalizing the evaluation process, allowed the development of a broader evaluative scope, and allowed for a more consistent evaluation with the potential for greater use of results in public health policies.

We recommend caution regarding the use of technology and adjustments of the software for evaluation needs, planning, training, digital inclusion, research integration, technology, and health services management. Further studies focusing on potential effects of using technology for new theoretical and methodological evaluation pathways are needed. The literature also lacks studies regarding the parameters used to determine the type of IT to be applied in evaluation processes, strategies for achieving interoperability of health information systems, and how technology could and should empower stakeholders to follow up evaluative processes.

Acknowledgments

The authors acknowledge the assistance provided by the Graduate Program in Collective Health of the Federal University of Rio Grande do Norte. We would like to express our sincere thanks to all stakeholders in this study for their willingness to share their experiences. This work was supported by the DAB-MS (Departamento de Atenção Básica–Ministério da Saúde), Brazil, and by the National Council for Science and Technology Development (grant 207493 / 2014-6). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors thank Probatus Academic Services for providing scientific language translation, revision, and editing.

Authors' Contributions

OdGBJ, SAdCU, CRDVS, LMdFM, and CSM led the study, supervised all aspects of its conduction, and wrote the manuscript. OdGBJ, MBdS, NdSPDR, TXdAP, and MdSL collected the data, assisted with data analysis and interpretation of data, and wrote the manuscript. AAC, CRDVS, CSM, MdSL, MBdS, and TXdAP assisted with the study design and questionnaire item development. OdGBJ, SAdCU, CRDVS, LMdFM, CSM, and PdMR assisted with data analysis and interpretation of data. All authors helped to conceptualize ideas, analyzed and interpreted the data, drafted and revised the paper, and gave their final approval for the paper to be published.

Conflicts of Interest

None declared.

References

1. Williams F, Oke A, Zachary I. Public health delivery in the information age: The role of informatics and technology. Perspect Public Health 2019 Sep;139(5):236-254 [FREE Full text] [doi: 10.1177/1757913918802308] [Medline: 30758258]
2. Bamberger M, Raftree L, Olazabal V. The role of new information and communication technologies in equity-focused evaluation: Opportunities and challenges. Evaluation 2016 Apr 20;22(2):228-244. [doi: 10.1177/1356389016638598]
3. Materia FT, Miller EA, Runion MC, Chesnut RP, Irvin JB, Richardson CB, et al. Let's get technical: Enhancing program evaluation through the use and integration of internet and mobile technologies. Eval Program Plann 2016 Jun;56:31-42. [doi: 10.1016/j.evalprogplan.2016.03.004] [Medline: 27018831]
4. Colasante E, Benedetti E, Fortunato L, Scalese M, Potente R, Cutilli A, et al. Paper-and-pencil versus computerized administration mode: Comparison of data quality and risk behavior prevalence estimates in the European school Survey Project on Alcohol and other Drugs (ESPAD). PLoS One 2019;14(11):e0225140 [FREE Full text] [doi: 10.1371/journal.pone.0225140] [Medline: 31747446]
5. Scharbatke-Church C, Patel AG. Technology for Evaluation in Fragile and Conflict Affected States: An Introduction for the Digital Immigrant Evaluator. Working Paper. Boston, MA: The Fletcher School, Tufts University and Besa; 2016 Apr. URL: https://sites.tufts.edu/ihs/files/2018/02/Technology-and-Evaluation-Hitachi-Paper.pdf [accessed 2022-06-23]
6. Raftree L, Bamberger M. Emerging Opportunities: Monitoring and Evaluation in a Tech-Enabled World. New York, NY: The Rockerfeller Foundation; 2014 Sep. URL: https://www.rockefellerfoundation.org/wp-content/uploads/Monitoring-and-Evaluation-in-a-Tech-Enabled-World.pdf [accessed 2022-06-23]
7. Azzam T, Jacobson M. Reflections on the future of research on evaluation. New Dir Eval 2015 Dec 11;2015(148):103-116. [doi: 10.1002/ev.20160]
8. da Costa Uchôa SA, Santos Martiniano C, Rêgo de Queiroz AA, de Goes Bay Júnior JO, Gonçalves do Nascimento W, Albuquerque Diniz IV, et al. [Innovation and utility: External evaluation of the National Program for Access and Quality Improvement in Primary Health Care] [Article in Portuguese]. Saúde 2018 Sep;42:100-113 [FREE Full text] [doi: 10.1590/0103-11042018s107]
9. Manual Instrutivo Para As Equipes De Atenção Básica e NASF. Programa Nacional de Melhoria do Acesso e da Qualidade da Atenção Básica (PMAQ) – Terceiro ciclo – (2015-2017). Brasília, Brasil: Ministério da Saúde; 2017. URL: http://189.28.128.100/dab/docs/portaldab/documentos/Manual_Instrutivo_3_Ciclo_PMAQ.pdf [accessed 2022-06-23]
10. Nota Metodológica da Certificação das Equipes de Atenção Básica. Programa Nacional de Melhoria do Acesso e da Qualidade da Atenção Básica (PMAQ-AB) – Terceiro ciclo – 2ª Edição. Brasília, Brasil: Ministério da Saúde; 2019 May. URL: http://189.28.128.100/dab/docs/portaldab/documentos/nota_metodologica_2_ed_ab_absb_certificacao.pdf [accessed 2022-06-23]
11. PMAQ, Sistema de Gestão da Avaliação Externa. 2015. URL: http://pmaq.lais.huol.ufrn.br/gestao/ab/ [accessed 2022-06-23]
12. Manual Para o Trabalho de Campo PMAQ - 3º Ciclo (Avaliação Externa). Orientações Gerais. Versão Preliminar. Brasília, Brasil: Ministério da Saúde; 2017. URL: http://189.28.128.100/dab/docs/portaldab/documentos/manual_de_campo_pmaq_3ciclo.pdf [accessed 2022-06-23]
13. Feder SL. Data quality in electronic health records research: Quality domains and assessment methods. West J Nurs Res 2018 May;40(5):753-766. [doi: 10.1177/0193945916689084] [Medline: 28322657]
14. Bennington TL. Ethical implications of computer-mediated evaluation. New Dir Eval 2004;1999(84):87-103. [doi: 10.1002/ev.1155]
15. Jamieson V, Azzam T. The use of technology in evaluation practice. J Multidiscip Eval 2012;8(18):1-15 [FREE Full text]
16. Azzam T, Robinson D. GIS in evaluation. Am J Eval 2013 Feb 15;34(2):207-224. [doi: 10.1177/1098214012461710]
17. Galen M, Grodzicki D. Utilizing emerging technology in program evaluation. New Dir Eval 2011 Sep 21;2011(131):123-128. [doi: 10.1002/ev.389]
18. Elliot LG. Meta-avaliação: Das abordagens às possibilidades de aplicação. Ensaio 2011;19(73):941-964 [FREE Full text] [doi: 10.1590/s0104-40362011000500011]

19. Figueiró AC, Thuler LC, Dias ALF. Padrões internacionais dos estudos de linha de base. In: de Araújo Hartz ZM, Felisberto E, Vieira da Silva LM, editors. Meta-Avaliação da Atenção Básica à Saúde: Teoria e Prática. Rio de Janeiro, Brasil: Fiocruz; 2008:49-70.
20. Gill S, Kuwahara R, Wilce M. Through a culturally competent lens: Why the program evaluation standards matter. Health Promot Pract 2016 Jan;17(1):5-8 [FREE Full text] [doi: 10.1177/1524839915616364] [Medline: 26679506]
21. Scriven M. Meta-evaluation revisited. J Multidiscip Eval 2009;6(11):iii-viii [FREE Full text]
22. Stufflebeam DL. Meta-evaluation. J Multidiscip Eval 2011 Feb;7(15):99-158 [FREE Full text]
23. Kautz K, Cecez-Kecmanovic D. Sociomateriality and information systems success and failure. In: Proceedings of the International Working Conference on Transfer and Diffusion of IT. 2013 Presented at: The International Working Conference on Transfer and Diffusion of IT; June 27-29, 2013; Bangalore, India p. 1-20 URL: https://link.springer.com/content/pdf/10.1007/978-3-642-38862-0_1.pdf [doi: 10.1007/978-3-642-38862-0_1]
24. Yin RK. Estudo de Caso: Planejamento e Métodos. 3ª Edição. Porto Alegre, Brasil: Bookman; 2005.
25. Leonardi PM. Theoretical foundations for the study of sociomateriality. Inf Organ 2013 Apr;23(2):59-76. [doi: 10.1016/j.infoandorg.2013.02.002]
26. da Costa Uchoa SA, de Goes Bay Junior O, de Medeiros Rocha P, Santos Martiniano Sousa C, da Silva Lopes M, Alves Coelho A, et al. Item Matrix for Evaluating the External Evaluation of Primary Health Care. figshare. 2021. URL: https://doi.org/10.6084/m9.figshare.14676504.v5 [accessed 2022-06-23]
27. Kitzinger J. Grupos focais. In: Pope C, Mays N, editors. Pesquisa Qualitativa na Atenção à Saúde. 3ª Edição. Porto Alegre, Brasil: Artmed; 2009:33-43.
28. Bardin L. Análise de Conteúdo. São Paulo, Brasil: Edições; 2011.
29. The World Bank Staff. ICT for Data Collection and Monitoring & Evaluation: Opportunities and Guidance on Mobile Applications for Forest and Agricultural Sectors. Washington, DC: The World Bank; 2013 Dec. URL: https://documents1.worldbank.org/curated/pt/124251468339606709/pdf/833050WP0ICT0Report0Box0382086B00PUBLIC0.pdf [accessed 2022-06-23]
30. Corlazzoli V, King Wale M, Nowak M, Parsons J. ICTs for Monitoring & Evaluation of Peacebuilding Programmes. London, UK: Department for International Development; 2014 May. URL: https://www.sfcg.org/wp-content/uploads/2014/05/CCVRI-SSP-_ICT-and-ME-_Final.pdf [accessed 2022-06-23]
31. McCullough JS. An introduction to the health IT issue. Am J Manag Care 2015 Dec:866 [FREE Full text]
32. Zeleke AA, Naziyok T, Fritz F, Röhrig R. Data quality and cost-effectiveness analyses of electronic and paper-based interviewer-administered public health surveys: Protocol for a systematic review. JMIR Res Protoc 2019 Jan 30;8(1):e10678 [FREE Full text] [doi: 10.2196/10678] [Medline: 30698530]
33. Singh T, Roy P, Jamir L, Gupta S, Kaur N, Jain DK, et al. Assessment of universal healthcare coverage in a district of North India: A rapid cross-sectional survey using tablet computers. PLoS One 2016;11(6):e0157831 [FREE Full text] [doi: 10.1371/journal.pone.0157831] [Medline: 27351743]
34. O'Leary DE. Technology life cycle and data quality: Action and triangulation. Decis Support Syst 2019 Nov;126:113139. [doi: 10.1016/j.dss.2019.113139]
35. Berenger F, Penna M, da Rocha-Pinto SR, Lima L. A tecnologia gerando novos arranjos organizacionais: Análise do modelo DAO sob a ótica da sociomaterialidade. Revista Vianna Sapiens 2019 Oct 29;10(2):28. [doi: 10.31994/rvs.v10i2.611]
36. Leonardi PM, Barley SR. Materiality and change: Challenges to building better theory about technology and organizing. Inf Organ 2008 Jan;18(3):159-176. [doi: 10.1016/j.infoandorg.2008.03.001]
37. Harnar MA, Hillman JA, Endres CL, Snow JZ. Internal formative meta-evaluation: Assuring quality in evaluation practice. Am J Eval 2020 Sep 02;41(4):603-613. [doi: 10.1177/1098214020924471]
38. Jacobson MR, Azzam T. The effects of stakeholder involvement on perceptions of an evaluation's credibility. Eval Program Plann 2018 Jun;68:64-73. [doi: 10.1016/j.evalprogplan.2018.02.006] [Medline: 29486426]
39. Perrin B. How to manage pressure to change reports: Should evaluators be above criticism? Am J Eval 2018 Sep 09;40(3):354-375. [doi: 10.1177/1098214018792622]
40. Picciotto R. Evaluation and the big data challenge. Am J Eval 2019 Sep 24;41(2):166-181. [doi: 10.1177/1098214019850334]
41. Galport N, Azzam T. Evaluator training needs and competencies: A gap analysis. Am J Eval 2016 Jul 09;38(1):80-100. [doi: 10.1177/1098214016643183]

Abbreviations

D: document (in the context of data collection)
DAB: Department of Primary Care (Departamento de Atenção Básica)
EEMS: External Evaluation Management System
FG: focus group (in the context of data collection)
I: interview (in the context of data collection)
IT: information technology
MS: Ministry of Health (Ministério da Saúde)

PHC: primary health care
PMAQ-AB: National Program for Improving Primary Healthcare Access and Quality (Programa Nacional de Melhoria do Acesso e da Qualidade da Atenção Básica)

Edited by A Mavragani; submitted 27.12.21; peer-reviewed by C Güthlin, M Dohan, J Walsh; comments to author 15.03.22; revised version received 29.04.22; accepted 22.05.22; published 29.07.22.

Please cite as:
Bay Júnior ODG, Diniz Vieira Silva CR, Santos Martiniano C, de Figueiredo Melo LM, Barros de Souza M, Lopes MDS, Coelho AA, de Medeiros Rocha P, de Albuquerque Pinheiro TX, de Sá Pinto Dantas Rocha N, da Costa Uchôa SA
Using the PMAQ-AB Mobile App and Management System to Evaluate the Quality of Primary Health Care in Brazil: Qualitative Case Study
JMIR Form Res 2022;6(7):e35996
URL: https://formative.jmir.org/2022/7/e35996
doi:10.2196/35996
PMID:35904848

Original Paper

Deployment of an End-to-End Remote, Digitalized Clinical Study Protocol in COVID-19: Process Evaluation

Nicole Zahradka[1], PhD; Juliana Pugmire[2], MPH, DrPH; Jessie Lever Taylor[1], BA, MSc; Adam Wolfberg[1], MD, MPH; Matt Wilkes[2], MBChB, PhD

[1]Current Health Inc, Boston, MA, United States

[2]Current Health Ltd, Edinburgh, United Kingdom

Corresponding Author:
Nicole Zahradka, PhD
Current Health Inc
294 Washington Street
Suite 510
Boston, MA, 02108
United States
Phone: 1 516 526 3395
Email: nicole.zahradka@currenthealth.com

Abstract

Background: The SARS-CoV-2 (COVID-19) pandemic may accelerate the adoption of digital, decentralized clinical trials. Conceptual recommendations for digitalized and remote clinical studies and technology are available to enable digitalization. Fully remote studies may break down some of the participation barriers in traditional trials. However, they add logistical complexity and offer fewer opportunities to intervene following a technical failure or adverse event.

Objective: Our group designed an end-to-end digitalized clinical study protocol, using the Food and Drug Administration (FDA)–cleared Current Health (CH) remote monitoring platform to collect symptoms and continuous physiological data of individuals recently infected with COVID-19 in the community. The purpose of this work is to provide a detailed example of an end-to-end digitalized protocol implementation based on conceptual recommendations by describing the study setup in detail, evaluating its performance, and identifying points of success and failure.

Methods: Primary recruitment was via social media and word of mouth. Informed consent was obtained during a virtual appointment, and the CH-monitoring kit was shipped directly to the participants. The wearable continuously recorded pulse rate (PR), respiratory rate (RR), oxygen saturation (SpO_2), skin temperature, and step count, while a tablet administered symptom surveys. Data were transmitted in real time to the CH cloud-based platform and displayed in the web-based dashboard, with alerts to the study team if the wearable was not charged or worn. The study duration was up to 30 days. The time to recruit, screen, consent, set up equipment, and collect data was quantified, and advertising engagement was tracked with a web analytics service.

Results: Of 13 different study advertisements, 5 (38.5%) were live on social media at any one time. In total, 38 eligibility forms were completed, and 19 (50%) respondents met the eligibility criteria. Of these, 9 (47.4%) were contactable and 8 (88.9%) provided informed consent. Deployment times ranged from 22 to 110 hours, and participants set up the equipment and started transmitting vital signs within 7.6 (IQR 6.3-10) hours of delivery. The mean wearable adherence was 70% (SD 19%), and the mean daily survey adherence was 88% (SD 21%) for the 8 participants. Vital signs were in normal ranges during study participation, and symptoms decreased over time.

Conclusions: Evaluation of clinical study implementation is important to capture what works and what might need to be modified. A well-calibrated approach to online advertising and enrollment can remove barriers to recruitment and lower costs but remains the most challenging part of research. Equipment was effectively and promptly shipped to participants and removed the risk of illness transmission associated with in-person encounters during a pandemic. Wearable technology incorporating continuous, clinical-grade monitoring offered an unprecedented level of detail and ecological validity. However, study planning, relationship building, and troubleshooting are more complex in the remote setting. The relevance of a study to potential participants remains key to its success.

(JMIR Form Res 2022;6(7):e37832) doi:10.2196/37832

KEYWORDS

evaluation study; telemedicine; remote consultation; digital divide; research design; virtual clinical trial; decentralized; COVID-19; primary recruitment; social media; virtual care; heart rate; wearable; health care cost; health technology

Introduction

Clinical researchers have leveraged emerging technologies to increase study efficiency and accuracy for over 20 years [1]. Digitalized and decentralized clinical study design allows investigators to recruit more heterogeneous populations, reduce the burdens of participation, and capture the experience of participants in real-world settings [2-4]. The SARS-CoV-2 (COVID-19) pandemic may accelerate the shift toward digitalized and decentralized clinical studies [5], forcing researchers to implement remote solutions that limit in-person interaction, while preserving clinical study integrity [1,5]. These mitigations are essential to prevent clinical study disruption, which have detrimental immediate and long-term effects on outcomes, treatment, and cost [6].

Technology already exists to enable digitalization of most aspects of a successful clinical study from recruitment through outcome collection [7]. Online platforms, such as social media, have been shown to be time-efficient and cost-effective methods of recruitment [2,8]. Teleconsent coupled with e-consenting resources can ensure the 3 elements of consent (information, comprehension, and voluntariness) are met, and have moderate-to-high levels of user satisfaction and ease of use across different populations [3]. Remote data collection tools range from online platforms and custom apps for self-reporting outcomes to wearables that continuously collect physiological measurements [4]. Such observations can be collected at high frequencies, increasing the granularity of data to improve the capture of clinically relevant outcomes in ecologically valid settings, compared to traditional clinical studies that typically involve less frequent observations collected during in-person study visits [9,10]. Ultimately, continuously worn wearable data sources may enable digital biomarkers and predictive models that translate detailed data into trial endpoints, clinically actionable insights, and effective diagnoses [11-13].

A fully remote clinical study protocol requires consideration of external factors that have typically been more easily eliminated, or controlled for, in traditional protocols. Strategies for participant education and "nudges" must be adapted for digital delivery when the underlying research question relies on the data and is not focused on capturing voluntary engagement with the data collection instruments. There is less participant visibility and fewer opportunities to intervene and correct during remote data collection compared to in-person study visits. Remote observational studies, therefore, are less reliable because data collected in this manner are more vulnerable to inconsistency and reliant on participant compliance. The effect of human-computer interaction on data collection, for better or for worse, must be considered during analysis. To an extent, these effects can be monitored (or at least, contextualized) by collecting measures of adherence alongside the primary study outcomes. Logistical considerations, such as equipment shipment duration, become factors when eligibility and data collection are time sensitive or following a technical failure or adverse event.

In this study, we designed an end-to-end digitalized clinical study protocol using the Current Health (CH) remote monitoring platform (Current Health Ltd.) to collect symptoms and continuous physiological data to build novel predictive algorithms of COVID-19 progression and severity in individuals who were recently infected in the community. Risk scores based on demographics and risk scores for hospitalized patients already exist [14]. However, by combining continuous remote patient monitoring with machine learning, the goal was to predict the risk of hospitalization, intensive care unit (ICU) treatment, or death for an individual infected with COVID-19 based on their vital signs while still in the community. The CH wearable and software platform were Food and Drug Administration (FDA) 510(k)–cleared for vital sign collection; therefore, confidence in the quality of vital sign observations captured by the CH wearable was higher than the unregulated general wellness wearables that are commercially available [9]. It was hoped that future CH platform integration of the risk algorithm developed from the vital sign data collected in this clinical study might improve resource allocation for patients after a COVID-19 diagnosis and enable more patient-centered management, increasing confidence for low- and high-risk patients and for those managing their care.

Given the need to recruit individuals positive for COVID-19 within 48 hours, key recruitment methods were social media and word of mouth. Recruitment through social media facilitated rapid reach, to an audience most likely to be eligible, in geographical locations associated with high numbers of COVID-19–positive cases and low vaccination rates [15-17].

Two recruitment methods, in person and pairing with test centers, were considered but not pursued. In-person recruitment was eliminated due to the increased risk of exposure and transmission of the infectious disease to the study team. Test center pairing was explored but was unsuccessful because test sites already had existing partnerships with academic institutions or were discouraged from advertising research studies that might deter COVID-19 testing.

Conceptual recommendations for digitalized and remote clinical studies have been outlined in the literature [7,18,19]. However, detailed examples, reviews, and learnings from actual implementation of these recommendations are limited. Our study implemented best-practice recommendations, although we were unable to progress beyond the pilot stage due to low recruitment. The goal of this paper is to describe the study setup in detail, evaluate its performance, and identify points of success and failure.

Methods

Trial Methodology

A scalable end-to-end digitalized protocol was designed for an observational clinical study in individuals who tested positive for COVID-19 in the community. Following the pilot phase, target recruitment was to be 2000 participants and enrollment was time sensitive following a positive COVID-19 test. The protocol eliminated in-person interaction and limited person-to-person interaction by utilizing commercially available technology**.** Each part of the research study, including recruitment, screening, consent, equipment setup, data collection, and follow-up, was automated, when possible.

Ethical Considerations

The study was advertised on social media platforms, including Facebook, Instagram, and LinkedIn, from March 2021 through May 2021. The language in the recruitment material ranged from general and inclusive, such as "positive COVID-19 test," to emphasizing the time-sensitive inclusion criteria of "tested positive for COVID-19 in the past 48 hours" as a call to action (Multimedia Appendix 1). Emails were also distributed internally to employees of CH, and study information was shared with family and friends through word of mouth. The study was conducted according to the guidelines of the Declaration of Helsinki, and the study protocol and recruitment materials were approved by the Advarra Institutional Review Board (Protocol Pro00047371, December 15, 2020; Advarra, Columbia, MD, USA).

Advertisements and emails directed interested individuals to Community by Current Health, a central resource for all CH research studies [20]. The COVID-19 study page included the time-sensitive inclusion criteria (positive COVID-19 test in the past 48 hours), a study overview, frequently asked questions, and a button to "Volunteer Today," which led to a web-based eligibility form (Jotform), and eligible subjects could schedule an appointment with the study team online. Individuals were also asked whether they were agreeable to be contacted for future studies. Eligibility questions were accompanied by a rule set based on inclusion/exclusion criteria (Multimedia Appendix 2) to automate most of the screening process. After the appointment was scheduled, the individual received a copy of the informed consent document. Study team members contacted eligible individuals using the contact number provided. If not eligible, a popup window indicated they were ineligible.

The virtual appointment with a study team member was typically the first point of interaction between the eligible individual and study team member. This appointment, while remote, was a chance to build rapport with the potential participant while taking them through the informed consent process. Teleconsent was similar to an in-person experience, where the details of the study, benefits, risks, and potential conflicts of interest were explained, and time was provided to address any questions [8]. If the eligible individual was still interested in participating, the informed consent form (ICF) was sent to them through software that enabled signature verification and was Health Insurance Portability and Accountability Act (HIPAA) compliant (DocuSign, Inc.). The consent process took approximately 10-15 minutes, and the study team received an updated ICF with the digital signature of the participant. The study team member who obtained the informed consent was designated as the point of contact with the participant throughout the study duration to maintain consistency and engagement. A central study team contact number was used to monitor communication and provide responses within 24 hours. The communication platform was flexible and included text, email, or phone contacts based on participant preference. A system was developed between the study team and the shipping team to track package movement. A logistics partner (Seko Worldwide LLC, Itasca, Illinois, USA) was engaged to distribute and return equipment. The tracking information was also sent to the participants to increase their engagement and to engender a sense of responsibility. Return labeling and packaging were provided. Once consent was obtained, study equipment was delivered to the participants within 1-2 business days.

Data Collection

During the pilot phase, the CH remote monitoring platform was used to collect vital signs and symptoms from 8 individuals who tested positive for COVID-19 (mean age 35.6 years (SD 10 years); 6 (75%) female; 7 (87.5%) White non-Hispanic; 1 (12.5%) Black or African American) for up to 30 days. Study endpoints included recovery (as defined by the Centers for Disease Control and Prevention Clinical Criteria), hospitalization, or death [21]. Therefore, study participation duration varied (mean 27.1 days, SD 5.4 days). The CH kit included a clinical-grade wearable that continuously recorded pulse rate (PR), respiratory rate (RR), oxygen saturation (SpO_2), skin temperature, and step count and a tablet configured to the local time zone of the participant that administered surveys and task reminders. Participants were spread across 3 US time zones (Eastern, Central, and Mountain). Vital signs collected by the wearable and survey responses recorded in the tablet were transmitted to the CH cloud-based platform as raw waveforms and displayed in the web-based dashboard where compliance to study procedures could be monitored. Vital signs were sent when the CH wearable was in range of the CH transmitter and stored for up to 10 hours on the wearable if out of range. The transmitter was enabled for both home Wi-Fi and cellular communication, broadening participation to those without home internet.

An email was sent parallel to study equipment delivery to prompt the participants to set up the equipment as soon as possible. The email contained recommendations regarding wearing and charging the wearable and answering the daily survey. There was a prompt to complete a welcome survey using a unique weblink (Jotform). The welcome survey included questions about the participant and took less than 5 minutes to complete (Multimedia Appendix 3). Although participants were encouraged to complete the welcome survey at the beginning of the study, survey responses were accepted at any point during the study duration.

The CH kit was set up independently by the participants in their home using the tablet-guided instructions. The study team was available to provide remote assistance if there were difficulties setting up the equipment. The participant wore the CH wearable

on the upper arm. The PR, SpO_2, motion, and skin temperature were recorded at up to 30 samples per minute, and RR was recorded at up to 15 samples per minute when the wearable was on-arm. Two notifications were received on the tablet each day at 9:00 a.m. and 10:00 a.m. (local time zone). The first was a reminder to charge the wearable, and the second was a reminder to complete a brief series of questions about symptoms and decisions each day (Multimedia Appendix 4). The participants were not able to see vital sign data or survey responses in real time. Participant adherence was remotely monitored by applying a threshold alarm to vital sign data. The study team were alerted when a participant did not transmit vital sign data to the CH platform for 4 or more hours. The study team escalated the alert by contacting the participant to see whether they were experiencing technical issues or were away from home or to remind the participant to wear the CH device. The vital signs were not monitored in real time, which was made explicitly clear during the consent process. Participants were encouraged to act as they normally would if they felt unwell. At the end of the study duration, the participants returned the CH kit via mail, receiving up to US $100 (US $25/week) if they successfully adhered to the study protocol, in recognition of their time.

Evaluation Methodology

Metrics were created to quantify each phase of the study that was delivered remotely: recruitment, screening, consent, equipment setup, and data collection (Multimedia Appendix 5). Facebook advertising engagement and CH website traffic were tracked with a web analytics service (Google LLC). Advertisement clicks and website views were counted. Metrics of data collected during the trial from the 8 participants, such as daily wearable adherence, vital signs per day (PR, RR, SpO_2), and symptoms, were calculated, in addition to quantitative metrics of trial evaluation. All metrics were assessed for distribution through visual inspection and the Shapiro-Wilk test. Where metrics were normally distributed, they are presented as the mean (SD), and where they were nonparametric (Shapiro-Wilk significant) they are summarized as the median (IQR).

Results

Enrollment Funnel

Of 13 different study advertisements, 5 (38.5%) were live on Facebook at any one time (Multimedia Appendix 1). There were 8852 clicks on the Facebook advertisements for a total spend of US $6770.35. Community by Current Health and its COVID-19 study page were viewed 8932 and 618 times, respectively. There was a decrease in the mean unique advertisement clicks per day from 100.83 in March 2021 to 28.97 in May 2021. In total, 38 eligibility forms were completed, and 19 (50%) respondents met the eligibility criteria. Of these, 9 (47.4%) were contactable (the remainder were uncontactable or deemed themselves “too sick” to take part). Informed consent was obtained from these 9 individuals, and 8 (88.9%) signed the ICF within 9 (SD 8) minutes of the study team sending the ICF via DocuSign (Table 1).

Deployment times ranged from 22 to 110 hours, with a median time of 41 (IQR 28-68) hours. Participants set up the study equipment and started transmitting vital signs within 7.6 (IQR 6.3-10) hours of delivery, and 5 (62.5%) of 8 participants completed the welcome survey in a median of 7.4 (IQR 7.2-53) hours of receiving the welcome email. In addition, 2 (25%) participants completed the welcome survey after the 30-day study period (time to task completion was 32 and 51 days, respectively), while 1 (12.5%) participant never completed the welcome survey. Welcome survey responses indicated that 7 (87.5%) participants did not have asthma, cancer, chronic obstructive pulmonary disease (COPD), diabetes, heart conditions, high blood pressure, sickle cell disease, or kidney disease; undergo an organ transplant; or take beta blockers. In addition, 3 (37.5%) of the participants were smokers, and 7 (87.5%) participants lived with other people but only 1 (12.5%) participant had another member in the household currently COVID-19 positive.

Study participation varied from 17 to 30 days, with 6 (75%) participants completing 30 days and 2 (25%) participants released at 17 and 20 days, respectively, having met the definition for recovery. The mortality rate was 0%, and none of the participants were hospitalized. The mean wearable adherence was 70% (SD 19%), and the mean daily survey adherence was 88% (SD 21%). The median daily wearable adherence ranged from 52% (IQR 29.2%-82.0%) to 92.7% (IQR 82.6%-96.1%). The median PR per day ranged from 65.4 (IQR 59.3-75.5) to 96.5 (IQR 89.9-100.1) beats per minute. The median RR per day ranged from 15.5 (14.1-18.2) breaths/min to 19.3 (15.9-23.0) breaths per minute. The median SpO_2 per day ranged from 95.8% (IQR 93.0%-97.2%) to 98.0% (IQR 96.7%-98.5%). Reported symptoms decreased over time (Multimedia Appendix 6). Participants triggered 87 technical alarms (12 [13.8%], for low battery and 75 [86.2%], for no data for >4 hours, although in many cases the data were buffered and transfer resumed once they had returned home). The median alarms per patient per day was 0.68 (IQR 0.57-1.0, range 0.25-1.6); these were predominantly in the morning. In addition, 4 (50%) participants met the criteria for wearable adherence (wearable worn for at least 20 hours a day and at least 6 days a week, up to 30 days). However, 4 (50%) missed 2 consecutive days of surveys and 1 (12.5%) failed to return the kit at the end. So, 2 (25%) of 8 participants met the strict criteria for full adherence to the study.

Table 1. Enrollment funnel.

Enrollment step	Participants, n (%)
Assessed for eligibility (N=38)	38 (100)
Eligible (N=38)	19 (50)
Informed consent call (N=19)	9 (47.4)
Enrolled (N=9)	8 (88.9)
Completed trial (N=8)	8 (100)

Discussion

Principal Findings

This work evaluated the performance of an end-to-end digitalized clinical study implemented using best-practice recommendations. In total, 8 participants were enrolled into the study; symptoms and continuous physiological data were collected using the CH platform for up to 30 days. Metrics associated with enrollment, deployment, and adherence to the study procedures were reported, and study data were summarized per day. During study participation, vital signs remained within their normal ranges and symptoms decreased over time (Multimedia Appendix 6). This study demonstrated both the advantages and the compromises inherent in decentralized, digitalized clinical trials, and we hope this experience will be valuable to other research groups considering a similar approach.

Throughout the trial, communication strategies were optimized to maintain engagement with participants without being invasive or time-consuming. A central number was created via Google Voice so that multiple study team members could be in communication with 1 participant from the same source. Email communication was timed around equipment delivery dates, and study milestones, to offer avenues for support at the times when most likely to be needed. Regular contact, by multiple channels, helped maintain engagement and reduced the likelihood of a participant being lost to follow-up. The study spanned several time zones, and rather than being a hindrance, this facilitated recruitment as study staff could offer longer “office hours” for participants. Investigators were nonetheless conscious to strike a balance between forming a relationship with a participant but allowing them sufficient space and anonymity for participant to re-engage if they forgot a study task.

The study logistics had to run parallel with participant communication. Engaging a logistics partner offered a level of flexibility in returning study equipment, and the participants were only paid their accrued study incentives once the equipment had been safely received. Collectively, these measures were reflected in most participants receiving equipment and beginning to transmit data within 48 hours of consent and all but 1 of the kits being returned at the close.

Decentralized, digitalized clinical trials also have unique challenges. Study planning, including multidisciplinary working and participant review, can either be facilitated or made more complex, depending on the circumstances of the work. All software must be HIPAA-compliant. Electronic documents must be organized into a study master file that maintains collaboration but also the integrity of study and participant confidentiality. This is typically accomplished by keeping participant information separate from study identifiers. In the CH platform, a random token unique to each participant was generated and stored along with personally identifiable information (PII) in a secure PII enclave. All clinical data processing was then performed on de-identified data, including only that token.

Wearable devices must be validated in the study setting, including being used by the population in question without contemporaneous instruction, and there should be appropriate internet or cell coverage for data transmission (which may exclude certain populations). Wearable continuous monitoring provides vastly more data than single study visits, so consideration should be given to data storage, triage, and quality. A decision must be made a priori between storing raw data, which maintains maximum fidelity and flexibility for future investigation (but is costly), and committing to a level of aggregation. One argument for choosing wearables that record and store raw waveforms is that wearable sensors vary in quality (only some have achieved FDA 510k clearance for clinical-grade monitoring) and participants are out of sight of investigators. Storing the raw waveforms allows for retrospective audit of vital sign quality, which can be reassuring. At the very least, devices should be selected that supply an indication of data quality as metadata.

Given this potential to collect vast amounts of data, definitions of adherence must be set around the goals of the study and be realistic for the participants throughout their participation. More data are not always more informative, but more of the “right data” will be. Our definition of full adherence was strict, with a high bar set for wearable adherence and survey completion. Although this bar may have been appropriate and achievable in the more severe phases of the illness, once participants felt well again, they started to leave home and re-engage with work, and it proved too high for some. Equally, if technical alerts for no data are to be set, they should also be set at a threshold and cadence mindful of the study goals and likely participant behavior. Although a proportion of our no-data alarms reflected genuine technical difficulty or nonadherence, most were simply triggered by symptom-free participants leaving the house for more than 4 hours at a time.

A remote trial clearly facilitates some aspects of research in an infectious disease population and reduces the likelihood of disease transmission to participants and investigators. However, for those quarantining within their own homes, troubleshooting device issues can be made more complex. It can be hard for

them to retrieve deliveries of study components when the initial delivery has been missed. Patients who are acutely unwell may be less likely to see study advertisements. As discussed earlier, they may also be harder to connect with or may feel too sick or disinclined to participate. This may skew results toward less severe illness. Infections, by their nature, are time sensitive. Logistical delays or issues with equipment may leave some participants ineligible. Relationship building can be more difficult to foster when the participant and investigator never meet in person. It is harder to gauge understanding during the informed consent process, and some study data, particularly sensitive information or demographic details, may be harder to acquire. A balance must be struck between creating appropriate minor hurdles to ensure the participant is serious about completing the study, and ensuring overall ease of participation and appropriate incentives to recognize the minor inconvenience associated with data collection.

Comparison With Prior Work

Social media platforms hold much promise as avenues for recruitment. Their advertising models are designed and priced to target particular demographics in specific locations, and their advertisements are readily amenable to A-B testing. The platforms' broad reach and pricing models such as cost-per-click offer fine control over costs and potential savings to investigators. Ali et al [2] reported enrollment of 6602 participants with 9609 advertisement clicks and a total spending of US $906 over the recruitment period. Although there were 8852 advertisement clicks (US $6770.35) in our study, enrollment was much lower (9 participants). However, without a bricks-and-mortar institution, online recruitment can still struggle to achieve a signal amidst the noise. In the context of COVID-19 studies and the pandemic, ring fencing of some language on the platforms made this signaling more challenging. For example, Facebook prohibits advertisements containing content that asserts or implies personal attributes, including physical health.

Indeed, in our study, the initial phase of recruitment remained the most difficult. Advertisements and their amendments were often delayed, while their content was manually reviewed for language. Although 19 participants were eligible, only 9 agreed to a phone call, with many citing a worsening of their condition as a reason not to participate. In studies of rapidly evolving diseases, such as COVID-19, it would be prudent to tie the process of information and consent directly to testing in order to enroll potential participants once proven positive but before their symptoms become overwhelming.

However, once the challenges of initial recruitment were overcome and participants were in the funnel, all patients except 1 who received a phone call consented to the study. This success at converting interest to consent and participation may have been because online recruitment removed perceived barriers to participation or because word of mouth was also used as a recruitment tool. A systematic review and meta-analysis of online patient recruitment in clinical trials found that traditional offline recruitment strategies (eg, word of mouth) result in higher conversion rates than online recruitment strategies [8]. Proximity to a study site was not required. There was no lead time to appointments, and scheduling and coordination of information sharing and consent were simplified. The economic burden on participants, in terms of travel, time, and opportunity cost, was reduced. In the context of COVID-19, there was no concern about viral transmission to participants or investigators from visiting an institution. Removing these barriers, real and perceived, may allow remote clinical trials to reach more marginalized communities, often the hardest for researchers to enroll and, in the case of COVID-19, disproportionately affected by the pandemic.

Limitations

Ultimately, the perceived relevance of a study to its participants, and the landscape in which it is deployed, will hold the key to its success. This pilot study was opened in February 2021 just as infection rates were falling and vaccination rates were rising. Despite incorporating recommendations for best practices in decentralized trials, and our success at enrolling participants once in the funnel and removing barriers to participation, we still missed our recruitment targets and evaluation of the implementation of a scalable end-to-end digitalized protocol was limited to a small sample size. Decentralized and virtual studies hold enormous promise, and may indeed revolutionize the way studies are conducted, but they will still likely remain 1 of many tools in the clinical trials toolkit.

Conclusion

Our pilot study demonstrated the advantages, challenges, and compromises inherent in digitalized, decentralized remote clinical trials. With a well-calibrated approach to online advertising and enrollment, barriers to recruitment can be removed and the cost reduced. Equipment can be effectively and promptly shipped to participants, without risk of illness transmission during a pandemic. Wearable technology incorporating continuous, clinical-grade monitoring can offer an unprecedented level of detail and ecological validity. However, study planning, relationship building, and troubleshooting are more challenging, and definitions of adherence must be crafted around anticipated participant behavior. The relevance of a study to potential participants, be it in person or remote, remains key to its success, particularly during a pandemic.

Acknowledgments

Conceptualization and funding acquisition were performed by AW and MW; methodology, NZ, JP, JLT, and MW; formal analysis and writing—original draft preparation, NZ and MW; writing—review and editing, JP, JLT, and AW; and project administration, NZ and JLT. All authors have read and agreed to the published version of the manuscript. The authors would like to express their

gratitude and appreciation toward the participants for their involvement in this study. We would also like to thank Kimberly Sciaretta and Marie Wax for their manuscript suggestions.

This project was funded in part with federal funds from the US Department of Health and Human Services, Office of the Assistant Secretary for Preparedness and Response, Biomedical Advanced Research and Development Authority, Division of Research Innovation and Ventures (Contract 75A50120C00190).

Conflicts of Interest

NZ, JP, JLT, AW, and MW are employees of Current Health Inc. (Boston, MA, USA). The Biomedical Advanced Research and Development Authority (BARDA, US Department of Health and Human Services) funders had no role in the design of the study; in the collection, analyses, or interpretation of data; or in the decision to publish the results.

Multimedia Appendix 1
Social media.
[PDF File (Adobe PDF File), 163 KB - formative_v6i7e37832_app1.pdf]

Multimedia Appendix 2
Trial eligibility.
[PDF File (Adobe PDF File), 10 KB - formative_v6i7e37832_app2.pdf]

Multimedia Appendix 3
Welcome survey.
[PDF File (Adobe PDF File), 90 KB - formative_v6i7e37832_app3.pdf]

Multimedia Appendix 4
Daily survey.
[PDF File (Adobe PDF File), 40 KB - formative_v6i7e37832_app4.pdf]

Multimedia Appendix 5
Quantitative metrics of trial evaluation.
[PDF File (Adobe PDF File), 13 KB - formative_v6i7e37832_app5.pdf]

Multimedia Appendix 6
Metrics of data per day during the 30-day study duration. (A) Median (IQR) adherence per day. The red line indicates the number of participants enrolled on each day. (B) Symptom count per day collected from the daily survey administered on the CH tablet. The count is of the number of participants who reported the symptom. Colors represent different symptoms. The numeral above each bar represents the number of participants who completed the daily survey. (C-E) Median (IQR) vital signs per day collected by the CH wearable.
[PDF File (Adobe PDF File), 118 KB - formative_v6i7e37832_app6.pdf]

References

1. Saberi P. Research in the time of coronavirus: continuing ongoing studies in the midst of the COVID-19 pandemic. AIDS Behav 2020 Aug 18;24(8):2232-2235 [FREE Full text] [doi: 10.1007/s10461-020-02868-4] [Medline: 32303924]
2. Ali SH, Foreman J, Capasso A, Jones AM, Tozan Y, DiClemente RJ. Social media as a recruitment platform for a nationwide online survey of COVID-19 knowledge, beliefs, and practices in the United States: methodology and feasibility analysis. BMC Med Res Methodol 2020 May 13;20(1):116 [FREE Full text] [doi: 10.1186/s12874-020-01011-0] [Medline: 32404050]
3. Skelton E, Drey N, Rutherford M, Ayers S, Malamateniou C. Electronic consenting for conducting research remotely: a review of current practice and key recommendations for using e-consenting. Int J Med Inform 2020 Nov;143(2):104271-104279 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104271] [Medline: 32979650]
4. 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 29;27(1):73-77. [doi: 10.1038/s41591-020-1123-x] [Medline: 33122860]
5. McDermott MM, Newman AB. Preserving clinical trial integrity during the coronavirus pandemic. JAMA 2020 Jun 02;323(21):2135-2136. [doi: 10.1001/jama.2020.4689] [Medline: 32211830]
6. Anderson M. How the COVID-19 pandemic is changing clinical trial conduct and driving innovation in bioanalysis. Bioanalysis 2021 Aug;13(15):1195-1203 [FREE Full text] [doi: 10.4155/bio-2021-0107] [Medline: 34275327]
7. Inan OT, Tenaerts P, Prindiville SA, Reynolds HR, Dizon DS, Cooper-Arnold K, et al. Digitizing clinical trials. NPJ Digit Med 2020 Jul 31;3(1):101 [FREE Full text] [doi: 10.1038/s41746-020-0302-y] [Medline: 32821856]

8. Brøgger-Mikkelsen M, Ali Z, Zibert JR, Andersen AD, Thomsen SF. Online patient recruitment in clinical trials: systematic review and meta-analysis. J Med Internet Res 2020 Nov 04;22(11):e22179 [FREE Full text] [doi: 10.2196/22179] [Medline: 33146627]
9. Liao Y, Thompson C, Peterson S, Mandrola J, Beg MS. The future of wearable technologies and remote monitoring in health care. Am Soc Clin Oncol Educ Book 2019 Jan;39:115-121 [FREE Full text] [doi: 10.1200/EDBK_238919] [Medline: 31099626]
10. Ferrar J, Griffith GJ, Skirrow C, Cashdollar N, Taptiklis N, Dobson J, et al. Developing digital tools for remote clinical research: how to evaluate the validity and practicality of active assessments in field settings. J Med Internet Res 2021 Jun 18;23(6):e26004 [FREE Full text] [doi: 10.2196/26004] [Medline: 34142972]
11. Joshi M, Ashrafian H, Aufegger L, Khan S, Arora S, Cooke G, et al. Wearable sensors to improve detection of patient deterioration. Expert Rev Med Devices 2019 Feb 06;16(2):145-154. [doi: 10.1080/17434440.2019.1563480] [Medline: 30580650]
12. Burnham JP, Lu C, Yaeger LH, Bailey TC, Kollef MH. Using wearable technology to predict health outcomes: a literature review. J Am Med Inform Assoc 2018 Sep 01;25(9):1221-1227 [FREE Full text] [doi: 10.1093/jamia/ocy082] [Medline: 29982520]
13. Seshadri DR, Davies EV, Harlow ER, Hsu JJ, Knighton SC, Walker TA, et al. Wearable sensors for COVID-19: a call to action to harness our digital infrastructure for remote patient monitoring and virtual assessments. Front Digit Health 2020 Jun 23;2:8 [FREE Full text] [doi: 10.3389/fdgth.2020.00008] [Medline: 34713021]
14. Galloway JB, Norton S, Barker RD, Brookes A, Carey I, Clarke BD, et al. A clinical risk score to identify patients with COVID-19 at high risk of critical care admission or death: an observational cohort study. J Infect 2020 Aug;81(2):282-288 [FREE Full text] [doi: 10.1016/j.jinf.2020.05.064] [Medline: 32479771]
15. Ritchie H, Mathieu E, Rodés-Guirao L. Coronavirus Pandemic (COVID-19). URL: https://ourworldindata.org/coronavirus [accessed 2021-08-25]
16. COVID-19 United States Cases by County. URL: https://coronavirus.jhu.edu/us-map [accessed 2022-07-22]
17. Centers for Disease Control and Prevention (CDC). COVID-19 Case Surveillance Public-Use Data. URL: https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data/vbim-akqf [accessed 2021-08-25]
18. Cummings SR. Clinical trials without clinical sites. JAMA Intern Med 2021 May 01;181(5):680-684. [doi: 10.1001/jamainternmed.2020.9223] [Medline: 33646281]
19. Apostolaros M, Babaian D, Corneli A, Forrest A, Hamre G, Hewett J, et al. Legal, regulatory, and practical issues to consider when adopting decentralized clinical trials: recommendations from the clinical trials transformation initiative. Ther Innov Regul Sci 2020 Jul 9;54(4):779-787 [FREE Full text] [doi: 10.1007/s43441-019-00006-4] [Medline: 32557302]
20. Community by Current Health. Let's Shape The Future of Healthcare Together. URL: https://community.currenthealth.com/ [accessed 2022-07-22]
21. Centers for Disease Control and Prevention (CDC). Ending Isolation and Precautions for People with COVID-19: Interim Guidance. URL: https://www.cdc.gov/coronavirus/2019-ncov/hcp/duration-isolation.html [accessed 2022-05-01]

Abbreviations

CH: Current Health
FDA: Food and Drug Administration
HIPAA: Health Insurance Portability and Accountability Act
ICF: informed consent form
PII: personally identifiable information
PR: pulse rate
RR: respiratory rate
SpO_2: Oxygen saturation

Edited by A Mavragani; submitted 08.03.22; peer-reviewed by K Ho, G Kernohan, B Johnson; comments to author 30.03.22; revised version received 24.06.22; accepted 14.07.22; published 29.07.22.

Please cite as:
Zahradka N, Pugmire J, Lever Taylor J, Wolfberg A, Wilkes M
Deployment of an End-to-End Remote, Digitalized Clinical Study Protocol in COVID-19: Process Evaluation
JMIR Form Res 2022;6(7):e37832
URL: https://formative.jmir.org/2022/7/e37832
doi:10.2196/37832
PMID:35852933

Original Paper

Changes in Resilience Following Engagement With a Virtual Mental Health System: Real-world Observational Study

Grant Graziani[1], PhD; Brandon S Aylward[1], PhD; Sarah Kunkle[1], MPH; Emily Shih[1], PhD

Ginger, San Francisco, CA, United States

Corresponding Author:
Grant Graziani, PhD
Ginger
116 New Montgomery St. Suite 500
San Francisco, CA, 94105
United States
Phone: 1 855 446 4374
Email: ggraziani@ginger.io

Abstract

Background: Digital health services can serve as scalable solutions to address the growing demand for mental health care. However, more research is needed to better understand the association between engagement with care and improvements in subclinical outcomes.

Objective: This study aims to fill this research gap by examining the relationship between members' engagement with the Ginger platform and changes in their psychological resilience.

Methods: We conducted a retrospective observational study of 3272 members who accessed Ginger, an on-demand mental health service, between January 2021 and November 2021. Each member completed the 10-item Connor-Davidson Resilience Scale questionnaire, a measure of psychological resilience, at baseline and again during a 6- to 16-week follow-up window. Depression and anxiety symptoms (9-item Patient Health Questionnaire and 7-item Generalized Anxiety Disorder) were also measured. Linear regression was used to identify the association between engagement with Ginger's multiple care modalities and changes in resilience. Moderator analysis was conducted to test whether clinical depression or anxiety at baseline moderated the relationship between engagement level and changes in resilience.

Results: Of the 3272 members, 2683 (82%) reported low resilience at baseline. The mean change in resilience was 0.77 (SD 5.50) points. Linear regression models showed that age and census region did not predict changes in resilience; however, male members showed larger improvements (coefficient=0.58; P=.04). Baseline mental health outcomes, including resilience and depression and anxiety symptoms, were strong predictors of changes in resilience. Every point decrease in baseline resilience is associated with a 0.28-point increase in change in resilience (P<.001), and members with no or mild depression and anxiety at baseline saw changes in resilience that were 1.44 points (P<.001) larger than their clinical counterparts. Engagement with the Ginger system predicted changes in resilience. Members who engaged with Ginger coaching, clinical services, or both improved their resilience by 1.82, 1.55, and 1.40 points, respectively (P<.001), more than those who only engaged with Ginger content. Screening negative for moderate to severe depression and anxiety at baseline was associated with larger improvements in resilience (coefficient=1.30; P<.001); however, subclinical status was not shown to be a moderator for the association between level of engagement and changes in resilience.

Conclusions: Engagement with Ginger services was associated with improvements in resilience. Members who engaged in coaching or clinical care had significantly larger improvements compared with those who only engaged in self-guided content, regardless of whether a member screened positive for clinical depression or anxiety at baseline.

(JMIR Form Res 2022;6(7):e37169) doi:10.2196/37169

KEYWORDS

behavioral coaching; psychological resilience; mental health; telehealth

Introduction

Background

Resilience is a multidimensional construct that may be viewed as one's ability to cope with stress or represent personal qualities that enable one to thrive in the face of adversity [1-3]. Resilience has been studied in a variety of disciplines, including psychology, psychiatry, sociology, genetics, and neuroscience among others [4,5]. The definition of resilience has evolved along with increased scientific knowledge in these disciplines. As such, researchers have argued that resilience is "the process of multiple biological, psychological, social, and ecological systems interacting in ways that help individuals to regain, sustain, or improve their mental wellbeing when challenged by one or more risk factors" [6]. Furthermore, research has shown that resilience can vary with factors such as time, context, gender, and cultural origin [7]. From a strength-based perspective, resilience can contribute to positive functioning and help prevent negative emotions, thoughts, and behaviors [8]. Studies have demonstrated that resilience is a mitigating factor for heightened stress and adverse mental health outcomes in life circumstances.[9]. Specifically, individuals with low or normal levels of resilience have been found to be more likely to experience mental distress than those with high resilience [10,11].

Although many researchers' conceptualization of resilience is somewhat related to the notions mentioned earlier, namely, individuals' ability to cope with stress and adverse circumstances, there is no consensus on the operational definition of resilience according to a recent meta-analysis [12]. Given this, studies such as this one must be explicit about the conceptualization of resilience used as well as the precise measurement strategy. For the purpose of this study, psychological resilience is defined as how well one is able to adapt to change and *bounce back* after stressful events, tragedy, or trauma. This framework is consistent with the 10-item Connor-Davidson Resilience Scale (CD-RISC-10), which is the survey instrument used to measure resilience in this study. The 10 topics included in the CD-RISC-10 were confidence, determination, flexibility, focus, grit, perseverance, personal growth, positivity, self-reliance, and weathering emotions. This measure is explained in more detail in the *Data* section.

In light of the current pandemic, there has been increased focus on employee mental health and recognition of the importance of resilience in daily stress management and overall well-being [13]. For many, the COVID-19 pandemic fundamentally changed how and where one works as well as the daily demands of their jobs, and many reported increased loneliness, anxiety, depression, and suicidal ideation [14-16]. Recent studies have found that individuals with lower resilience scores are more likely to experience mental distress and express greater difficulty coping with the emotional challenges of the COVID-19 pandemic [10,16]. Similarly, a population-based study in China at the peak of the pandemic found that psychological resilience was significantly negatively correlated with depression, anxiety, and somatization symptom scores [17]. In addition, a study among public workers found that resilience mediated the effect of depression in public workers and their stress and anxiety levels during the pandemic [18].

Behavioral health coaching draws from several theoretical approaches that can effectively impact resilience or well-being. Coaching interventions address a variety of day-to-day challenges by identifying and working toward concrete and actionable goals. Resilience is seen as a proactive capability that supports the attainment of such goals and enhances overall mental health [19]. During the pandemic and postpandemic return to the office, many organizations have allocated increased resources toward the mental health and well-being of their employees, including resilience training and coaching-based interventions. However, a review of resilience intervention studies found that most in-person trainings were of short duration and had limited follow-up periods [13].

Previous studies have indicated a positive relationship between resilience and well-being, with higher resilience in the workplace setting associated with better mental health, reduced stress, and greater well-being [13]. Furthermore, the authors found that individuals who participated more often in the web-based resilience training program achieved the greatest improvements [13]. A recent multilevel meta-analysis found that resilience-promoting interventions yielded a small but significant overall effect on resilience [12]. One key finding was that ambiguity in conceptualizing and operationalizing resilience, in turn, leads to variability both between and within treatment effect sizes [12].

In addition to resilience-focused interventions, studies have found that general coaching interventions have also demonstrated that coaching can support resilience, even in the absence of it being the focus or aim of services. For example, a randomized controlled study of executives in a public health agency found that individual coaching sessions enhanced goal attainment, increased resilience and workplace well-being, and decreased depression and stress compared with controls [20]. Similarly, Lee et al [21] found that a health coaching program is an effective strategy for improving resiliency in youth.

Ginger offers various types of care designed to provide mental health support, including self-guided content, text-based behavioral health coaching, teletherapy, and telepsychiatry. Theoretically, each modality of care has a different effect on resilience. This hypothesis was tested in this study. Furthermore, given the clinical focus of therapy and psychiatry, these modalities may have different impacts on resilience depending on whether a member presents with clinical symptoms. It could be that these modalities are more effective at impacting resilience for these members if interventions designed to impact clinical symptoms are more impactful on resilience. Alternatively, if addressing clinical symptoms is the focus of care before addressing subclinical outcomes such as resilience, we could expect to see that therapy and psychiatry have a smaller impact on resilience (but perhaps have an equal or larger impact on a time horizon beyond the scope of this study). Given the unique Ginger context that offers multiple care modalities, testing whether clinical symptoms moderate the impact of engagement on resilience is possible. We are not aware of any existing studies that explicitly test this moderator hypothesis.

Overall, literature supports the relationship between resilience and other mental health and well-being outcomes and the fact that interventions, including coaching, can bolster resilience. Given that most of these studies have been conducted in controlled research settings, it is important to supplement this knowledge to better understand what is happening in real-world settings, particularly when a global pandemic introduces unique challenges to resilience.

Study Objectives

The purpose of this study was to examine changes in resilience among members seeking on-demand mental health treatment. We explicitly tested the following three hypotheses:

1. Change in resilience is associated with member characteristics at baseline, including demographic characteristics and baseline mental health outcomes (baseline resilience, depression symptoms, and anxiety symptoms).
2. Engagement with Ginger care is associated with larger improvements in resilience.
3. Baseline depression and anxiety symptoms moderate the association between engagement and changes in resilience.

Consistent with previous literature [10,11,13,17], we hypothesize that resilience will increase over the follow-up time points of treatment, and those with higher anxiety or depression symptom scores at intake will evidence smaller improvements or worsening in resilience over time.

Study Contributions

This study contributes to the literature on resilience in several ways. First, we present the results of one of the largest longitudinal studies of resilience. Our sample includes 3272 individuals. Second, to our knowledge, this is the first study to specifically examine resilience in the context of a digital mental health system that offers self-guided content, text-based behavioral health coaching, telepsychotherapy, and telepsychiatry. Third, by leveraging our rich data on Ginger members, we tested specific hypotheses that relate resilience to clinical depression and anxiety symptoms. In particular, we were able to test for the first time whether depression and anxiety symptoms moderate the impact of coaching and clinical interventions on resilience.

Methods

Overview

This was a retrospective observational study of Ginger members: individuals who joined Ginger, an on-demand mental health system. Data were collected between January 1, 2021, and November 13, 2021, from Ginger members residing in the United States. As part of its measurement-based care system, Ginger used the CD-RISC-10 as an indicator of resilience at intake as well as to track treatment progress beyond anxiety and depression symptom scores. By leveraging a retrospective design, this study contributes to the growing literature using real-world evidence. Although such studies often lack clear causal inference, they offer increased feasibility, larger samples, and robust external validity.

Participants

Study participants had access to the Ginger system as part of their employee or health plan benefits. Internal clinical protocols include the following exclusionary criteria, where self-directed telehealth is not likely appropriate and more specialized and urgent psychiatric services are required: (1) active suicidal ideation; (2) active high-risk self-harm behavior; (3) 2 or more hospitalizations within the past 6 months or 1 hospitalization in the past month for psychiatric reasons; (4) certain symptoms of psychosis that are poorly managed (eg, member is not medication compliant or symptoms are unresponsive to treatment) and are likely incompatible with telehealth; (5) a primary diagnosis of a substance use disorder or moderate to severe substance abuse issues, owing to the high complexity, severity, and risk frequently associated with such members, as well as the need for specialized care; (6) active eating disorders with symptoms considered high-risk; (7) ongoing grave disability, including certain patients who are bipolar with active mania or hypomania or mixed episodes who are unmedicated or have poor compliance with a medication regimen over time; and (8) two or more medical hospitalizations in the last month, owing to the high likelihood that the individual has a poorly controlled medical condition that requires close monitoring. For this study, we included Ginger users aged ≥18 years who downloaded the app during the data collection period.

Data

The Ginger System

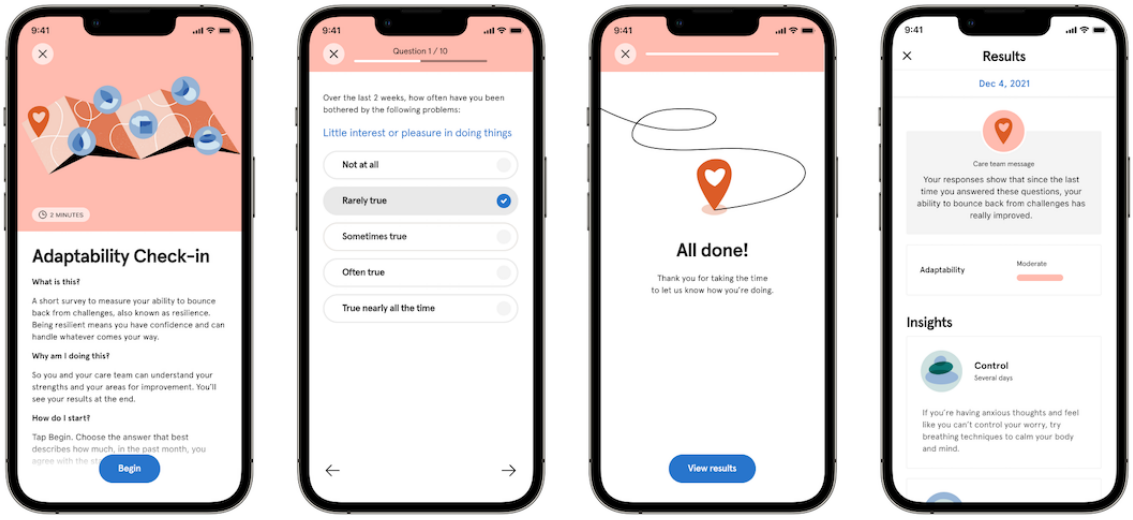
Ginger provides virtual on-demand mental health services, primarily through employee or health plan benefits. Using a mobile app platform, Ginger members can access text-based behavioral health coaching, teletherapy, and telepsychiatry, as well as self-guided content and assessments. Individuals who are eligible for Ginger can download the mobile app, complete an onboarding process, and begin texting with a behavioral health coach within minutes. Members who are interested in or have been determined to be in need of a higher level of care can meet with a therapist or psychiatrist via video. All participants had access to self-care activities via mobile apps. Additional details regarding the Ginger system can be found in prior publications evaluating depression and anxiety outcomes as measured by the 9-item Patient Health Questionnaire (PHQ-9) and 7-item Generalized Anxiety Disorder (GAD-7) surveys [22,23].

Data Collection

Ginger uses various assessments including the PHQ-9 and GAD-7 surveys as part of its measurement-based care system. Since December 2020, Ginger has used the CD-RISC-10 survey (referred to as an adaptability check-in within the app) to track progress beyond depression and anxiety symptom scores. This is particularly relevant to understand the needs of *subclinical* members (ie, members who do not exhibit clinically significant levels of depression or anxiety at intake). The CD-RISC-10 was selected because of its focus on behavioral health coaching to build resilience and its strength-based focus. A total of 7 CD-RISC-10 surveys were sent to members 1 week after enrollment, and follow-up surveys were sent to members every

30 days. Importantly, members who signed up but did not engage with the app past the 1-week mark did not complete the baseline survey. In this way, members with a low likelihood of meaningful engagement (a proxy for behavioral health needs) were excluded from the sample. A visual depiction of how the CD-RISC-10 survey appears to the members is shown in Figure 1.

Figure 1. The 10-item Connor-Davidson Resilience Scale survey in the Ginger mobile app.



Measures

CD-RISC-10 Assessment Tool

A common assessment tool for resilience is the CD-RISC-10. As mentioned earlier, Ginger uses the CD-RISC-10 as a proxy measure of an individual's psychological resilience level. The original researchers initially developed a 25-item scale to measure resilience or how well one is able to adapt to change and *bounce back* after stressful events, tragedy, or trauma. A total of 2 brief versions, the 10-item (CD-RISC-10) [24] and the 2-item (CD-RISC-2) [25], were subsequently developed by other research teams. The CD-RISC-10 contains 10 of the original 25 items from the CD-RISC scale and has demonstrated robust validity, reliability, and practicality [24]. The 10 topics included in the CD-RISC-10 were confidence, determination, flexibility, focus, grit, perseverance, personal growth, positivity, self-reliance, and weathering emotions. For each of the 10 items, respondents were asked to rate items on a 5-point scale: not true at all (0), rarely true (1), sometimes true (2), often true (3), and true nearly all the time (4). A respondent's total score ranges from 0 to 40, with higher scores indicating greater resilience. Results from the US population indicate that the quartiles for this measure are as follows: Q1: 0 to 29, Q2: 30 to 32, Q3: 33 to 36 and Q4: 37 to 40 [22].

PHQ-9 Assessments

The PHQ-9 is a 9-item self-report questionnaire that assesses the frequency and severity of depression symptomatology within the previous 2 weeks. Each of the 9 items is based on the Diagnostic and Statistical Manual of Mental Disorders, 4th edition criteria for major depressive disorder and is scored on a 0 (not at all) to 3 (nearly every day) scale. Items include *Little interest or pleasure in doing things* and *Feeling down, depressed, or hopeless.* Total scores range from 0 to 27, with higher scores indicating more depressive symptoms. A score of 10 is used as the clinical threshold [26].

GAD-7 Assessments

The GAD-7 is a valid, brief self-report tool used to assess the frequency and severity of anxious thoughts and behaviors over the past 2 weeks. Each of the 7 items is based on the Diagnostic and Statistical Manual of Mental Disorders, 4th edition diagnostic criteria for generalized anxiety disorder and is scored on a 0 (not at all) to 3 (nearly every day) scale, with total scores ranging from 0 to 21. Items include *Feeling nervous, anxious, or on edge* and *Not being able to stop or control worrying*. Consistent with existing literature [27], a score of 10 was used as the clinical threshold for this study.

Levels of Engagement

Coaching sessions were operationalized as the number of unique days on which members and coaches each exchanged at least five text messages, the minimum we believe is needed to capture a productive conversation between members and their coaches. Clinical sessions were operationalized as the number of video sessions completed with a clinician.

For this study, 5 different levels of engagement were considered based on members' engagement with self-guided content, text-based coaching, and teletherapy sessions. Specifically, members who engaged only with self-guided content and did not complete any coaching or clinical sessions were categorized as the *self-guided* group. *Low engagement* was defined as members who completed >4 coaching or >4 teletherapy sessions. Members who completed ≥4 text-based coaching sessions and

no clinical sessions were categorized as the *coaching only* group. Members who completed 4 or more clinical teletherapy sessions (with a therapist) and at most one coaching session were categorized as the *clinical only* group. We allow for *clinical only* members to have completed at most one coaching session, given that many members' point of entry to Ginger is a coaching session, after which they could be escalated to therapy and not continue coaching. Finally, members who completed 4 or more coaching sessions coupled with at least one clinical session or members who completed 4 or more teletherapy sessions coupled with more than 1 coaching session were categorized as the *hybrid care* group. The descriptions and rationales for the creation of these groups are described in Table 1.

Table 1. Engagement levels.

Engagement level	Definition	Rationale
Self-guided	Engagement only with self-guided content; 0 coaching and 0 clinical sessions	These are members who have engaged with the app's self-guided content but have not interacted with any coaching or clinical care. This group serves as the primary reference group in the linear regression models.
Low engagement	Between 0 and 3 total sessions comprising coaching or clinical care	Internal analyses suggest that 4 sessions are an inflection point for meaningful symptom reduction. These are members who have not reached this threshold.
Coaching only	≥4 coaching sessions and 0 clinical sessions	These are members who have completed at least the internally established threshold of 4 sessions but exclusively with coaching care.
Clinical only	≥4 clinical sessions and ≤1 coaching session	These are members who have completed at least the internally established threshold of 4 sessions but exclusively with clinical care.
Hybrid care	>1 coaching session and ≥4 clinical sessions or ≥1 clinical session and ≥4 coaching sessions	These are members who have completed more than the internally established threshold of 4 sessions using a combination of coaching and clinical care.

Baseline Characteristics

For each member, the following data were either collected at baseline or were fixed characteristics of members: age group, gender, geographic region, PHQ-9 score, and GAD-7 score. The demographic and location data were not self-reported. Instead, they were reported by a member's parent organization, which is either their employer or health insurance plan. Baseline PHQ-9 and GAD-7 data were collected using the Ginger system. The baseline PHQ-9 and GAD-7 scores were selected within 1 week before and after a member's baseline CD-RISC-10 score was collected, and the first PHQ-9 and GAD-7 scores were chosen. Members without baseline PHQ-9 and GAD-7 scores were excluded from analysis.

For many of our participants, the baseline characteristics were missing. The data were missing owing to 1 of 2 reasons. First, a member's parent organization may not share members' demographic information. Thus, missing demographic data are a signal of a member's parent organization and not necessarily a signal of information specific to a given member. For example, of the 197 parent organizations represented in this study, 118 (59.9%) reported all their members' gender information, 76 (38.6%) reported no gender information, and the remaining 3 (1.5%) organizations reported gender information for some but not all of their members.

Analyses

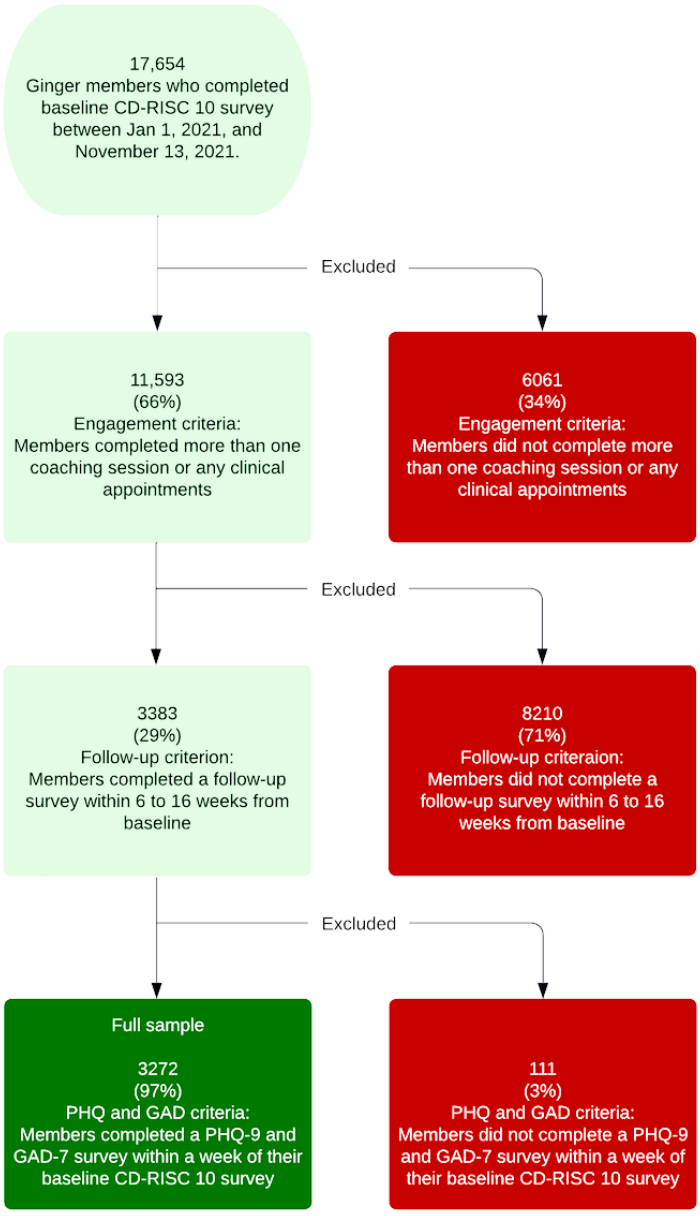
Sample

The sample for this study included Ginger members residing in the United States who completed a baseline survey between January 1, 2021, and November 13, 2021. Members were excluded from the analysis if they satisfied any of the following criteria:

1. Engagement criterion: Members completed more than 1 coaching session or any number of clinical appointments before their baseline resilience scores. The members' baseline survey was sent after their first coaching session.
2. Follow-up criterion: A member did not have a follow-up resilience score between 6 and 16 weeks from baseline.
3. PHQ-9 and GAD-7 criteria: Members without valid PHQ-9 and GAD-7 scores within a week of their baseline resilience score.

There were 17,654 members in the baseline resilience survey, of whom 6061 (34.33%) were excluded for meeting the engagement exclusion criteria. Of the remaining 11,593 participants, 3383 (29.18%) completed a follow-up survey between 6 and 16 weeks from their baseline survey. Of these, 3.28% (111/3383) did not have a valid PHQ-9 or GAD-7 score and were thus excluded. The resulting 3272 members comprised the full sample for this study and were used for our descriptive analysis. Of these 3272 members, 2674 (81.72%) had a low baseline score (CD-RISC-10 score<30) and comprised the low-resilience subsample for the study [24]. Given that the Ginger intervention is intended to improve resilience among members with low resilience at baseline, this subsample was used to analyze the association between engagement level and changes in resilience. Figure 2 outlines the sample construction process.

Figure 2. Sample flowchart. CD-RISC-10: 10-item Connor-Davidson Resilience Scale; GAD-7: 7-item Generalized Anxiety Disorder; PHQ-9: 9-item Patient Health Questionnaire.



Summary Statistics and Subgroup Analysis

Our descriptive analysis summarized the changes in resilience scores by presenting the mean, median, and SD of these change scores. To analyze differences across subgroups of members, we used a 2-tailed Welch 2-tailed *t* test for differences in mean changes across groups with unequal variances when a category had 2 groups (eg, gender). For categories with more than 2 groups (eg, census regions), we used an *F* test as part of an ANOVA to test for significant differences in mean changes across the groups. Furthermore, to understand whether members with missing data had significantly different outcomes than those without missing data, we performed 2-tailed Welch *t* tests to compare mean changes across the missing and nonmissing groups.

The subgroups of focus in this study were based on baseline resilience, depression, and anxiety symptom scores. In particular, members are grouped by their presence at baseline with low resilience (CD-RISC-10 score<30) and moderate to severe depression or anxiety (ie, PHQ-9 or GAD-7 score≥10) [26-28].

Descriptive Multivariate Regressions

To understand the association between the demographic and baseline survey responses and changes in resilience, we estimated a multivariate linear ordinary least squares (OLS) regression. The dependent variable for this model was the change in resilience scores. The following categorical independent variables were included in the model: gender, age group, census region, and interacted indicators for whether a member's baseline PHQ-9 or GAD-7 score was ≥10. For depression and anxiety at baseline, interacted indicators were

included in the regression model to account for possible differences based on combinations of depression and anxiety clinical status. Given that missing demographic data are highly dependent on whether a member reports such data for any of their members, indicators for a member's parent organization were included as independent variables. In addition, given the relatively wide follow-up period (6-16 weeks), indicator variables for the number of weeks between a member's baseline and follow-up scores were included to account for secular time trends. The coefficients of the indicators for parent organization and weeks between scores have not been reported. For each of these independent variables, a category of members with missing data was included. Homoscedasticity was not assumed, and robust SEs were computed.

Moderator Analysis

To understand the association between engagement with Ginger coaching and changes in resilience scores, we leveraged a moderator model with baseline depression and anxiety clinical status as the moderator and the level of engagement category (eg, self-guided, coaching only, and clinical only) as the independent variable. For our moderator categorization, members with either moderate to severe depression or anxiety at baseline were included in the clinical group (ie, PHQ-9 or GAD-7 score≥10), whereas all other members were included in the subclinical group. Clinical status at baseline was the hypothesized moderator of the association between the level of engagement and changes in resilience. We present the mean changes in resilience according to the clinical status for each engagement level. This analysis was restricted to members with low resilience at baseline (ie, CD-RISC-10 score<30).

To formally test whether clinical status at baseline was a moderator for engagement, we used a multivariate OLS regression model that included an indicator for engagement level interacting with an indicator of clinical status at baseline.

Ethics Approval

This study represents a secondary analysis of pre-existing deidentified data. The study team did not have access to participants or information to identify participants and did not intend to recontact participants. This study protocol was reviewed by Advarra institutional review board and determined to be exempt from institutional review board oversight, as deidentified secondary data analysis is generally not regarded as human subject research.

Results

Summary Statistics and Subgroup Analysis

Figures 3 and 4 show the distribution of baseline resilience and changes in resilience, respectively, for the full sample of 3272 members. Baseline resilience was centered at 24 out of 40 points (mean 23.83, SD 6.47; median 24). The distribution is similar to a normal distribution; however, there is an excess mass toward the upper limit of the distribution and a relatively long left tail. We did not observe any evidence of scores being concentrated at any particular part of the distribution. Change in resilience was centered at 1 out of 40 points (mean 0.77, SD 5.50; median 1). The distribution was roughly normal, with a small number of outliers at either end of the distribution. All subsequent analyses were conducted including these outliers and excluding members below the fifth percentile for baseline resilience, below the fifth percentile for changes in resilience, and above the fifth percentile for changes in resilience. All the results were robust to excluding these outliers. For the sake of transparency, we presented the results inclusive of outliers.

Table 2 presents the number of members and statistics for members' baseline resilience scores and score changes at follow-up for the overall sample and subgroups based on demographic characteristics and mental health outcomes at baseline.

Figure 3. Distribution of baseline resilience (full sample). N=3272; mean=23.83; SD=6.47; median=24.

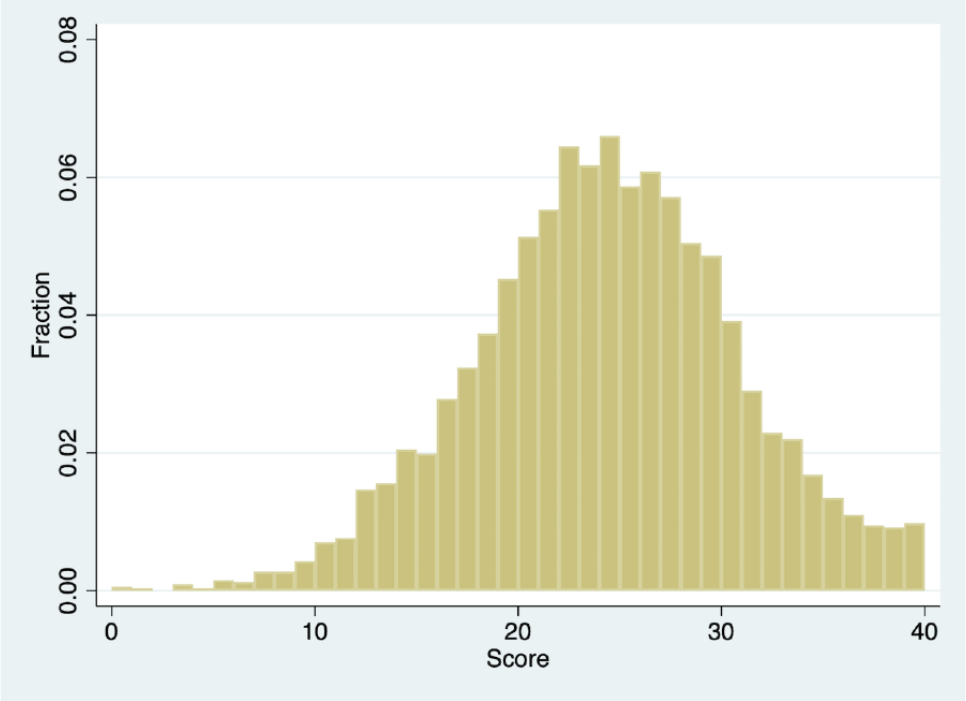


Figure 4. Distribution of changes in resilience at follow-up (full sample). N=3272; mean=.77; SD=5.5; median=1.

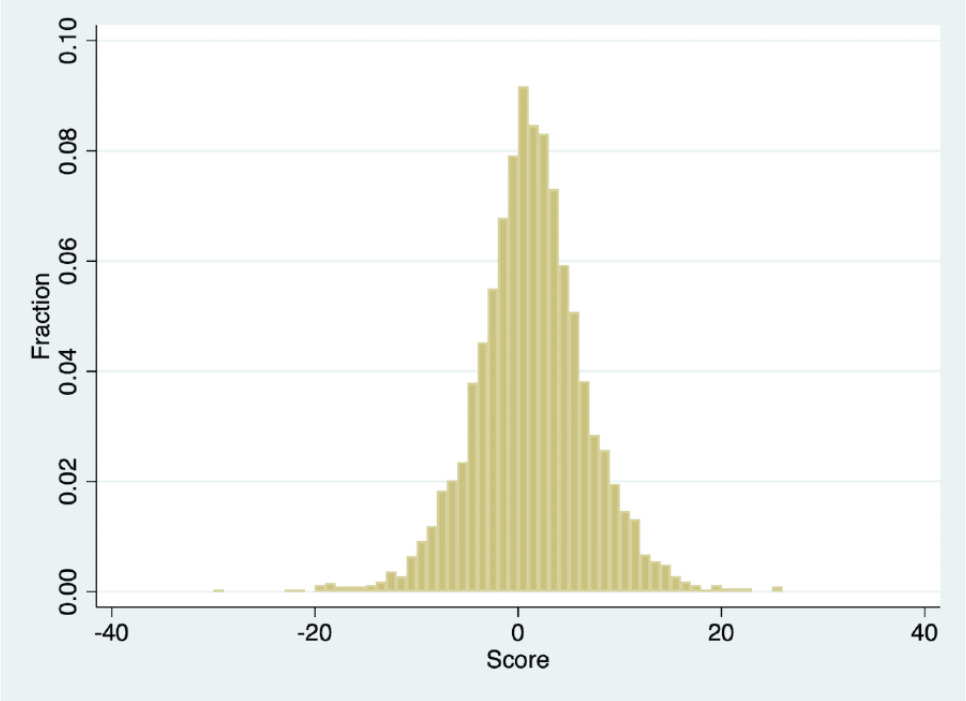


Table 2. Resilience score characteristics by overall sample and subgroups group.

Characteristics	Participants, n (%)	Resilience score		Resilience change	
		Mean (SD)	*P* value	Mean (SD)	*P* value
All	3272 (100)	23.83 (6.47)	—[a]	0.77 (5.50)	—
Gender			.19		.14
Female	1377 (42.08)	23.95 (6.49)		0.67 (5.49)	
Male	569 (17.38)	24.39 (6.84)		1.07 (5.56)	
Missing gender	1326 (40.52)	23.45 (6.26)	.005	0.74 (5.49)	.83
Age (years)			.01		.34
18 to 24	176 (5.37)	22.84 (5.94)		0.27 (5.16)	
25 to 34	844 (25.79)	23.76 (6.10)		0.54 (5.23)	
35 to 44	547 (16.71)	24.54 (6.76)		0.94 (5.37)	
45 to 64	461 (14.08)	24.29 (6.49)		0.89 (5.58)	
≥65	37 (1.13)	25.03 (6.49)		0.14 (5.77)	
Missing age	1207 (36.88)	23.48 (6.61)	.02	0.91 (5.75)	.28
Region			.54		.40
West	1052 (32.15)	23.69 (6.06)		0.64 (5.22)	
Midwest	290 (8.86)	23.93 (6.25)		0.80 (5.55)	
South	994 (30.37)	23.94 (6.83)		1.03 (5.81)	
Northeast	397 (12.13)	24.23 (6.46)		0.65 (5.33)	
Missing region	539 (16.47)	23.53 (6.69)	.26	0.60 (5.56)	.45
Baseline resilience			<.001		<.001
High resilience (CD-RISC-10[b]≥30)	598 (18.27)	33.15 (2.84)		–1.54 (5.11)	
Low resilience (CD-RISC-10<30)	2674 (81.72)	21.74 (5.06)		1.28 (5.45)	
Baseline depressive symptoms			<.001		.31
PHQ-9[c] score≥10	1477 (45.14)	21.71 (6.38)		0.66 (5.78)	
PHQ-9 score<10	1795 (54.85)	25.56 (6.01)		0.86 (5.26)	
Baseline anxiety symptoms			<.001		.55
GAD-7[d] score≥10	1327 (40.55)	21.70 (6.44)		0.70 (5.75)	
GAD-7 score<10	1945 (59.44)	25.28 (6.07)		0.82 (5.33)	
Engagement level			<.001		<.001
Self-guided	499 (15.25)	24.33 (6.66)		–0.11 (5.64)	
Low engagement	989 (30.22)	23.87 (6.64)		0.52 (5.60)	
Coaching only	544 (16.62)	24.61 (6.36)		1.26 (5.03)	
Clinical only	670 (20.47)	23.45 (6.29)		1.06 (5.35)	
Hybrid care	570 (17.42)	22.99 (6.18)		1.15 (5.70)	

[a]*P* values are reported only for testing differences in baseline resilience across categories of subgroups.

[b]CD-RISC-10: 10-item Connor-Davidson Resilience Scale.

[c]PHQ-9: 9-item Patient Health Questionnaire.

[d]GAD-7: 7-item Generalized Anxiety Disorder.

Demographics

Demographic data were missing for a large portion of the sample because of irregular reporting by members' employers or health plans. Of those without missing demographic data, most participants were female (1377/1946, 70.76%) and aged ≥35 years (1045/2065, 50.61%). Members were most likely to live in the West (1052/2733, 38.49%) and South (994/2733,

36.37%); however, all 4 census regions were represented in the baseline sample.

The baseline statistics and changes at follow-up are presented in Table 2. Columns 3 to 4 correspond to the baseline scores, and columns 5 to 6 correspond to the changes at follow-up. For each category (gender, age, etc), a *P* value is presented in the category's first row to test whether the difference in mean changes in scores across the category was statistically significant. In the row for groups with missing data, the *P* value corresponds to a 2-tailed *t* test of the difference in mean baseline scores between those with and without missing data.

For categories based on gender and census region, neither the mean baseline resilience score nor the mean changes at follow-up were statistically different across groups (all *P*>.05). The mean baseline resilience score across age groups was significantly different, with older members having higher baseline resilience scores. Differences in the mean change in resilience across age groups were not statistically significant. However, the mean baseline scores for members with missing gender and age data were significantly different from those without missing data. This pattern did not hold for mean changes at follow-up (ie, members with missing demographic data did not have significantly different mean changes at follow-up than those without missing data).

Mental Health Outcomes

The vast majority (2674/3272, 81.72%) of members reported low resilience at baseline (ie, CD-RISC-10 score<30). On the basis of the PHQ-9 and GAD-7 scores at baseline, 45.14% (1477/3272) of the members screened positive for clinical depression at baseline (PHQ-9 score≥10) and 40.55% (1327/3272) for clinical anxiety at baseline (GAD-7 score≥10). Consistent with prior work [10,11], the differences in mean baseline scores between the clinical and nonclinical groups were statistically significant. Specifically, the mean baseline scores for members who screened positive for either clinical depression (mean 21.71, SD 6.38) or anxiety (mean 21.7, SD 6.44) were significantly lower than for those who screened negative for depression (mean 25.56, SD 6.01) or anxiety (mean 25.28, SD 6.07). Members with subclinical depression and anxiety at baseline demonstrated an average resilience score improvement of 0.86 and 0.82 points, respectively; however, changes in resilience between those with clinical and nonclinical symptom scores was nonsignificant.

By construction, members with low baseline resilience scored below those with high resilience (mean 21.74 vs 33.15). On an average, members with low resilience at baseline demonstrated an increase of 1.28 points at follow-up. Conversely, members with high resilience at baseline evidenced decreasing scores (−1.54 points on average, SD 5.11) at follow-up. The difference in mean resilience score changes between these groups was significant (*P*<.001).

Engagement Level

The most common engagement level was low engagement, with 30.22% (989/3272) of members meeting the criteria. There were 15.25% (499/3272) of members in the *self-guided* group, 16.62% (544/3272) of members in the *coaching only* group, 20.47% (670/3272) of members in the *clinical only* group, and 17.42% (570/3272) of members in the *hybrid care* group. Both baseline resilience and changes in resilience differed significantly across the groups based on engagement levels (both *P*<.001). Specifically, members in the *coaching only* group had the highest mean baseline resilience scores (mean 24.61, SD 6.36) and the largest mean change in resilience (mean 1.26, SD 5.03). The *hybrid care* group had the lowest mean baseline resilience (mean 22.99, SD 6.18), and the *self-guided* group had the lowest mean change in resilience (mean −0.11, SD 5.64).

Table 3 presents the results of the multivariate OLS regressions predicting changes in resilience at follow-up.

Gender predicted changes in resilience scores when controlling for all baseline characteristics; male participants (mean 1.07, SD 5.56) had significantly larger mean improvements in resilience scores than females (mean 0.67, SD 5.49; coefficient=0.58; *P*=.04). Although Table 2 shows that the mean change in scores for members with missing gender data was significantly different from those without missing data (not controlling for other variables, importantly, a member's parent organization), Table 3 shows that these members are not associated with significantly different changes in scores. Similarly, age and census region did not predict changes in resilience scores when controlling for other variables.

Baseline resilience score was a strong predictor of changes in resilience scores at follow-up. Specifically, controlling for other variables, for each 1-point increase in baseline resilience score, the follow-up score decreased by 0.28 points, which was statistically significant at the 1% level. Baseline depression and anxiety were also strong predictors of changes at follow-up. Specifically, members without clinical depression or anxiety at baseline had mean resilience improvements of 1.44 points more than members with both clinical depression and anxiety (*P*<.001).

Table 3. Ordinary least squares regression of resilience change scores.

	β (95% CI)	*P* value
Gender		
Female	Reference	Reference
Male	0.58 (0.02 to 1.14)	.04
Missing gender	–1.44 (–6.44 to 3.56)	.57
Age (years)		
18 to 24	Reference	Reference
25 to 34	0.10 (–0.74 to 0.95)	.81
35 to 44	0.55 (–0.36 to 1.46)	.24
45 to 64	0.59 (–0.36 to 1.53)	.22
≥65	0.28 (–1.78 to 2.35)	.79
Missing age	0.22 (–2.46 to 2.89)	.87
Region		
West	Reference	Reference
Midwest	0.08 (–0.91 to 1.07)	.88
South	0.55 (–0.12 to 1.21)	.11
Northeast	–0.30 (–1.12 to 0.52)	.47
Missing Region	–0.30 (–1.53 to 0.94)	.64
Baseline resilience score	–0.28[b] (–0.31 to –0.25)	<.001
Baseline depressive and anxiety symptoms		
PHQ-9[a] score≥10; GAD-7[b] score≥10	Reference	Reference
PHQ-9 score≥10; GAD-7 score<10	0.07 (–0.57 to 0.71)	.83
PHQ-9 score<10; GAD-7 score ≥10	0.37 (–0.37 to 1.10)	.33
PHQ-9 score<10; GAD-7 score<10	1.44 (0.96 to 1.93)	<.001
R-squared	0.1568234 (—[c])	—
Adjusted R-squared	0.0972076 (—)	—
Observations	3272 (—)	—

[a]PHQ-9: 9-item Patient Health Questionnaire.

[b]GAD-7: 7-item Generalized Anxiety Disorder.

[c]Not available.

Engagement Level and Moderator Analysis

Figure 5 presents mean changes in resilience by the level of engagement received by a member between their baseline and follow-up scores. These results were restricted to members with low baseline resilience. Multimedia Appendix 1 presents similar results for the full sample. Members with self-guided engagement did not evidence significant improvements in mean resilience scores, based on a 5% significance level. Members with low engagement had significant improvements in resilience scores by approximately 1.0 points, whereas members with meaningful engagement (coaching only, clinical only, or hybrid care) had statistically significant improvements between 1.5 and 2.0 points between baseline and follow-up.

Table 4 reports the corresponding mean changes from a multivariate regression with the same set of independent variables included in the previous section. Table 4 includes members with low baseline resilience. Multimedia Appendix 2 Table S1 presents regression results for the full sample. Members with self-guided services had the smallest changes among all members. Members with low engagement saw changes in resilience scores that are 0.91 points larger than those who completed self-guided services (P=.01), whereas members engaged with coaching only, clinical only, or hybrid care services saw changes that are 1.82, 1.55, and 1.40 points larger than the self-guided group, respectively (all P<.001). Given the 95% CIs for the estimated coefficients on the indicators for low engagement, coaching only, clinical only, and hybrid care engagement levels, we cannot reject the null hypothesis at the 5% level that changes in resilience are the same across these levels of engagement.

Figure 6 presents the same mean changes as Figure 5 but separately for subclinical members (ie, GAD-7 and PHQ-9 scores<10) and clinical members (ie, GAD-7 or PHQ-9 scores≥10). Figure 6 includes members with low baseline resilience. Multimedia Appendix 3 presents similar results for the full sample. Mean changes in resilience for clinical members were smaller than their subclinical counterparts across all engagement levels, except clinical only, where clinical and subclinical members saw similar improvements. On the basis of the 95% CIs around these means, the differences between clinical and subclinical members' resilience score changes were not statistically different within each engagement level.

Table 5 presents the results from an interacted moderator model to formally test whether subclinical status at baseline is a moderator for the association between engagement level and changes in resilience. Table 5 includes members with low baseline resilience. Multimedia Appendix 2 Table S2 presents regression results for the full sample. Interacted terms refer to the interaction between the multinomial engagement level variable, which takes 1 of 5 values depending on a member's engagement and an indicator for members who screen negative for both depression and anxiety at baseline. The coefficients of the indicators for engagement level were similar to the model without interactions, indicating that engagement level predicted changes in resilience for clinical members. The coefficient of the indicator for subclinical status interacted with the indicator for self-guided engagement was positive and statistically significant (coefficient=1.39; *P*=.02), indicating that subclinical status is associated with larger improvements in resilience for that engagement level. The coefficients of the regressors interacting engagement level with subclinical status were not significant, indicating that subclinical members' improvements in resilience did not vary by engagement level. This indicates that subclinical members improved by roughly 1.4 points more than their clinical counterparts, regardless of their engagement level, thus rejecting the hypothesis that subclinical status is a moderator of the association between engagement level and changes in resilience.

Figure 5. Change in resilience by engagement level (low resilience sample).

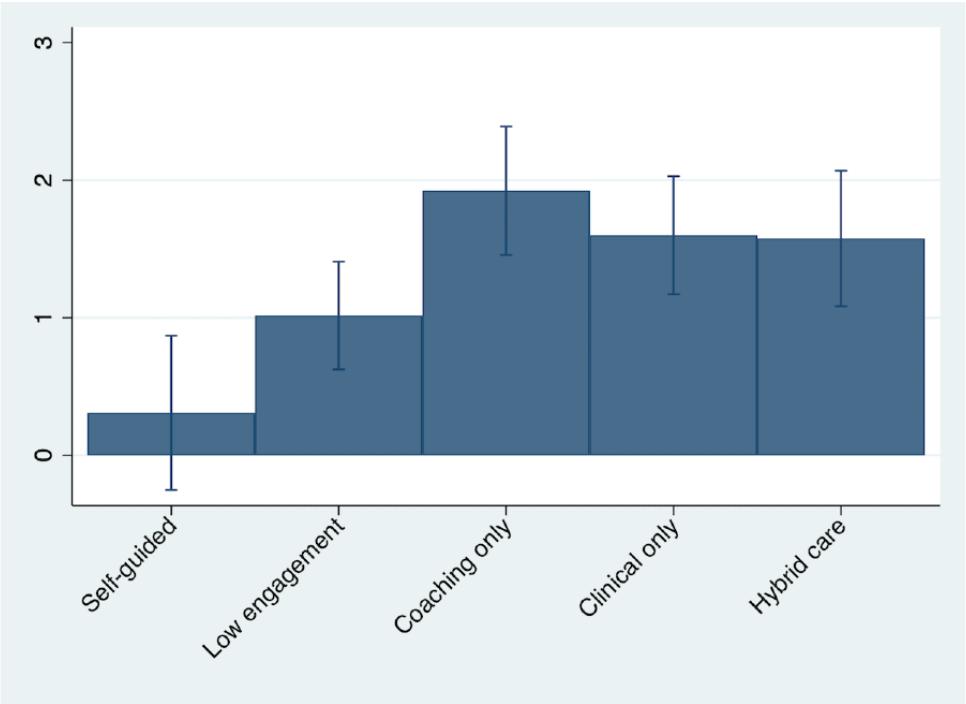


Table 4. Ordinary least squares regression of changes in resilience scores (low resilience sample).

	β (95% CI)	*P* value
Engagement level		
Self-guided	Reference	Reference
Low engagement	0.91 (0.20-1.63)	.01
Coaching only	1.82 (1.05-2.59)	<.001
Clinical only	1.55 (0.78-2.32)	<.001
Hybrid care	1.40 (0.61-2.19)	<.001
Subclinical, both	1.30 (0.85-1.75)	<.001
R-squared	0.1357372 (—[a])	—
Adjusted R-squared	0.0658413 (—)	—
Observations	2674 (—)	—

[a]Not available.

Figure 6. Change in resilience by engagement level and subclinical status (low resilience sample).

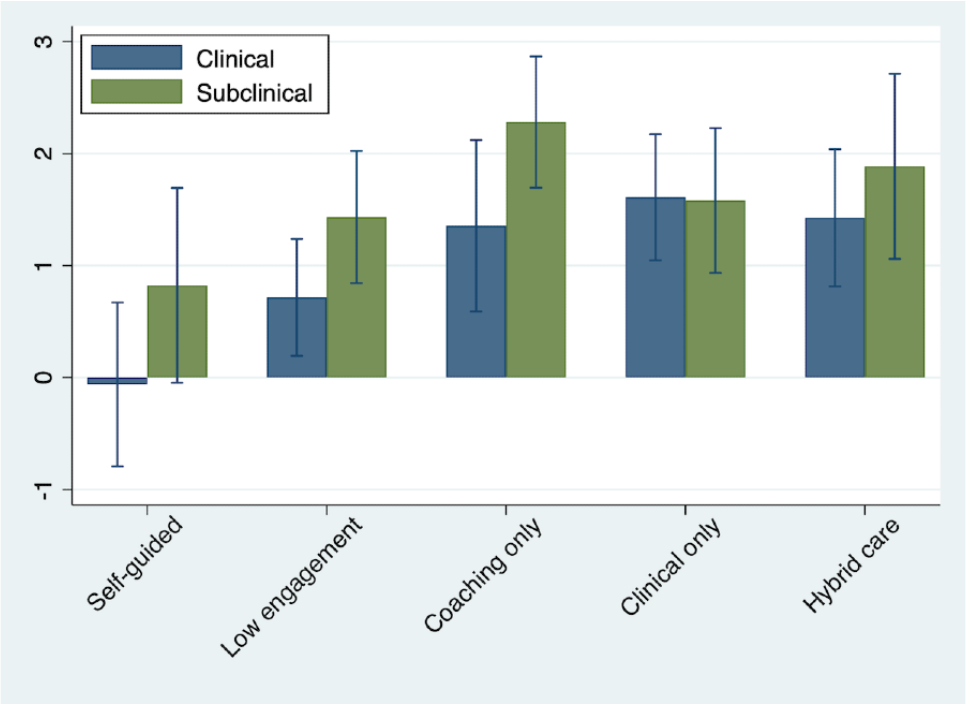


Table 5. Ordinary least squares regression of changes in resilience scores, interacted model (low resilience sample).

Engagement level	β (95% CI)	*P* value
Self-guided	Reference	Reference
Interacted with subclinical	1.39 (0.21 to 2.57)	.02
Low engagement	0.83 (–0.11 to 1.77)	.08
Interacted with subclinical	0.19 (–1.25 to 1.63)	.80
Coaching only	1.67 (0.55 to 2.79)	.003
Interacted with subclinical	0.21 (–1.33 to 1.76)	.79
Clinical only	1.74 (0.76 to 2.71)	<.001
Interacted with subclinical	−0.51 (–1.97 to 0.95)	.49
Hybrid care	1.53 (0.54 to 2.52)	.002
Interacted with subclinical	−0.39 (–1.98 to 1.20)	.63
R-squared	0.1364253 (—[a])	—
Adjusted R-squared	0.0650729 (—)	—
Observations	2674 (—)	—

[a]Not available.

Discussion

Principal Findings

The purpose of this study was to examine changes in resilience in members seeking on-demand mental health treatment as a function of both baseline symptoms of anxiety and depression, as well as engagement level. At baseline, most members (2674/3272, 81.72%) reported low resilience scores (ie, CD-RISC-10<30; mean 23.83, SD 6.47), which was well below the benchmarks of the US general population [29]. Overall, members experienced an average improvement in resilience of 0.77 points with a large variance around the mean (SD 5.50). According to the results by engagement level, members in the coaching only group had the largest changes in mean resilience between baseline and follow-up (mean 1.26, SD 5.03), followed by those in the hybrid care group (mean 1.15, SD 5.03). Across all levels of engagement, members who did not present with moderate to severe depression or anxiety at baseline saw larger improvements than their clinical counterparts (a difference of 1.3 points; *P*<.001, when controlling for covariates). Given that the difference between clinical and subclinical members' improvements did not vary significantly across levels of engagement, we conclude that subclinical status is not a moderator for the association of levels of engagement with changes in resilience. That is, although subclinical members improved more than clinical members, the gap between these groups was stable across all levels of engagement.

In subgroup analyses, members with low resilience at baseline, on average, demonstrated a 1.28-point improvement in resilience scores at follow-up. Resilience has been conceptualized as a dynamic process, and engagement with coaching or teletherapy may contribute to improvements in resilience, even in the absence of it being the focus of services [4,20,30]. Conversely, members with high resilience scores at baseline demonstrated a decrease in the CD-RISC-10 at follow-up (mean −1.54, SD 5.11). These findings could be a function of several factors, including regression to the mean or recalibration or self-discovery by members as part of treatment. Higher scores may leave less room for improvement. Prior research has shown that there is a ceiling effect in being able to detect improvement in individuals who self-report high resilience at baseline [31]. In addition, these findings could be the result of selection or attrition bias. Specifically, those with high resilience who continue to use treatment services may have a higher need for exogenous factors that impact resilience. On the basis of the results of the regression analyses, these findings could also be driven by members who reported high resilience (ie, CD-RISC-10 >30) but also presented with clinical depression or anxiety.

The construct of resilience can vary as a function of time, context, gender, and age [7]. However, our study did not find any statistically significant differences across demographic groups overall (ie, gender and location), with the exception of baseline resilience by age group. Specifically, those in the youngest age group demonstrated significantly lower baseline resilience (mean 22.84, SD 5.94) compared with other age groups. Resilience has been found to be greater in older adults, particularly those with emotional regulation ability and problem-solving ability [32]. When controlling for other baseline variables, our regression analysis found that male participants had a significantly larger mean change in resilience scores than female participants, suggesting that one must consider the socioecological context and treatment process when examining changes in resilience, including gender. Hirani et al [33] suggested that women typically score lower on measures of resilience because existing conceptualizations of resilience are not reflective of how gender roles, social expectations, perceptions, and environmental factors, among others, interact to differentially impact experiences for men and women and, in turn, their response to adversity.

We found that members with symptom scores within the clinical range (ie, PHQ-9 or GAD-7 score≥10) demonstrated

significantly lower baseline resilience scores. When controlling for other baseline characteristics, members without clinical depression or anxiety at baseline had improvements in resilience of 1.4 points more than members with both clinical depression and anxiety. Given the negative association between resilience and clinical symptom severity [10,11], ongoing symptomatology may negatively affect changes in resilience. Moreover, members with clinical symptomatology may focus on symptom management before transitioning to a focus on positive psychological outcomes such as resilience. One may expect that improvements in mental health symptom scores would parallel changes in resilience over the course of treatment. In addition, resilience-focused interventions may serve as a preventive approach to reduce the exacerbation of mental health symptoms that would meet the clinical threshold.

Regarding the association between care engagement level and both baseline and changes in resilience, there were significant differences in the mean baseline resilience and changes in resilience between those with and without meaningful engagement (ie, coaching only, clinical only, or hybrid care vs self-guided or minimal engagement). Specifically, members with self-guided engagement did not show any improvement in resilience, followed by those with minimal engagement having a 0.5-point improvement on average. Those with meaningful engagement (ie, coaching only, clinical only, or hybrid care) demonstrated improvements in resilience between 1.06 and 1.26 points. Within the Ginger context, these findings support the hypothesis that engagement with a human care provider (which is true for members in the coaching only, clinical only, and hybrid care groups) can lead to larger improvements in resilience than engagement with self-guided content alone. This is perhaps not surprising, given the more intensive nature of care delivered by a human provider. Behavioral health coaches, therapists, and psychiatrists tailor their care to the specific needs of a member, whereas it is incumbent on a member to find and engage with content that is applicable to their needs. This does not rule out that resilience-focused content could have an impact on resilience; however, given the wide variety of content available on the Ginger platform, this study was not designed to test whether resilience-specific content was associated with greater or lesser improvements in resilience than the care provided by a trained human provider. These findings are not surprising, as previous research has demonstrated that individuals who had greater engagement in a web-based resilience training program achieved the greatest improvements [10].

Controlling for baseline characteristics, members with subclinical status at baseline had larger improvements in resilience across all engagement levels; however, our moderation analysis found that the association between care engagement and changes in resilience did not significantly differ by baseline subclinical status.

Limitations

From a research perspective, there is much ambiguity in approaching the concept of resilience and no current consensus on the operational definition of resilience [12]. Moreover, the construct is sometimes assessed as a dynamic process and at other times as a personal trait or an outcome [4,13]. For example, some conceptualize resilience as the ability to *bounce back* from adversity, conflict, and failure [30]. Future research is needed to develop greater conceptual clarity around resilience, particularly as it relates to coaching interventions for depression and anxiety. This definition and conceptualization of resilience should fit the behavioral health coaching context [19]. There are many validated tools to assess resilience, in addition to the one used in this study (eg, the Resilience Scale for Adults, Psychological Capital Questionnaire, and Cognitive Hardiness Scale). Each of these tools reflects differing conceptualizations of resilience; thus, changes in resilience scores over time could capture different processes, with some being more static than others [19].

The construct of resilience can vary as a function of time, context, gender, age, and cultural origin [7]. A large percentage of the participants did not have age (1207/3272, 36.88%) or gender (1326/3272, 40.52%) reported and there was limited access to other demographic information of the study participants. Thus, we were unable to stratify analyses by key demographic factors or examine factors that could affect resiliency and treatment response (eg, marital status, family composition, significant life events, previous mental health treatment, sources of social support, and educational level). These analyses could provide additional insights into those who may best benefit from virtual care and those for whom additional support may be needed. In addition, because there was no comparison group in this retrospective observational study, we were unable to draw any causal inferences regarding the impact of the Ginger on-demand mental health platform.

Another limitation of this study is that baseline surveys for resilience, depression, and anxiety symptoms were conducted at different times. This could have led to measurement errors, as both types of outcomes could evolve before the other is measured. The lag between collecting CD-RISC-10 and PHQ-9 and GAD-7 (the latter 2 were collected at the same time) was intentional in an effort to avoid survey fatigue. Therefore, the results of this study should be interpreted with this lag in mind.

Conclusions

This study examined changes in resilience over time among members of an on-demand virtual mental health system. Overall, members with low baseline resilience and subclinical symptoms of anxiety and depression demonstrated the largest improvement in resilience over time. Even for interventions that did not focus on resilience, members could demonstrate improvements in resilience over the course of treatment with virtual-based treatment for anxiety and depression. Future studies should examine symptom scores of anxiety and depression over time in relation to resilience. In addition, the inclusion of measures such as perceived social support might provide additional insight into treatment, given its association with both mental health and resilience [4,34].

Conflicts of Interest

All authors are employees of Ginger.

Multimedia Appendix 1
Change in resilience by engagement level (full sample).
[PNG File , 471 KB - formative_v6i7e37169_app1.png]

Multimedia Appendix 2
Additional results using the full sample.
[DOCX File , 283 KB - formative_v6i7e37169_app2.docx]

Multimedia Appendix 3
Change in resilience by engagement level and subclinical status (full sample).
[PNG File , 557 KB - formative_v6i7e37169_app3.png]

References

1. Windle G. What is resilience? A review and concept analysis. Rev Clin Gerontol 2011;21(2):152-169. [doi: 10.1017/s0959259810000420]
2. Earvolino-Ramirez M. Resilience: a concept analysis. Nurs Forum 2007;42(2):73-82. [doi: 10.1111/j.1744-6198.2007.00070.x] [Medline: 17474940]
3. Wald J, Taylor S, Asmundson GJ, Jang K, Stapleton JA. Literature review of concepts: psychological resiliency. Defence R&D Canada – Toronto. 2006 Jul 1. URL: https://www.semanticscholar.org/paper/Literature-Review-of-Concepts%3A-Psychological-Wald-Taylor/dfc3fa7def8da2cb839f5944bb1ae8c11759a686 [accessed 2021-12-19]
4. Herrman H, Stewart DE, Diaz-Granados N, Berger EL, Jackson B, Yuen T. What is resilience? Can J Psychiatry 2011 May;56(5):258-265. [doi: 10.1177/070674371105600504] [Medline: 21586191]
5. Feldman R. What is resilience: an affiliative neuroscience approach. World Psychiatry 2020 Jun;19(2):132-150 [FREE Full text] [doi: 10.1002/wps.20729] [Medline: 32394561]
6. Ungar M, Theron L. Resilience and mental health: how multisystemic processes contribute to positive outcomes. Lancet Psychiatry 2020 May;7(5):441-448. [doi: 10.1016/S2215-0366(19)30434-1] [Medline: 31806473]
7. Hu T, Zhang D, Wang J. A meta-analysis of the trait resilience and mental health. Pers Individ Dif 2015 Apr;76:18-27. [doi: 10.1016/j.paid.2014.11.039]
8. Zadok-Gurman T, Jakobovich R, Dvash E, Zafrani K, Rolnik B, Ganz AB, et al. Effect of Inquiry-Based Stress Reduction (IBSR) intervention on well-being, resilience and burnout of teachers during the COVID-19 pandemic. Int J Environ Res Public Health 2021 Apr 01;18(7):3689 [FREE Full text] [doi: 10.3390/ijerph18073689] [Medline: 33916258]
9. Campbell-Sills L, Cohan SL, Stein MB. Relationship of resilience to personality, coping, and psychiatric symptoms in young adults. Behav Res Ther 2006 Apr;44(4):585-599. [doi: 10.1016/j.brat.2005.05.001] [Medline: 15998508]
10. Riehm KE, Brenneke SG, Adams LB, Gilan D, Lieb K, Kunzler AM, et al. Association between psychological resilience and changes in mental distress during the COVID-19 pandemic. J Affect Disord 2021 Mar 01;282:381-385 [FREE Full text] [doi: 10.1016/j.jad.2020.12.071] [Medline: 33421866]
11. Kermott CA, Johnson RE, Sood R, Jenkins SM, Sood A. Is higher resilience predictive of lower stress and better mental health among corporate executives? PLoS One 2019 Jun 11;14(6):e0218092 [FREE Full text] [doi: 10.1371/journal.pone.0218092] [Medline: 31185049]
12. Liu JJ, Ein N, Gervasio J, Battaion M, Reed M, Vickers K. Comprehensive meta-analysis of resilience interventions. Clin Psychol Rev 2020 Dec;82:101919. [doi: 10.1016/j.cpr.2020.101919] [Medline: 33045528]
13. Smith B, Shatté A, Perlman A, Siers M, Lynch WD. Improvements in resilience, stress, and somatic symptoms following online resilience training: a dose-response effect. J Occup Environ Med 2018 Jan;60(1):1-5 [FREE Full text] [doi: 10.1097/JOM.0000000000001142] [Medline: 28820863]
14. Pearman A, Hughes ML, Smith EL, Neupert SD. Age differences in risk and resilience factors in COVID-19-related stress. J Gerontol B Psychol Sci Soc Sci 2021 Jan 18;76(2):e38-e44 [FREE Full text] [doi: 10.1093/geronb/gbaa120] [Medline: 32745198]
15. Schuster C, Weitzman L, Sass Mikkelsen K, Meyer-Sahling J, Bersch K, Fukuyama F, et al. Responding to COVID-19 through surveys of public servants. Public Adm Rev 2020 May 27;80(5):792-796 [FREE Full text] [doi: 10.1111/puar.13246] [Medline: 32836447]
16. Killgore WD, 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]
17. Ran L, Wang W, Ai M, Kong Y, Chen J, Kuang L. Psychological resilience, depression, anxiety, and somatization symptoms in response to COVID-19: a study of the general population in China at the peak of its epidemic. Soc Sci Med 2020 Oct;262:113261 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113261] [Medline: 32758794]

18. Ju G, Lee J, Ahn MH, Lee J, Kim EJ, Suh S, et al. Effects of depression and resilience of public workers on work-related stress and anxiety in response to the COVID-19 pandemic. J Korean Med Sci 2021 Sep 13;36(36):e262 [FREE Full text] [doi: 10.3346/jkms.2021.36.e262] [Medline: 34519189]
19. Smith CL. Coaching for resilience and well-being. In: Bachkirova T, Spence G, Drake D, editors. The SAGE Handbook of Coaching. Thousand Oaks, CA, USA: Sage Publications; 2017:346-362.
20. Grant AM, Curtayne L, Burton G. Executive coaching enhances goal attainment, resilience and workplace well-being: a randomised controlled study. J Posit Psychol 2009 Sep;4(5):396-407. [doi: 10.1080/17439760902992456]
21. Lee JA, Heberlein E, Pyle E, Caughlan T, Rahaman D, Sabin M, et al. Evaluation of a resiliency focused health coaching intervention for middle school students: building resilience for healthy kids program. Am J Health Promot 2021 Mar;35(3):344-351. [doi: 10.1177/0890117120959152] [Medline: 32959670]
22. Kunkle S, Yip M, Ξ W, Hunt J. Evaluation of an on-demand mental health system for depression symptoms: retrospective observational study. J Med Internet Res 2020 Jun 18;22(6):e17902 [FREE Full text] [doi: 10.2196/17902] [Medline: 32554387]
23. Kunkle S, Yip M, Hunt J, Ξ W, Udall D, Arean P, et al. Association between care utilization and anxiety outcomes in an on-demand mental health system: retrospective observational study. JMIR Form Res 2021 Jan 26;5(1):e24662 [FREE Full text] [doi: 10.2196/24662] [Medline: 33496679]
24. Campbell-Sills L, Forde DR, Stein MB. Demographic and childhood environmental predictors of resilience in a community sample. J Psychiatr Res 2009 Aug;43(12):1007-1012. [doi: 10.1016/j.jpsychires.2009.01.013] [Medline: 19264325]
25. Vaishnavi S, Connor K, Davidson JR. An abbreviated version of the Connor-Davidson Resilience Scale (CD-RISC), the CD-RISC2: psychometric properties and applications in psychopharmacological trials. Psychiatry Res 2007 Aug 30;152(2-3):293-297 [FREE Full text] [doi: 10.1016/j.psychres.2007.01.006] [Medline: 17459488]
26. 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]
27. 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]
28. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. Gen Hosp Psychiatry 2016;39:24-31. [doi: 10.1016/j.genhosppsych.2015.11.005] [Medline: 26719105]
29. Wingo AP, Briscione M, Norrholm SD, Jovanovic T, McCullough SA, Skelton K, et al. Psychological resilience is associated with more intact social functioning in veterans with post-traumatic stress disorder and depression. Psychiatry Res 2017 Mar;249:206-211. [doi: 10.1016/j.psychres.2017.01.022] [Medline: 28119173]
30. Tugade MM, Fredrickson BL. Resilient individuals use positive emotions to bounce back from negative emotional experiences. J Pers Soc Psychol 2004 Feb;86(2):320-333 [FREE Full text] [doi: 10.1037/0022-3514.86.2.320] [Medline: 14769087]
31. Cameron D, Dromerick LJ, Ahn J, Dromerick AW. Executive/life coaching for first year medical students: a prospective study. BMC Med Educ 2019 May 22;19(1):163 [FREE Full text] [doi: 10.1186/s12909-019-1564-4] [Medline: 31118014]
32. Gooding PA, Hurst A, Johnson J, Tarrier N. Psychological resilience in young and older adults. Int J Geriatr Psychiatry 2012 Mar;27(3):262-270. [doi: 10.1002/gps.2712] [Medline: 21472780]
33. Hirani S, Lasiuk G, Hegadoren K. The intersection of gender and resilience. J Psychiatr Ment Health Nurs 2016 Aug;23(6-7):455-467. [doi: 10.1111/jpm.12313] [Medline: 27593204]
34. Li F, Luo S, Mu W, Li Y, Ye L, Zheng X, et al. Effects of sources of social support and resilience on the mental health of different age groups during the COVID-19 pandemic. BMC Psychiatry 2021 Jan 07;21(1):16 [FREE Full text] [doi: 10.1186/s12888-020-03012-1] [Medline: 33413238]

Abbreviations

CD-RISC-10: 10-item Connor-Davidson Resilience Scale
GAD-7: 7-item Generalized Anxiety Disorder
OLS: ordinary least squares
PHQ-9: 9-item Patient Health Questionnaire

Edited by A Mavragani; submitted 09.02.22; peer-reviewed by MM Abdul Kader Jilani, I Choi; comments to author 22.03.22; revised version received 21.04.22; accepted 09.05.22; published 29.07.22.

Please cite as:
Graziani G, Aylward BS, Kunkle S, Shih E
Changes in Resilience Following Engagement With a Virtual Mental Health System: Real-world Observational Study
JMIR Form Res 2022;6(7):e37169
URL: https://formative.jmir.org/2022/7/e37169
doi:10.2196/37169
PMID:35904875

Original Paper

Using a Brief Mental Imagery Competing Task to Reduce the Number of Intrusive Memories: Exploratory Case Series With Trauma-Exposed Women

Kristjana Thorarinsdottir[1], MSc; Emily A Holmes[2], Prof Dr; Johann Hardarson[1], MSc; Elin S Stephenssen[1], BSc; Marianna H Jonasdottir[1], BSc; Marie Kanstrup[2], PhD; Laura Singh[2,3], PhD; Arna Hauksdottir[4], Prof Dr; Thorhildur Halldorsdottir[4,5], PhD; Berglind Gudmundsdottir[6,7], Prof Dr; Edda Thordardottir[4], PhD; Unnur Valdimarsdottir[4,7,8], Prof Dr; Andri Bjornsson[1], Prof Dr

[1]Department of Psychology, University of Iceland, Reykjavík, Iceland

[2]Department of Psychology, Uppsala University, Uppsala, Sweden

[3]Swedish Collegium for Advanced Study, Uppsala University, Uppsala, Sweden

[4]The Center of Public Health Sciences, University of Iceland, Reykjavik, Iceland

[5]Department of Psychology, University of Reykjavik, Reykjavik, Iceland

[6]Faculty of Medicine, University of Iceland, Reykjavik, Iceland

[7]The National University Hospital of Iceland, University of Iceland, Reykjavik, Iceland

[8]Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, United States

Corresponding Author:
Kristjana Thorarinsdottir, MSc
Department of Psychology
University of Iceland
Sæmundargata 12
Reykjavík, 102
Iceland
Phone: 354 5254240
Email: kth35@hi.is

Abstract

Background: Novel interventions should be developed for people who have undergone psychological trauma. In a previous case study, we found that the number of intrusive memories of trauma could be reduced with a novel intervention. The intervention included a brief memory reminder, a visuospatial task and mental rotation, and targeted trauma memory hotspots one at a time in separate sessions.

Objective: This case series (N=3) extended the first case study with 3 new cases to determine whether a similar pattern of beneficial results is observed. We explored whether the brief intervention would result in reduced numbers of intrusive memories and whether it would impact symptoms of posttraumatic stress, depression and anxiety, and general functioning. Acceptability of the intervention was also explored.

Methods: A total of 3 women completed the study: 2 with posttraumatic stress disorder and other comorbidities and 1 with subthreshold posttraumatic stress disorder. The primary outcome was the change in the number of intrusive memories from the baseline phase to the intervention phase and at the 1-month follow-up, with an assessment of the intrusion frequency at 3 months. Participants monitored the number of intrusive memories in a daily diary for 1 week at baseline, for maximum of 6 weeks during the intervention phase and for 1 week at the 1-month and 3-month follow-ups. The intervention was delivered in person or digitally, with guidance from a clinical psychologist. A repeated AB design was used (*A* was a preintervention baseline phase and *B* intervention phase). Intrusions were targeted individually, creating repetitions of an AB design.

Results: The total number of intrusive memories was reduced from the baseline to the intervention phase for all participants. The total number for participant 3 (P3) reduced from 38.8 per week during the baseline phase to 18.0 per week in the intervention phase. It was 13 at the 3-month follow-up. The total number for P4 reduced from 10.8 per week at baseline to 4.7 per week in the intervention phase. It was 0 at the 3-month follow-up. The total number for P5 was reduced from 33.7 at baseline to 20.7 per week in the intervention phase. It was 8 at the 3-month follow-up. All participants reported reduction in posttraumatic stress

symptoms in the postintervention phase. Depression and anxiety symptoms reduced in 2 of the 3 participants in the postintervention phase. Acceptability was favorable.

Conclusions: We observed good compliance with the intervention and intrusive memory diary in all 3 cases. The number of intrusive memories was reduced for all participants during the intervention phase and at the 1-month follow-up, with some improvement in other symptoms and functioning. Further research should explore the remote delivery of the intervention and whether nonspecialists can deliver the intervention effectively.

(JMIR Form Res 2022;6(7):e37382) doi:10.2196/37382

KEYWORDS

trauma; intrusive memories; visuospatial task; Tetris gameplay; mental imagery; imagery competing task; case series; mobile phone; posttraumatic stress

Introduction

Background

Most people experience psychological trauma (eg, accidents or interpersonal violence) in their lives [1,2], and many (up to 37%) develop posttraumatic stress disorder (PTSD) after such experiences [3]. Intrusive memories are the core clinical symptoms of PTSD and are within the intrusion symptoms criterion of PTSD in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) [1,4]. Intrusive memories are persistent, unwanted upsetting memories of the traumatic event [1]. In their most extreme form, they can include reliving the traumatic event as if it were happening again (flashbacks). Other intrusion symptoms include dreams or nightmares about the traumatic event and emotional distress or physical reactivity after exposure to reminders of the traumatic event. Other symptom criteria include avoidance of memories or reminders of the trauma, along with negative alterations in cognition and mood [1]. Posttraumatic stress symptoms, even when subthreshold for a diagnosis of PTSD, can be associated with substantial distress, functional impairment, and comorbidity [5].

As noted previously [6], although evidence-based treatments for PTSD exist [7,8], the existing treatment options have some limitations. For example, current treatments require trauma survivors to talk in detail about the traumatic experience, which can be distressing, and many are reluctant to discuss their trauma in depth [9]. Dropout rates during PTSD treatment are high, up to 48% in clinical trials and approximately 18% overall and may be higher outside research trial settings in clinical practice [10-12]. Finally, existing options are time consuming, typically requiring numerous sessions, and there is often a lack of treatment providers specializing in empirically validated treatment of PTSD. Similar to numerous other countries, Iceland has mental health services that lack the capacity to offer treatments when needed by trauma survivors at the scale needed. These limitations to current treatments make the search for additional scalable treatment alternatives imperative.

A novel brief and simple intervention to reduce the number of intrusive memories after trauma has been developed based on cognitive science, as described elsewhere [13,14]. This intervention takes a single-symptom approach (not an entire disorder). The intervention includes a *brief memory reminder* cue for one *specific* intrusive memory of trauma, followed by a 25-minute Tetris gameplay with mental rotation (ie, actively rotating the blocks in one's mind eye to best make lines; [15]). The intervention was initially examined based on recent memories of trauma [16-18]. It has been further explored for *older memories* of trauma using case studies and case series approaches [6,19-21]. These studies involved in-person delivery, that is, face-to-face sessions guided by a clinical psychologist or researcher.

We adapted the intervention for women in Iceland who experienced intrusive memories of trauma, as reported in a recent case study [6]. Some of the details of this case are now summarized for context and comparison with the 3 new cases presented here. As previously reported, the participant was a woman in her fifties with 4 distinct intrusive memories from a traumatic event that happened in childhood, that is, the intrusive memories were decades old. Each specific memory was targeted in a session (in person) with a clinical psychologist with expertise in trauma. The memory reminder used was to *briefly* bring the visual content of the memory to mind without becoming emotionally overwhelmed by a method agreed with the participant (here, for example, choosing 1 of her 4 specific memories to be targeted using the diary, then thinking about the memory for a few seconds only, and letting the psychologist know when the memory was in their mind). Next, the participant was taught to use mental rotation. She then played Tetris for 25 minutes using mental rotation. She monitored her specific intrusive memories in a daily diary so that the impact of the intervention on a distinct intrusive memory could be easily observed. The total number of intrusive memories decreased from 12.6 per week at baseline to 6.1 per week in the postintervention phase. Furthermore, the number of intrusive memories continued to reduce to only 1 memory per week at the 3-month follow-up. Symptoms of posttraumatic stress and depression and anxiety reduced in the postintervention phase, whereas functioning improved. The participants considered the intervention to be an acceptable way to reduce the number of intrusive memories. The next step in exploring the effects of the intervention involves examining if they extend to other cases of women after trauma and whether remote (rather than in person) delivery is a feasible delivery method given restrictions occurring during the COVID-19 pandemic (eg, isolation) [22].

Objectives

In this case series, we aimed to extend our previous case study to a short case series of trauma-exposed women in Iceland, drawn from an epidemiological study of trauma experienced

by women in Iceland. The intervention sessions took place either in person in a university setting or remotely using a web-based platform. We examined whether the novel intervention approach could reduce the number of intrusive memories of trauma (primary outcome) and whether reductions were maintained at follow-up (1 and 3 months). As before, the brief intervention was guided by a clinical psychologist, targeting one distinct intrusive memory at a time per session. The acceptability of the intervention was also explored along with adaptions in intervention delivery. Again, we explored whether having fewer intrusive traumatic memories would also be associated with improvements in general functioning, posttraumatic stress, and depression and anxiety symptoms (secondary outcomes). The design adopted here can be described as a within-subject *repeated* AB design in which each specific memory is targeted in separate sessions, so that we can consider the effects of an individual intervention session on each specific memory over time [6,20].

As in our previous study [6], we predicted that participants would report fewer intrusive memories (primary outcome) during the intervention phase than in the preceding baseline phase and that the reduction in the number of intrusions would be maintained at the 1-month follow-up in the diary. In addition, we explored a 3-month follow-up using a diary. We expected that the number of targeted intrusive memories would decrease relative to that of nontargeted memories. We also examined whether having fewer intrusive memories would be associated with reductions in symptoms of posttraumatic stress and depression and anxiety and associated with improvements in general functioning (secondary outcomes). Furthermore, we explored the feasibility and acceptability of the intervention, alongside adaptions in intervention delivery format, that is, remote (web-based) delivery.

Methods

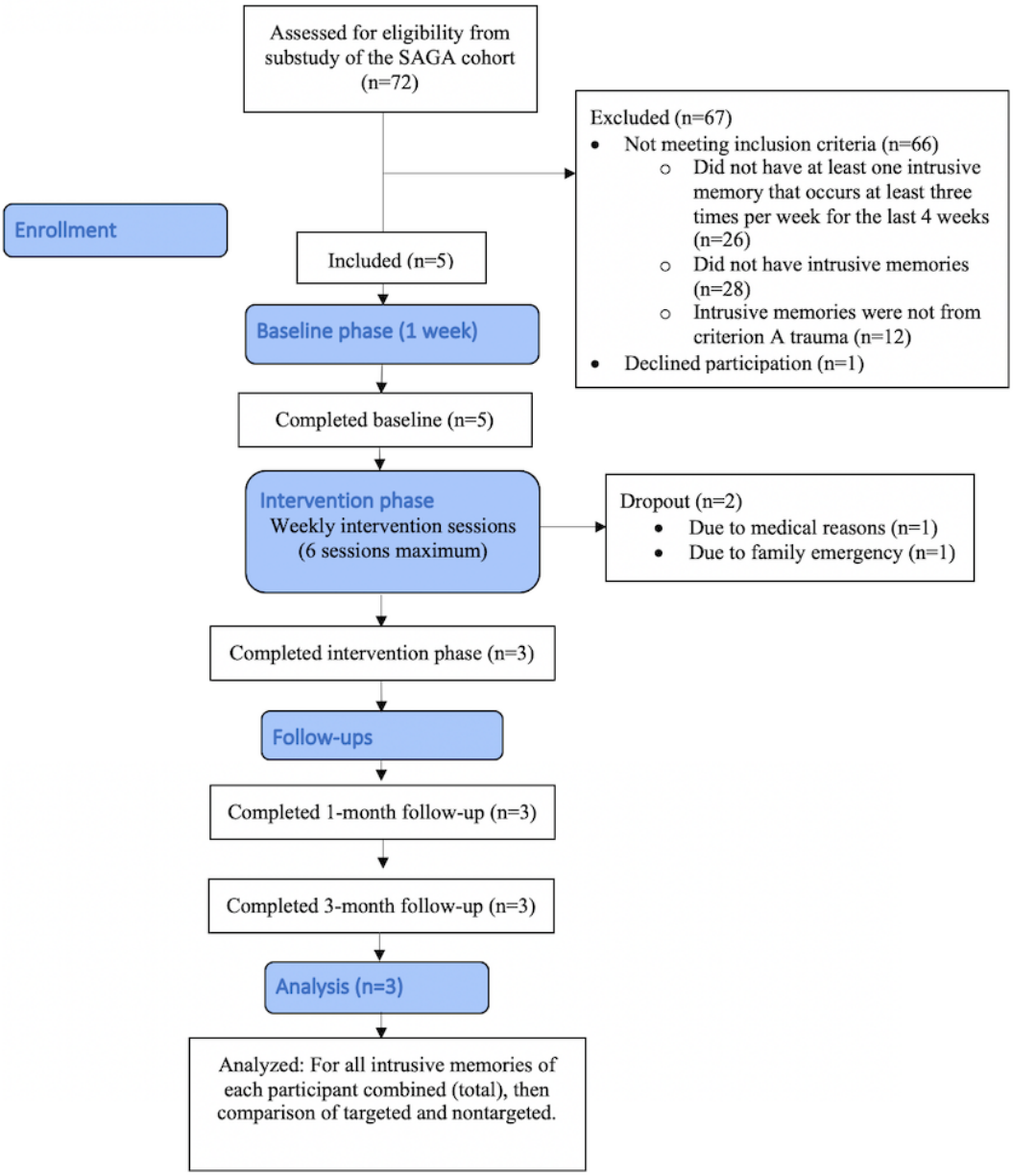
Participants

Participants were drawn from an epidemiological study of trauma experienced by women in Iceland (as in our previous case study [6]). As described previously [6], women who participated in a substudy of the Stress-And-Gene-Analysis (SAGA) cohort study were screened for eligibility. The SAGA cohort study was a population-based longitudinal cohort study of women in Iceland who completed an extensive questionnaire on trauma history and mental health (baseline data collection was completed on July 1, 2019). The Social Trauma Project substudy compared 2 samples from the SAGA cohort study, with a probable diagnosis of PTSD (ie, Posttraumatic Stress Disorder Checklist-5 [PCL-5] score of ≥33; see *Measures* section) or not likely PTSD (ie, PCL-5 score in the lowest fifth), using clinical interviews. In all, 2 semistructured interviews were administered in the substudy (ie, the Mini International Neuropsychiatric Interview [MINI], which was also used to assess the exclusion criteria for this study, and the Clinician-Administered PTSD Scale for DSM-5 (CAPS-5; see *Measures* section).

Women who took part in the Social Trauma Project substudy (both likely PTSD group and not likely PTSD group) were screened for the presence of intrusive memories of trauma, as in our previous case study [6]. As before, screening included a short description of intrusive memories (memories that include sensory impressions such as sight, sound, and so on; often pictures or a film clip that pops into the mind's eye; are distressing and occur involuntarily). Next, they were asked questions regarding the presence of this symptom to assess their eligibility for participation in this study ("Do you have intrusive memories of trauma? If yes, how often in the past week have you experienced such memories?" and "How often have you experienced intrusive memories a week in the past four weeks?").

In this study, 72 women from the substudy, who provided their consent to be contacted for further research were assessed for inclusion in this study (Figure 1).

Figure 1. Adapted CONSORT (Consolidated Standards of Reporting Trials) flow diagram for the study. SAGA: Stress-And-Gene-Analysis.



Inclusion criteria were (1) having experienced criterion A trauma as defined by DSM-5 [1] (using criterion A on CAPS-5), (2) having at least one intrusive memory that occurs at least three times per week for the last 4 weeks (How many intrusive memories have you experienced in the last 4 weeks on average?), (3) being able and willing to attend 3 to 8 sessions with the researcher, (4) being able and willing to monitor intrusive memories in daily life, (5) having access to a smartphone, and (f) being able to speak and read study materials in Icelandic. The exclusion criteria were (1) current psychotic disorder (assessed with the MINI), (b) current manic episode (assessed with MINI), and (c) being acutely suicidal (assessed with the MINI).

In all, 5 women who met the inclusion criteria were included, ranging in age from 39 to 66 years (mean 49, SD 11 years). Participants are referred to as P1, P2, P3, P4, and P5. The initial 2 participants recruited did not complete the intervention sessions due to nonsuitability of the study timing in their lives (medical and family issues). Both P1 and P2 were excluded from the data analyses, although data from the baseline were collected. After P1 and P2 wished to cease their participation, a question was added to the recruitment process asking if the participants foresaw any obstacles to their participation in the study. P1 did not meet any diagnostic criteria for psychological disorders according to the MINI. P2 met the diagnostic criteria for bipolar disorder and reported subthreshold PTSD symptoms.

A total of 3 patients completed the study, 2 (67%) with PTSD and other comorbidities and 1 (33%) with subthreshold PTSD. P3, woman aged ≥40 years, met the criteria for major depressive disorder and PTSD. P4, a woman aged ≥60 years, did not meet any diagnostic criteria but reported subthreshold PTSD according to the CAPS-5. P5, a woman aged ≥50 years, met the diagnostic criteria for major depressive disorder, social anxiety disorder, and PTSD. P3 reported having 10 different intrusive memories of physical violence that occurred in childhood. P4 reported having 3 different intrusive memories from childhood sexual abuse, and P5 reported 6 different intrusive memories from childhood sexual abuse.

Design

The case series used a single-symptom approach, in which each intrusive memory was targeted in different sessions, that is, one at a time [6,19]. Specifically, participants distinguished the content of their different intrusive memories and described them briefly; for example, for a participant who had four distinct intrusions, they may label them as (1) broken glass, (2) man's face, (3) blood on door, and (4) red car (the examples used are fictitious to preserve anonymity). Participants then monitored the occurrence of each distinct intrusion over time.

Thus, the design for each participant was a *repeated AB design* whereby the length of baseline (*A*, before intervention, monitoring only) and intervention (*B*) phases varied for each distinct intrusion, according to which intervention session the intrusive memory was targeted in. Thus, the baseline phases for each distinct intrusive memory were used as control periods in the comparison with the intervention, that is, the number before and number after being targeted.

A daily diary was used to monitor the number of each intrusive memory over time. That is, for 1 week before intervention, then over 6 weeks (maximum) of intervention; then for 1 week at the 1- and 3-month follow-ups. Intrusive memories were targeted individually in up to 6 intervention sessions. These sessions were guided by a clinical psychologist, who was a specialist in trauma-focused cognitive behavior therapy. After the first session, participants were able to self-administer the intervention if they wished for memories that they had already targeted in the session. P3 received 2 repetitions of an AB design, P4 only targeted 1 memory (and thus no repetitions of an AB design), and P5 received 4 repetitions of an AB design.

The primary outcome was the change in the number of intrusive memories per week from baseline to the intervention phase and to the long-term follow-ups (1 and 3 months). Participants further completed secondary outcomes—self-report measures for posttraumatic stress, depression and anxiety symptoms, and functional impairment—at baseline, the last intervention session, and 1- and 3-month follow-ups.

Procedure

Training

Researchers delivering the intervention (KT and JH; both licensed clinical psychologists trained in trauma-focused cognitive behavior therapy) underwent training and received clinical supervision, promoting adequate intervention delivery and protocol adherence. Training involved 2 in vivo workshops for 3 days and then again approximately 6 months later for 2 days, delivered by psychologists with expertise in developing the intervention and delivering it in other settings (MK and EAH). The theoretical background and practical aspects of the intervention (eg, how to explain and use the primary outcome measure) were covered in the workshops, as well as role-playing with trainers and feedback until sufficient performance was reached. While data collection was ongoing, the researchers received supervision. Such supervision included weekly supervision meetings, as well as more *in real time* support from a clinical supervisor via telephone directly after participant sessions regarding any case specific adaptions needed (EAH, MK, and AB). Twice a month, the researchers joined remote (Zoom; Zoom Video Communications) peer-group training meetings with other international researchers about the intervention (convened by EAH and LS).

Baseline Session

A similar procedure was followed as in our previous case study [6]; in the baseline session, the researcher explained the nature of intrusive memories (ie, memories that include sensory impressions such as sight, sound, and so on; are predominantly similar to pictures or a film clip in the mind's eye; and are distressing and occur involuntarily). Participants identified their intrusive memories by giving a brief verbal account of the intrusion's visual content using only a few words. Researchers noted down the image's description on a *hotspots* sheet in a way that the participant could also see it. Participants then labeled each intrusive memory with a symbol (the first memory was labeled *A*, the second memory *B*, etc) and were instructed on how to monitor their frequency each day in a pen-and-paper diary (primary outcome measure). To indicate when a certain memory was experienced, the participants recorded the symbol corresponding to a given memory for each time frame of that day. Each diary was divided into 7 days and each day, into 4 periods (see *Measures* section). The participants also completed baseline questionnaires (secondary outcomes) in the baseline session.

Intervention Sessions

In each intervention session (maximum 6 sessions), the participant chose 1 intrusive memory to target (by looking through their diary entries) and completed the intervention procedure (guided by the researcher). The memory chosen could be the most distressing or frequent or one chosen to be targeted by the participant for other reasons. As in our previous case study [6], the intervention consisted of a brief memory reminder, that is, briefly thinking about the intrusive memory to bring the image to mind without it becoming emotionally overwhelming. Please note that the approach to bringing the memory to mind here differs procedurally from the memory reminder method in the studies by Kessler et al [19] or Kanstrup et al [20]. Participants were told, "To make the game as useful as possible, we first had to make sure the memory was in your mind before using the intervention. So, I want to ask you what do you think would be the best way for you to bring this memory to mind without it becoming emotionally overwhelming?" They then discussed with the researcher options for the best way for them to bring the memory into mind without it emotionally becoming overwhelming. To do this, they were given examples of writing it down briefly and not discussing it with the psychologist, thinking about it briefly again and not talking about it in detail, or finding another method. Here, all participants chose to bring to mind the memory they had chosen to target by briefly thinking about it for a few seconds with their eyes open and telling the psychologist when it had come fully to mind (without talking about it in detail). The psychologist confirmed that the participants were able to picture their memory (ie, see it in their mind's eye) before moving to the instructions about the gameplay.

After the memory reminder procedure, participants were trained on how to use the Tetris game and practiced using mental rotation. They then played Tetris using mental rotation for 25 minutes [15]. For in-person meetings, the Tetris gameplay was delivered with a Nintendo DS 10.1-inch screen, set to *Marathon* mode with the ghost piece off. When an intervention session took place remotely with a video call, Tetris gameplay was performed on the participants' own computer with a shared screen so that the researcher could monitor the gameplay, especially regarding mental rotation. Only one distinct intrusion was targeted for each session. To select which intrusion was targeted, the participant (not the researchers) selected which intrusive memory it was. At the end of the last intervention session, secondary outcome measures were completed again.

Participants were invited to self-administer the intervention for memories already targeted in a session using a mobile version of the Tetris created by Electronic Arts [23]. For example, when the intrusion came to mind involuntarily in daily life, they were told to use a similar procedure as they had learned with the researcher in session.

When the COVID-19 pandemic started (the University of Iceland closed on March 19, 2020), researchers switched to remote (rather than in person) delivery through Kara Connect, which is a General Data Protection Regulation–compliant web-based platform certified by the Icelandic Directorate of Health. The last intervention session for P3 was performed remotely, intervention sessions 2 and onward were performed remotely for P4, and all sessions were remotely delivered for P5. Tetris was played on the web on the participants' computers (ghost piece off and sound set to 0%) with a shared screen so that the researcher could monitor participants' gameplay via Kara Connect, which increased the likelihood of instruction adherence, especially regarding the use of mental rotation.

Follow-up

At both the 1- and 3-month follow-up time points, participants recorded the number of intrusions in their diary daily again for 1 week and completed the secondary outcome measures. Data were entered via laptop into a REDCap (Research Electronic Data Capture; Vanderbilt University) database (REDCap), an encrypted electronic software stored on a secure server [24].

Measures

Eligibility Assessments (Part of the SAGA Cohort Substudy)

Please note that the measures described here have already been described in our previous case study [6] and are repeated here for clarity.

The Clinician-Administered PTSD Scale (CAPS-5) is a 30-item semistructured interview used to assess symptoms of PTSD from an index of trauma and symptom severity (in the past month) according to DSM-5 [1]. Items are scored on a 5-point Likert scale (0=mild or subthreshold; 4=extreme or incapacitating). A symptom rating of 2 (ie, moderate) was the threshold for a possible diagnosis. For each symptom, frequency and intensity were assessed and rated separately. The CAPS-5 has excellent internal consistency (Cronbach α=.88), test-retest reliability (α=.83), and good convergent validity (α=.83) [25]

The MINI for the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, assesses Axis I psychiatric disorders according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, using a structured diagnostic interview. For most diagnoses, the MINI has good sensitivity and specificity [25]. Interrater and test-retest reliability is good, with κs in the high to very high range (κs=0.79-1.00) [26].

Primary Outcome Measure

The intrusive memory diary, a pen-and-paper diary similar to that used in Thorarinsdottir et al [6], was adapted from previous experimental and clinical studies [16,20]. It involves daily recording, for 4 time points each day (morning, afternoon, evening, and night) for 1 week. The diary instructions defined the nature of intrusive memories as distressing and involuntary mental images (such as visual images or a film in the mind's eye). Participants were asked not to record voluntary (ie, deliberately recalled) thoughts or involuntary verbal thoughts without sensory content (intrusive verbal thoughts that had an imagery component could be included). Participants monitored the occurrence (or nonoccurrence) of their intrusive memories in the daily diary for 1 week before the intervention (baseline phase), for a maximum of 6 weeks during the intervention phase and then again for 1 week at both the 1- and 3-month follow-ups (note this was a daily diary, see retrospective amendment to clinical trial registration [CTR] NCT04209283). When indicating having an intrusion, participants used the symbol corresponding to that specific memory (eg, A or B as noted earlier), and therefore, it was possible to examine change in frequency (here calculated as the number per week) for each distinct memory individually.

The primary outcome was the change in the number of intrusive memories from the baseline to the intervention phase and at the 1-month follow-up. The original CTR additionally prespecified a measure of intrusive memories at the 3-month follow-up ("Change in the number of intrusive memories of trauma from baseline to 3 month follow-up") but incorrectly stated that the measure was "Questions about the frequency of intrusive memories for the past day or week (for each intrusive memory, to be tallied to arrive at a mean frequency for the memories for the previous day and for the week)," specifically "How often did this memory come up yesterday?" and "How often did this memory come up per day in the past week?" However, this was incorrect, as we had changed this measure at the study start to use the same diary as for the earlier time points (ie, daily diary), that is the "Number of intrusive memories of traumatic event recorded by participants in a diary daily (morning, afternoon, evening, night) per week over the baseline phase and during one week at three month follow-up." The measure has been updated retrospectively in the CTR for this 3-month period and should be interpreted with caution.

Secondary Outcome Measures

PTSD symptoms in the past month were evaluated using the PCL-5, a self-report scale with 20 items. Each symptom is rated on a 5-point Likert scale (0=not at all; 4=extremely), with higher

scores indicating greater severity. The PCL-5 evaluates the severity of PTSD symptoms according to DSM-5 criteria. It has strong internal and test-retest reliability and good convergent and discriminant validity [27]. Criteria for clinical significance are not available for the PCL-5; however, posttreatment scores of ≤24 can be interpreted as clinically significant change [28]. The Icelandic version of this measure in the SAGA cohort study had excellent internal consistency (α=.95).

Depression symptoms in the past 2 weeks were evaluated using the Patient Health Questionnaire-9 (PHQ-9), a self-report measure with 9 items rated on a 4-point Likert scale (0=not at all; 3=nearly every day) [29]. The PHQ-9 evaluates depression symptoms and their severity and has good internal reliability and test-retest reliability [29]. The Icelandic version of the SAGA cohort study had good internal consistency (α=.89). A 5-point change in the total PHQ-9 score was considered clinically significant [30].

Anxiety symptoms in the past 2 weeks were evaluated using the Generalized Anxiety Disorder-7 scale (GAD-7). In this self-report questionnaire, each item is rated on a 4-point Likert scale (0=not at all; 3=nearly every day). The GAD-7 assesses symptoms of general anxiety disorder and its severity and has great internal reliability and good test-retest reliability [31]. In general, the GAD-7 has presumably been useful for screening anxiety disorders [32]. The Icelandic version of the SAGA cohort study had good internal consistency (α=.90). A 4-point change on the GAD-7 is considered clinically significant [33].

Functional impairment in the previous week was evaluated using the Sheehan Disability Scale (SDS). This self-report measure has three domains: (1) work or school, (2) social, and (3) family life, assessing functional impairment using a 11-point scale (0=not at all; 10=extremely) [34]. A 3-point change on the SDS scale has been considered indicative of response to treatment [35]. To assess the impairment associated with intrusive memories, scale adjustments were made. The SDS has been found to have good internal and test-retest reliability and good construct validity [34]. The Icelandic version of the SDS has been found to have good internal consistency in clinical groups (α=.84) [36].

Self-guided adherence for daily life using the gameplay part of the intervention was rated "How often did you manage to play Tetris after you experienced an intrusive memory?" (11-point scale: 0=not at all; 10=every time).

Feasibility and acceptability ratings for the intervention were assessed with 2 ratings "Would you recommend playing Tetris to a friend?" and "Do you consider gameplay to be an acceptable way to reduce the daily frequency of intrusive memories?" The scores ranged from 0 to 10. Higher scores indicated greater acceptability. Open-ended questions included (1) "How did you feel about playing Tetris after you had an intrusive memory?" and (2) "Did you find the intervention helpful? If yes, how?"

The impact of intrusive memories on concentration, sleep, and stress in the past week was evaluated using 6 self-rated items: a total of 2 items assessed general concentration difficulties and difficulties in concentration due to intrusive memories (11-point scales; higher scores indicated greater difficulties), 1 item assessed concentration disruption in minutes in the past week, and 2 items assessed the impact of intrusive memories on sleep (11-point scale; higher scores indicated greater sleep disturbance). An item assessed the impact of intrusive memories on stress levels (0=not at all; 10=affected very much).

General impact of intrusive memories was assessed using 2 ratings of intensity and vividness of the intrusions on a 11-point scale (0=not at all; 10=very distressing or vivid).

Intrusion diary adherence item was assessed using the rating "How accurately did you fill out the diary?" (0=not at all; 10=very accurate).

Impact of intrusive memories on daily functioning was evaluated via an open-ended question: "How have the intrusive memories affected your ability to function in your daily life in the past week?" and a self-rated question, "Have the intrusive memories affected your ability to function in your daily life?" (11-point scale, a higher score indicated greater impact).

Data Analysis

Number of Intrusive Memories

The number of intrusive memories of trauma was recorded by participants in the diary daily (morning, afternoon, evening, and night) during the baseline phase and each week during the intervention phase (weeks 1-6) and during 1 week at the 1-month follow-up. The primary outcome was the change in the number of intrusive memories of the trauma. The timeframe was baseline week to the intervention phase (weeks 1-6) and follow-up (1 month). In practice, owing to scheduling reasons, the baseline phase was longer than 1 week, and as anticipated, the number of intervention weeks varied. Therefore, because these periods had different time lengths, the mean number *per week* was calculated for comparability. Missing data were dealt with by excluding these time points from the calculations and using available data (see *Results* section). For example, a participant had a baseline period of 14 days, but data were present for only 6.5 days; thus, the total number of intrusions per week was calculated as 10 intrusions / 6.5 days × 7 = 10.8 intrusive memories per week during baseline.

When examining change over time, the percentage reduction in total intrusions per week was calculated from the baseline phase to the intervention phase to other periods as follows: 1– (mean number per week during intervention phase / mean number per week during baseline) × 100. For example, for the same participant there were 4.7 intrusions per week in the intervention phase, which was calculated as 1 – (4.7 / 10.8) ×100 = 57% reduction in the intervention phase compared with the baseline.

Please note that at the 3 month-follow-up, we did not use the telephone questions on the CTR (NCT04209283) about the frequency of intrusive memories for the past day or week (eg, "How often did this memory come up per day in the past week?") but instead replaced this with the same 1 week diary used at earlier time points. We noted the use of a diary at 3 months in our ethics submission, but we incorrectly specified it in our CTR and did not update the CTR on this point until the submission of this paper.

Change in the Number of Targeted Intrusive Memories Relative to Nontargeted Memories

We examined the number per week of targeted memories in comparison with nontargeted memories. This was done by calculating the number in the same way as described above for targeted memories (ie, each of targeted memories have different baseline and intervention periods). However, standard baseline and intervention periods were established for nontargeted memories (ie, same length of periods for all nontargeted memories), as they were not targeted by the intervention. The baseline for each nontargeted memory was 1 week (ie, before any memory was targeted), and the intervention phase was determined as the period from when any memory was targeted with the intervention.

Other Symptoms and Functioning

A descriptive approach was used to explore whether clinically meaningful changes were observed in the overall symptoms of posttraumatic stress, depression, anxiety, and functional impairment.

Ethics Approval

This study was approved by the National Bioethics Committee of Iceland (VSNb2017110046/03.01). The participants provided written informed consent before the start of the study. All the sessions followed a written protocol. No adverse events related to the intervention were reported. Participants were asked to briefly consider the visual content of the traumatic memories they selected, which might have resulted in some distress. Previous research has indicated that this intervention approach is well tolerated, including in inpatients with complex PTSD [19]. However, given the early stage of this research, an arrangement was made with an independent clinical psychologist who specializes in trauma for an interview free of charge to the participant and to be referred to a licensed clinical psychologist for treatment if needed. None of the participants had used these services.

Open Science Statement

The study was registered before the start of the study on ClinicalTrials.gov (NCT04209283) on December 24, 2019. The manuscript contains anonymized summary-level data. The study materials may be made available upon reasonable request with an appropriate material transfer agreement with the University of Iceland or Uppsala. We note that delivery of this intervention at present requires prerequisite training and supervision by psychologists with experience of developing it (see *Procedure*, *Training* section).

Results

Overview

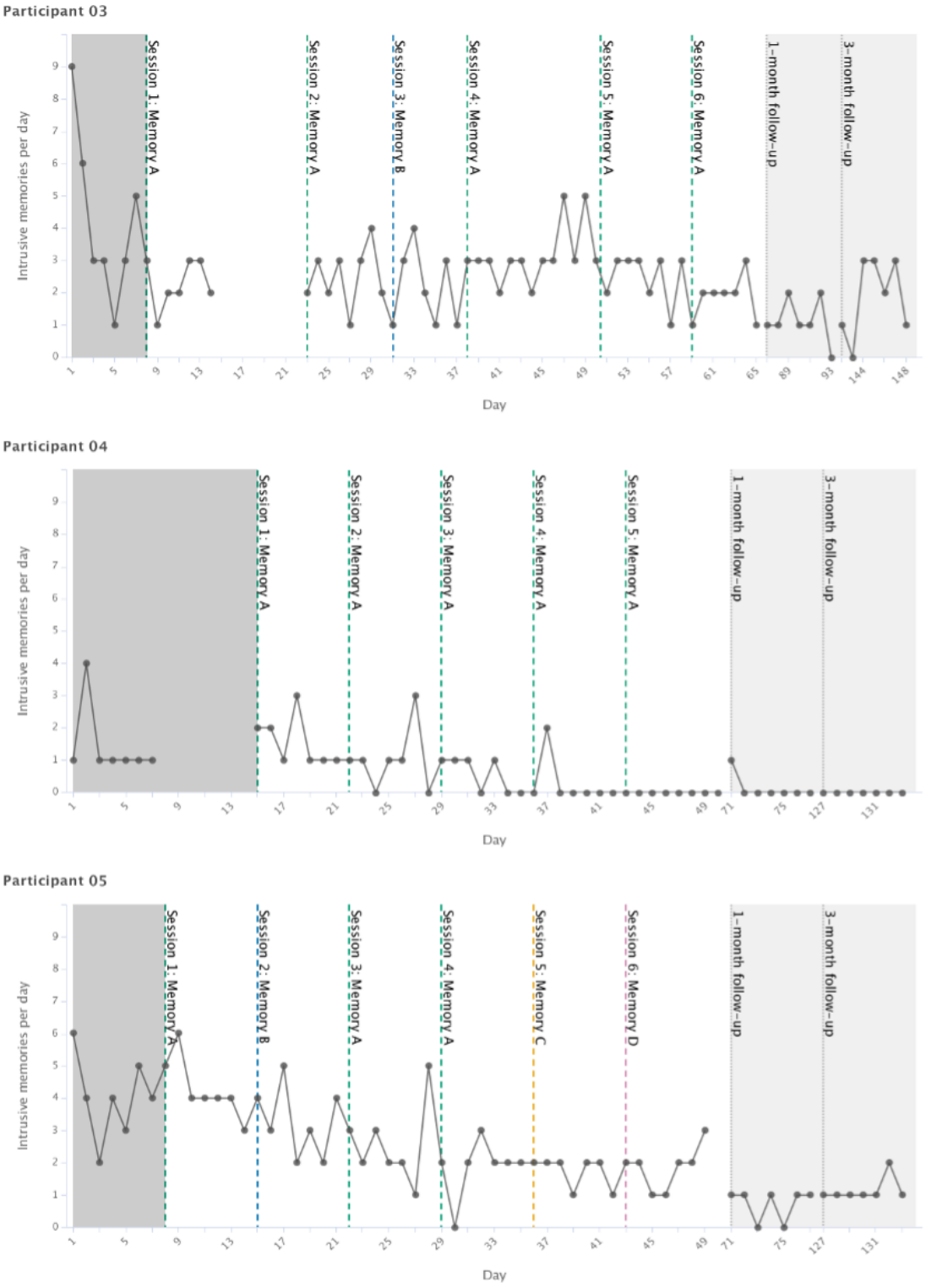
The number of intrusive memories and how many were targeted varied among participants. P3 reported having 10 different intrusive memories of physical violence during childhood. In all, 2 of the intrusions were targeted with the intervention; one of the memories (Memory A) was targeted 5 times and the other (Memory B) was targeted once. P3 monitored the memories quite accurately, but data were missing for days 15 to 22 during the intervention phase. P4 reported having 3 intrusive memories of childhood sexual abuse, of which only 1 (Memory A) was targeted with the intervention 5 times. P4 monitored the memories accurately, and data were missing for half of day 1 and for days 8 to 15 during the baseline phase. P5 reported having 6 intrusive memories of childhood sexual abuse, 4 of which were targeted. One of the memories (Memory A) was targeted 3 times, whereas the others (Memories B, C, and D) were targeted once. P5 accurately monitored their memories with no missing data. No attempts were made to retrieve the missing data. On average, 90% of the diary data were completed for all 3 participants. No data were missing for the secondary outcome measures.

Primary Outcome

Change in the Total Number of Intrusive Memories

The total number of intrusive memories per day throughout all phases for each participant (baseline, intervention, and 1 month) is shown in Figure 2. In addition, as it is on the same measure, the diary used at 3 months is also shown in Figure 2. Diary compliance was good for the outcome phases, with most missing data in the baseline phase. The number of intrusive memories per day fluctuated during the baseline phase for all the participants (N=3). P3 had 38.8 intrusive memories (summed across all 10 distinct intrusive memories) per week during the baseline phase. The number was reduced to 18.0 per week (54% reduction from baseline) during the intervention phase and further reduced to 8 at the 1-month follow-up week (80% reduction from baseline). P4 had 10.8 intrusions (summed across all 3 memories) per week during the baseline phase, and the number reduced to 4.7 per week (57% reduction from baseline) during the intervention phase and was further reduced to 1 (91% reduction from baseline) at the 1-month follow-up week (Figure 2). P5 had 33.7 intrusive memories (summed across all 6 distinct intrusive memories) per week during the baseline phase, which reduced to 20.7 per week (39% reduction from baseline) during the intervention phase. The number further reduced at the 1-month follow-up to 5 that week (85% reduction from baseline).

Figure 2. Graphs for visual inspection of primary outcome data for each participant (on the y-axis as total number of intrusive memories per day, ie, for all distinct memories combined). Days since study start are shown on the x-axis, which includes baseline (gray), intervention (white), and follow-up periods (light gray). Dashed vertical lines show when each intervention session was administered and which specific traumatic memory was targeted. Memories are labeled in order of when they were targeted (eg, Memory A was targeted in the first intervention session). Dotted vertical lines show the 1-month and 3-month follow-up periods. Gaps in the time series; for example, in the baseline, represent missing data.



We now consider each of the participant graphs shown in Figure 2. Visual inspection of P3 showed that relative to baseline after the first intervention session, the number of intrusions decreased. The number persisted with some fluctuation through the intervention period (days 8-65) and the 1-month and 3-month follow-ups. Visual inspection for P4 showed that relative to baseline (which included missing data), the number of intrusions remained relatively steady until the fourth intervention session targeting the same memory (day 36), when there was a noticeable drop in occurrence to 0 that was maintained in the last intervention session and 1-month (and 3-month) follow-up. Visual inspection of P5 showed a slight drop in the frequency

of intrusions after the second intervention session with some fluctuations, until intervention session 5 (day 36), where the reduction in frequency became more stable (Figure 2). The frequency decreased even further at the 1-month follow-up (and at 3 months).

Data for the 3-Month Follow-up Diary

P3 had 13 intrusive memories in their diary at the 3-month follow-up (67% reduction from baseline). P4 had 0 intrusive memories (100% reduction from baseline) at the 3-month follow-up, whereas P5 had 8 in the week of the 3-month follow-up (76% reduction from baseline; Figure 2). Patterns in relation to diaries at earlier time points are noted in the *Change in the Total Number of Intrusive Memories* section; however, please see earlier notes about the change in measures at this time point in contrast to the original CTR.

Change in the Number of Targeted Intrusive Memories Relative to Nontargeted Memories

The number of targeted and nontargeted intrusions per week at baseline, in the postintervention phase, and at the 1- and 3-month follow-ups are displayed in Table 1. The mean number of individual targeted memories was 7.6 (SD 4.3) per week in the baseline phase and reduced to 5.8 (SD 2.7) per week in the intervention phase. For individual memories, refer to Table 1. However, for nontargeted intrusions, the baseline rate for individual memories was very low, such as 1 per week. Therefore, these percentages are not included in Table 1. The mean number of nontargeted memories was 2.5 (SD 3.3) per week in the baseline phase and reduced to 0.2 (SD 0.4) memories per week in the intervention phase. The number of targeted memories continued to decrease at the 1-month follow-up week to 2.0 (SD 2.5), and nontargeted memories were reduced to 0 memories. At the 3-month follow-up, the mean number of targeted memories was 3.0 (SD 3.4), and the frequency for nontargeted memories was again 0 that week.

Table 1. Number of intrusive memories per week in the baseline phase, postintervention phase, and at 1 and 3-month follow-ups from baseline phase for each participant (P)[a].

Participant and intrusions	Baseline phase	Postintervention phase	1-month follow-up	3-month follow-up
P3				
Memory 1 (A[b])	14	11.2	6	3
Memory 2 (B[b])	8.8	6.4	2	10
Memory 3 (C)	6	0.3	0	0
Memory 4 (D)	1	0	0	0
Memory 5 (E)	0	0	0	0
Memory 6 (F)	3	0	0	0
Memory 7 (G)	1	0	0	0
Memory 8 (H)	3	0.1	0	0
Memory 9 (I)	0	0	0	0
Memory 10 (J)	2	0	0	0
Total	38.8	18.0	8	13
Total targeted[b]	22.8	17.6	8	13
Total nontargeted	16	0.4	0	0
P4				
Memory 1 (A)[b]	8.6	3.9	1	0
Memory 2 (B)	1.1	0.8	0	0
Memory 3 (C)	1.1	0	0	0
Total	10.8	4.7	1	0
Total targeted[b]	8.6	3.9	1	0
Total nontargeted	2.2	0.8	0	0
P5				
Memory 1 (A)[b]	10.6	6.5	0	2
Memory 2 (B)[b]	7	2.8	0	2
Memory 3 (C)[b]	2.6	4.2	0	4
Memory 4 (D)[b]	2.0	5.6	5	0
Memory 5 (E)	11.6	0.5	0	0
Memory 6 (F)	0	1	0	0
Total	33.7	20.7	5	8
Total targeted[b]	22.1	19.2	5	8
Total nontargeted	11.6	1.5	0	0

[a]Total intrusions do not equate to the sum of intrusions for each memory here because the length of the baseline and intervention phases differ across each memory and the total. See the *Data analysis* section for further details.

[b]Specific intrusive memories targeted by the intervention.

Secondary Outcomes

Ratings of Adherence to Completing the Diary, General Impact of Intrusive Memories (Vividness and Distress), and Use of the Intervention in Daily Life

Self-reported accuracy for filling out the daily intrusion diaries was high throughout the study period (Table 2). For the general impact of intrusive memories, the ratings for vividness of the intrusions did not change considerably, although if anything, showed some decline (Table 2). Distress associated with intrusions more clearly diminished during the intervention and follow-up phases for all 3 participants (Table 2). For use of the gameplay intervention in daily life, ratings of the self-guided

adherence to Tetris gameplay after experiencing an intrusive memory between sessions are also shown in Table 2.

Table 2. Ratings for adherence to intrusive memory diary, general impact of intrusive memories, and use of the gameplay intervention in daily life (N=3).

Ratings and participant	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	1-month follow-up	3-month follow-up
Diary accuracy[a]								
P[b]3	8	7	8	7	8	7	7	7
P4	9	8	7	9	8	__[c]	8	10
P5	8	9	8	8	8	8	7	9
Intrusions' vividness[d]								
P3	8	8	7	8	7	6	6	7
P4	7	6	5	7	4	—	5	0
P5	7	8	7	7	7	6	6	5
Intrusions' distress[e]								
P3	7	9	7	7	7	6	6	4
P4	6	5	4	4	2	—	1	0
P5	8	8	7	7	7	6	5	3
Intervention use in daily life[f]								
P3	—	3	3	1	4	3	1	1
P4	—	3	2	7	0	—	0	0
P5	—	7	7	6	8	8	6	2

[a]How accurately did you complete the diary? 0=not at all; 10=very accurately.

[b]P: participant.

[c]Missing data.

[d]During the last week, how vivid was your intrusive memory? 0=not at all; 10=very vivid.

[e]During the last week, how distressing were your intrusive memories? 0=not at all; 10=very distressing.

[f]How often did you manage to play Tetris after you experienced an intrusive memory? 0=never; 10=every time.

Feasibility and Acceptability of Using a Computer Gameplay Intervention

Participants were asked if they would recommend the intervention to a friend. P3 rated this item at 9, P5 gave it a rating of 8 (highly likely to recommend to a friend), and P4 rated the item at 3 (unlikely to recommend to a friend). They also rated if they considered gameplay to be an acceptable way to reduce intrusive memories. P3 and P5 rated the item at 8 (high acceptability), and P4 rated it at 3 (low acceptability). When participants were asked to rate how they felt about playing Tetris after having an intrusion, P3 noted that "It reduced the emotion, sometimes I was able to concentrate and think my way through it. Sometimes I experienced a kind of peace within." P4 said, "I played so there was no room for other thoughts," and P5 reported, "Sometimes it's difficult to play for 25 minutes, but I played as many times as I could, for 5 to 25 minutes." When the participants were asked if they found the intervention helpful, P3 reported, "When I was able to plan ahead while playing the game, my brain could not interrupt me. The emotion that causes distress fades." P4 said, "Yes, I could not think about anything else whilst playing," and P5 reported, "I felt a physical calmness, like the pit in my stomach was shrinking."

Self-report Measures for Posttraumatic Stress, Depression and Anxiety Symptoms, and General Functioning

There was a clear reduction in posttraumatic stress symptoms (PCL-5) from baseline to the postintervention phase, which tended to continue to drop further during follow-up for all 3 participants who completed the intervention, suggesting clinical improvement (Table 3). Depression symptoms (on the PHQ-9) and anxiety symptoms (on the GAD-7) seemed to follow a similar pattern for 2 (P3 and P5) out of the 3 participants, showing reductions in the postintervention phase and during follow-up. Functional impairment (as measured by the SDS) was reduced for the same 2 out of 3 participants (P3 and P5) from baseline to the postintervention phase and at the 1-month follow-up, but for P3, it increased at 3 months. It should be noted that P4's ratings for all of these measures were low at baseline (Table 3).

Table 3. Self-report measures of secondary outcomes (posttraumatic stress, depression and anxiety symptoms, and general functioning) and impact of intrusive memories on concentration, sleep, stress, and daily functioning for each participant (P).

Item and participant	Baseline interview	Postintervention interview	1-month follow-up	3-month follow-up
PCL-5[a]				
P1	6	__[b]	—	—
P2	54	—	—	—
P3	54	35	20	35
P4	26	14	6	11
P5	64	49	51	31
PHQ-9[c]				
P1	3	—	—	—
P2	19	—	—	—
P3	13	8	6	9
P4	5	3	3	2
P5	19	11	14	11
GAD-7[d]				
P1	2	—	—	—
P2	8	—	—	—
P3	14	7	6	10
P4	2	3	5	1
P5	18	9	17	4
SDS[e]				
P1	7	—	—	—
P2	20	—	—	—
P3	20	16	13	21
P4	4	3	0	0
P5	21	16	8	7
Concentration disruption related to intrusive memories[f]				
P1	3	—	—	—
P2	3	—	—	—
P3	7	5	3	5
P4	2	2	1	0
P5	7	4	4	3
General concentration[g]				
P1	0	—	—	—
P2	6	—	—	—
P3	7	6	5	6
P4	2	2	2	0
P5	7	4	3	—
Duration of disruption[h]				
P1	1	—	—	—
P2	3	—	—	—
P3	4	4	3	4

Item and participant	Baseline interview	Postintervention interview	1-month follow-up	3-month follow-up
P4	2	1	2	1
P5	4	4	2	2
Sleep[i]				
P1	0	—	—	—
P2	5	—	—	—
P3	4	4	3	4
P4	1	1	1	0
P5	10	8	7	5
Nightmares[j]				
P1	0	—	—	—
P2	4	—	—	—
P3	7	2	2	4
P4	0	1	1	0
P5	10	8	8	5
Stress[k]				
P1	0	—	—	—
P2	4	—	—	—
P3	6	2	0	2
P4	1	1	2	0
P5	7	3	3	3
Daily functioning[l]				
P1	0	—	—	—
P2	8	—	—	—
P3	7	5	2	5
P4	0	0	0	0
P5	7	0	2	0

[a]PCL-5: Posttraumatic Stress Disorder Checklist-5; score range 0 to 80.

[b]Missing data.

[c]PHQ-9: Patient Health Questionnaire-9; score range 0 to 27.

[d]GAD-7: Generalized Anxiety Disorder-7 scale; score range 0 to 21.

[e]SDS: Sheehan Disability Scale; score range 0 (unimpaired) to 30 (highly impaired).

[f]In the past week, how much did your intrusive memories disrupt your concentration? 0=not at all disruptive; 10=extremely disruptive.

[g]In the past week, how much difficulty did you have concentrating generally? 0=no concentration difficulty at all; 10=extreme concentration difficulty.

[h]When you had an intrusive memory, how long did it disrupt your concentration (in minutes) in the past week? 0 (<1 min) to 5 (>60 min).

[i]Did your intrusive memories interfere with sleep during the night in the past week? 0=not at all; 10=interfered very much.

[j]Did you experience any nightmares that interfered with your sleep during the night in the past week? 0=did not experience any nightmares; 10=experienced a lot of nightmares.

[k]In the past week, did your intrusive memories affect how stressed you felt? 0=not at all; 10=affected very much.

[l]Have the intrusive memories affected your ability to function in your daily life? 0=not at all; 10=affected very much.

Impact of Intrusive Memories on Concentration, Sleep, Stress, and Daily Functioning

The ratings of the impact of intrusive memories on concentration, stress, and sleep are included in Table 3. P5 reported improved concentration after her intrusions over the course of the intervention and follow-up period. At baseline, when P5 experienced an intrusive memory, her concentration was disrupted for 10 to 30 minutes on average, and this time was reduced to 1 to 5 minutes at follow-up. P3 reported a slight improvement in concentration with respect to intrusive memories specifically but not in general. Ratings on concentration disruption from intrusive memories were low at baseline for P4 and did not change in the postintervention phase. The effect of

intrusive memories on sleep did not change for P3 or P4, whereas P5 showed some reduction. Nightmares were reduced for both P3 and P5, whereas nightmares were almost nonexistent for P4. The stress levels associated with intrusions were reduced during the intervention and follow-up periods for P3 and P5.

The reported impact of intrusive memories on the ability to function in daily life over the intervention and follow-up periods showed improvement (except for P4, who scored 0 at baseline). At baseline, the participants were asked to respond to an open question about how their intrusive memories impacted their ability to function in daily life. At baseline, P3 said, "When I have this overwhelming feeling, I find it difficult to be around other people, the worst thing is how it affects my ability to stay present for my daughters." P4 said, "I have anger inside me, I constantly think back to how no-one noticed what was done to me." P5 reported, "I get very stressed and anxious; it takes a lot of energy to get out of the emotion..." In the last intervention session, P3 reported, "The intrusions disrupt my concentration," and P5 said, "I experience distress and nightmares." At the 1-month follow-up, P3 disclosed, "When I am under a lot of stress, they disturb me more." P5 reported, "The images no longer have color, cause less disruption and are less frequent." At the 3-month follow-up, P3 said, "I am under a lot of stress and dealing with a certain communication problem which triggers my PTSD and the intrusions" and that "I can now comprehend that this is only a memory, and I don't feel as distressed." P5 reported, "They mostly impact my anxiety" and "I am not a person that easily believes in things, but this intervention works." For all phases following the baseline, P4 reported, "The memories have no impact anymore."

Discussion

Principal Findings

The aim of this case series was to extend our previous case study by evaluating a novel visuospatial intervention designed to reduce the number of intrusive memories of trauma [6,14,37]. The intervention was adapted to the Icelandic setting from previous clinical studies [6,19,20]. The total number of intrusive memories per week (primary outcome) was reduced between 38% and 56% from baseline to the intervention phase, in line with earlier results found by Kessler et al [19] and what we found in our earlier case study [6]. Importantly, we also found in this case series that the frequency of intrusive memories continued to decrease at the 1-month follow-up (the reduction from baseline was approximately 85%). In the diary used at the 3-month follow-up, the reduction from baseline was from 66% to 100%; although this measure replaced another one and was not prespecified, further investigation is required. These results are similar with what we found in the earlier case study [6] in which the frequency also continued to reduce from 52% in the postintervention phase to 76% at the 1-month follow-up and to 92% at the 3-month follow-up. These results indicate that the reduction in the frequency of intrusions might continue rather than rebound, which could be because of the simplicity of self-administered use of the intervention, giving participants the chance to use it independently, if needed. Reductions in distress related to intrusive memories were evident in all participants. A limitation of the study is that the 3-month diary data must be treated as exploratory, as it was not preregistered in the CTR (though it was in our ethics approval), and further studies should include this.

Not only did the targeted intrusions reduce in this case series but nontargeted intrusions were also reduced in the intervention phase. Although this appears relatively more so than the targeted ones, results must be treated with caution because of the potential floor effects on the low baseline number of nontargeted intrusions rendering comparisons misleading. Kessler et al [19] found that overall, targeted memories were reduced by 64% and nontargeted memories by 11%. The sample of participants in the study by Kessler et al [19] was inpatients with a diagnosis of complex PTSD with a larger number of different memories and baseline symptom rates, whereas the participants in this case series were non–treatment seeking with less symptom severity. Furthermore, targeted memories were reported to be much more distressing than nontargeted ones and may, in some cases, need more time to reduce in frequency (hence, the long-term effect described earlier in this section). It would be clinically interesting in future studies to see if there were links between treating a memory (say from the same trauma) that could generalize to reductions in nontargeted memories of the same episode.

We examined whether a reduction in the number of intrusions would have an impact on the symptoms of posttraumatic stress, depression and anxiety symptoms, and general functioning (secondary outcomes). The general pattern was that posttraumatic stress symptoms were reduced for all participants in the postintervention phase (cutoffs for clinical significance were not available for this measure). Depression and anxiety symptoms were reduced at times suggestive of a clinically significant change for 2 (P3 and P5) of the 3 participants in the postintervention phase [30,33], whereby a 5-point change in the PHQ-9 total score and a 4-point change on the GAD-7 can be considered clinically significant. Symptoms tended to be reduced further at follow-up. It should be noted that P4 had very low levels of distress, depression and anxiety symptoms, and impaired function at baseline. The overall findings are similar to those of our previous case study [6] and the study by Kessler et al [19]. These results provide preliminary evidence that a reduction in the number of intrusions from using this intervention could possibly reduce other symptoms connected to intrusive memories after trauma and improve functioning, with important implications for the quality of life. For the 2 (out of 3) participants who had a reduction in posttraumatic stress symptoms, depression and anxiety, and impairments in concentration and other factors related to intrusive memories, the overall pattern for secondary measures was that some improvements tended to continue for a longer term.

P3 and P5 rated the intervention as acceptable; using Tetris gameplay was an acceptable method to reduce the frequency of their intrusive memories and noted that they would recommend the intervention to a friend. This is similar to what Holmes et al [38] found among refugees and what we found in our earlier case study [6]. However, P4 did not rate the intervention as acceptable and was unlikely to recommend it to a friend, although she noted that it was helpful. It is important

to further develop how the intervention can be made more feasible and acceptable to a range of users.

Two of the participants ceased participation in the study after the baseline phase; the reasons were unrelated to the intervention but were related to scheduling issues in their daily life. However, it is important to determine who is most likely to benefit from the intervention and how to educate individuals about it in a way that increases the chance of people making an informed choice of whether they are able to try it or have the time to take part in a course of treatment. It could also be explored whether the intervention could have an impact on the frequency of intrusive memories with fewer guided intervention sessions, which would reduce the participation load.

Owing to the COVID-19 pandemic, the delivery of the intervention was changed from face to face to remote (eg, web-based communication via Kara Connect). Originally, our plan was to gradually move toward remote delivery in future studies. When the pandemic struck, it forced the change to occur more quickly during the intervention phase of the study. This turned out to be a positive development, as the data did not suggest that the intervention became less effective by being delivered remotely. Recent research by Singh et al [39] indicated that this novel intervention delivered remotely could be an acceptable method to reduce the number of intrusive memories among health care staff. The number of intrusive memories was reduced to 0 at the 5-week follow-up in all 3 participants [39]. Continued remote delivery using a web-based platform instead of face-to-face delivery will be important in future studies [40], thereby removing geographical restraints and making it possible to reach people regardless of where they live or whether they are in quarantine (eg, owing to the COVID-19 pandemic) [22].

Conclusions

Targeting established intrusive memories of trauma that participants had been experiencing for some years (eg, from childhood sexual abuse) with a brief visuospatial intervention, involving a brief memory reminder and Tetris gameplay with mental rotation, seems to show promise for further exploration as a method to reduce their frequency. These early data suggest that the intervention might also result in symptom reduction related to posttraumatic stress, anxiety and depression, and improved functioning; however, further studies are needed. Because of its simplicity, this intervention might be capable of removing common barriers to existing treatment options after trauma, such as for PTSD, including some patients' reluctance to talk about and describe their trauma in detail to a therapist, high costs, and a limited number of qualified therapists [9]. The intervention might even be delivered by nonexperts in evidence-based trauma-focused therapy after brief training, with ongoing supervision—something that should be further examined.

The results of this study are encouraging, and the effects of the intervention on the number of intrusive memories need to be further explored. Continuing to develop this kind of scalable intervention is crucial to reach a large number of people in need of treatment after experiencing trauma. Future research should further examine the feasibility and acceptability of remote delivery by nonexperts in mental health (rather than only qualified clinical psychologists) and whether fewer intervention sessions can yield similar results. Randomized controlled trials are required to assess the intervention.

Acknowledgments

AB received funding from the Icelandic Research Fund (11709-0270). EAH received funding from the Oak Foundation (OCAY-18-442) and the Swedish Research Council (VR 2020-00873) in support of this study; EAH also received funding from AFA Insurance (200342) and the Lupina Foundation. UV received funding to establish the Stress-And-Gene-Analysis (SAGA) cohort from the European Research Council (Stress Gene, grant 726413) and Icelandic Research Fund (grant 163362-051). EBT reports funding from the Icelandic Research Fund (185287-051). LS received funding from the Swiss National Science Foundation (P2BEP1_184378) and a Thunberg Fellowship from the Swedish Collegium for Advanced Study. The authors would like to thank Beau Gamble for statistical assistance and for making graphs for the visual inspection of the primary outcome data.

Conflicts of Interest

EAH served on the board of charity MQ: Transforming Mental Health (United Kingdom). She received book royalties from Oxford University Press (Imagery in Cognitive Therapy) and Guilford Press (Imagery-Based Cognitive Therapy for Bipolar Disorder and Mood Instability) and occasional fees from clinical workshops and conference keynotes.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5). 5th edition. Washington, DC, USA: American Psychiatric Association; 2013.
2. Kessler RC, Aguilar-Gaxiola S, Alonso J, Benjet C, Bromet EJ, Cardoso G, et al. Trauma and PTSD in the WHO world mental health surveys. Eur J Psychotraumatol 2017 Oct 27;8(sup5):1353383 [FREE Full text] [doi: 10.1080/20008198.2017.1353383] [Medline: 29075426]
3. Santiago PN, Ursano RJ, Gray CL, Pynoos RS, Spiegel D, Lewis-Fernandez R, et al. A systematic review of PTSD prevalence and trajectories in DSM-5 defined trauma exposed populations: intentional and non-intentional traumatic events. PLoS One 2013 Apr 11;8(4):e59236 [FREE Full text] [doi: 10.1371/journal.pone.0059236] [Medline: 23593134]

4. Kupfer DJ, Regier DA. Neuroscience, clinical evidence, and the future of psychiatric classification in DSM-5. Am J Psychiatry 2011 Jul;168(7):672-674. [doi: 10.1176/appi.ajp.2011.11020219] [Medline: 21724672]
5. McLaughlin KA, Koenen KC, Friedman MJ, Ruscio AM, Karam EG, Shahly V, et al. Subthreshold posttraumatic stress disorder in the world health organization world mental health surveys. Biol Psychiatry 2015 Feb 15;77(4):375-384 [FREE Full text] [doi: 10.1016/j.biopsych.2014.03.028] [Medline: 24842116]
6. Thorarinsdottir K, Holmes EA, Hardarson J, Hedinsdottir U, Kanstrup M, Singh L, et al. Reducing intrusive memories of childhood trauma using a visuospatial intervention: case study in Iceland. JMIR Form Res 2021 Nov 04;5(11):e29873 [FREE Full text] [doi: 10.2196/29873] [Medline: 34734830]
7. Lewis C, Roberts NP, Andrew M, Starling E, Bisson JI. Psychological therapies for post-traumatic stress disorder in adults: systematic review and meta-analysis. Eur J Psychotraumatol 2020 Mar 10;11(1):1729633 [FREE Full text] [doi: 10.1080/20008198.2020.1729633] [Medline: 32284821]
8. Post-traumatic stress disorder. NICE Guidance. 2018 Dec 5. URL: https://www.nice.org.uk/guidance/ng116 [accessed 2021-03-17]
9. Kantor V, Knefel M, Lueger-Schuster B. Perceived barriers and facilitators of mental health service utilization in adult trauma survivors: a systematic review. Clin Psychol Rev 2017 Mar;52:52-68 [FREE Full text] [doi: 10.1016/j.cpr.2016.12.001] [Medline: 28013081]
10. Lewis C, Roberts NP, Gibson S, Bisson JI. Dropout from psychological therapies for post-traumatic stress disorder (PTSD) in adults: systematic review and meta-analysis. Eur J Psychotraumatol 2020 Mar 9;11(1):1709709 [FREE Full text] [doi: 10.1080/20008198.2019.1709709] [Medline: 32284816]
11. Imel ZE, Laska K, Jakupcak M, Simpson TL. Meta-analysis of dropout in treatments for posttraumatic stress disorder. J Consult Clin Psychol 2013 Jun;81(3):394-404 [FREE Full text] [doi: 10.1037/a0031474] [Medline: 23339535]
12. Najavits LM. The problem of dropout from "gold standard" PTSD therapies. F1000Prime Rep 2015 Apr 2;7:43 [FREE Full text] [doi: 10.12703/P7-43] [Medline: 26097716]
13. Singh L, Espinosa L, Ji JL, Moulds ML, Holmes EA. Developing thinking around mental health science: the example of intrusive, emotional mental imagery after psychological trauma. Cogn Neuropsychiatry 2020 Sep;25(5):348-363. [doi: 10.1080/13546805.2020.1804845] [Medline: 32847486]
14. Holmes EA, James EL, Coode-Bate T, Deeprose C. Can playing the computer game "Tetris" reduce the build-up of flashbacks for trauma? A proposal from cognitive science. PLoS One 2009;4(1):e4153 [FREE Full text] [doi: 10.1371/journal.pone.0004153] [Medline: 19127289]
15. Holmes EA, Hales SA, Young K, Di Simplicio M. Imagery-Based Cognitive Therapy for Bipolar Disorder and Mood Instability. New York, NY, USA: The Guilford Press; 2019.
16. Iyadurai L, Blackwell SE, Meiser-Stedman R, Watson PC, Bonsall MB, Geddes JR, et al. Preventing intrusive memories after trauma via a brief intervention involving Tetris computer game play in the emergency department: a proof-of-concept randomized controlled trial. Mol Psychiatry 2018 Mar;23(3):674-682 [FREE Full text] [doi: 10.1038/mp.2017.23] [Medline: 28348380]
17. Horsch A, Vial Y, Favrod C, Harari MM, Blackwell SE, Watson P, et al. Reducing intrusive traumatic memories after emergency caesarean section: a proof-of-principle randomized controlled study. Behav Res Ther 2017 Jul;94:36-47 [FREE Full text] [doi: 10.1016/j.brat.2017.03.018] [Medline: 28453969]
18. Kanstrup M, Singh L, Göransson KE, Widoff J, Taylor RS, Gamble B, et al. Reducing intrusive memories after trauma via a brief cognitive task intervention in the hospital emergency department: an exploratory pilot randomised controlled trial. Transl Psychiatry 2021 Jan 11;11(1):30 [FREE Full text] [doi: 10.1038/s41398-020-01124-6] [Medline: 33431807]
19. Kessler H, Holmes EA, Blackwell SE, Schmidt AC, Schweer JM, Bücker A, et al. Reducing intrusive memories of trauma using a visuospatial interference intervention with inpatients with posttraumatic stress disorder (PTSD). J Consult Clin Psychol 2018 Dec;86(12):1076-1090. [doi: 10.1037/ccp0000340] [Medline: 30507232]
20. Kanstrup M, Kontio E, Geranmayeh A, Olofsdotter Lauri K, Moulds ML, Holmes EA. A single case series using visuospatial task interference to reduce the number of visual intrusive memories of trauma with refugees. Clin Psychol Psychother 2021 Jan;28(1):109-123. [doi: 10.1002/cpp.2489] [Medline: 32525244]
21. Iyadurai L, Hales SA, Blackwell SE, Young K, Holmes EA. Targeting intrusive imagery using a competing task technique: a case study. Behav Cogn Psychother 2020 Nov;48(6):739-744. [doi: 10.1017/S1352465820000296] [Medline: 32594968]
22. 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. Lancet Psychiatry 2020 Jun;7(6):547-560 [FREE Full text] [doi: 10.1016/S2215-0366(20)30168-1] [Medline: 32304649]
23. Tetris Holding. Electronic Arts. 2014. URL: https://www.ea.com/de-de/games/tetris/tetris [accessed 2021-03-17]
24. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
25. Lecrubier Y, Sheehan DV, Weiller E, Amorim P, Bonora I, Sheehan KH, et al. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. Eur Psychiatr 1997;12(5):224-231. [doi: 10.1016/S0924-9338(97)83296-8]

26. Sheehan DV, Lecrubier Y, Harnett Sheehan KH, Janavs J, Weiller E, Keskiner A, et al. The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. Eur Psychiatr 1997;12(5):232-241. [doi: 10.1016/s0924-9338(97)83297-x]
27. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The posttraumatic stress disorder checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. J Trauma Stress 2015 Dec;28(6):489-498. [doi: 10.1002/jts.22059] [Medline: 26606250]
28. Wortmann JH, Jordan AH, Weathers FW, Resick PA, Dondanville KA, Hall-Clark B, et al. Psychometric analysis of the PTSD Checklist-5 (PCL-5) among treatment-seeking military service members. Psychol Assess 2016 Nov;28(11):1392-1403. [doi: 10.1037/pas0000260] [Medline: 26751087]
29. 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]
30. Kroenke K. Enhancing the clinical utility of depression screening. CMAJ 2012 Feb 21;184(3):281-282 [FREE Full text] [doi: 10.1503/cmaj.112004] [Medline: 22231681]
31. 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]
32. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 06;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
33. Toussaint A, Hüsing P, Gumz A, Wingenfeld K, Härter M, Schramm E, et al. Sensitivity to change and minimal clinically important difference of the 7-item Generalized Anxiety Disorder Questionnaire (GAD-7). J Affect Disord 2020 Mar 15;265:395-401. [doi: 10.1016/j.jad.2020.01.032] [Medline: 32090765]
34. Leon AC, Olfson M, Portera L, Farber L, Sheehan DV. Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. Int J Psychiatry Med 1997;27(2):93-105. [doi: 10.2190/T8EM-C8YH-373N-1UWD] [Medline: 9565717]
35. Coles T, Coon C, DeMuro C, McLeod L, Gnanasakthy A. Psychometric evaluation of the Sheehan Disability Scale in adult patients with attention-deficit/hyperactivity disorder. Neuropsychiatr Dis Treat 2014 May 19;10:887-895 [FREE Full text] [doi: 10.2147/NDT.S55220] [Medline: 24899807]
36. Bjornsson AS, Hardarson JP, Valdimarsdottir AG, Gudmundsdottir K, Tryggvadottir A, Thorarinsdottir K, et al. Social trauma and its association with posttraumatic stress disorder and social anxiety disorder. J Anxiety Disord 2020 May;72:102228 [FREE Full text] [doi: 10.1016/j.janxdis.2020.102228] [Medline: 32361167]
37. Holmes EA, James EL, Kilford EJ, Deeprose C. Key steps in developing a cognitive vaccine against traumatic flashbacks: visuospatial Tetris versus verbal Pub Quiz. PLoS One 2010 Nov 10;5(11):e13706 [FREE Full text] [doi: 10.1371/journal.pone.0013706] [Medline: 21085661]
38. Holmes EA, Ghaderi A, Eriksson E, Lauri KO, Kukacka OM, Mamish M, et al. 'I can't concentrate': a feasibility study with young refugees in Sweden on developing science-driven interventions for intrusive memories related to trauma. Behav Cogn Psychother 2017 Mar;45(2):97-109 [FREE Full text] [doi: 10.1017/S135246581600062X] [Medline: 28229806]
39. Singh L, Kanstrup M, Depa K, Falk A, Lindström V, Dahl O, et al. Digitalizing a brief intervention to reduce intrusive memories of psychological trauma for health care staff working during COVID-19: exploratory pilot study with nurses. JMIR Form Res 2021 May 26;5(5):e27473 [FREE Full text] [doi: 10.2196/27473] [Medline: 33886490]
40. Gamble B, Depa K, Holmes EA, Kanstrup M. Digitalizing a brief intervention to reduce intrusive memories of psychological trauma: qualitative interview study. JMIR Ment Health 2021 Feb 22;8(2):e23712 [FREE Full text] [doi: 10.2196/23712] [Medline: 33616540]

Abbreviations

CAPS-5: Clinician-Administered Posttraumatic Stress Disorder Checklist-5 Scale for Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
CTR: clinical trial registration
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
GAD-7: Generalized Anxiety Disorder-7 scale
MINI: Mini International Neuropsychiatric Interview
PCL-5: Posttraumatic Stress Disorder Checklist-5
PHQ-9: Patient Health Questionnaire-9
PTSD: posttraumatic stress disorder
REDCap: Research Electronic Data Capture
SAGA: Stress-And-Gene-Analysis
SDS: Sheehan Disability Scale

Edited by A Mavragani; submitted 18.02.22; peer-reviewed by H Kessler, R Lystad; comments to author 27.04.22; revised version received 18.05.22; accepted 18.05.22; published 20.07.22.

Please cite as:
Thorarinsdottir K, Holmes EA, Hardarson J, Stephenssen ES, Jonasdottir MH, Kanstrup M, Singh L, Hauksdottir A, Halldorsdottir T, Gudmundsdottir B, Thordardottir E, Valdimarsdottir U, Bjornsson A
Using a Brief Mental Imagery Competing Task to Reduce the Number of Intrusive Memories: Exploratory Case Series With Trauma-Exposed Women
JMIR Form Res 2022;6(7):e37382
URL: https://formative.jmir.org/2022/7/e37382
doi:10.2196/37382
PMID:35857368

Original Paper

The Stroke Recovery in Motion Implementation Planner: Mixed Methods User Evaluation

Jessica Reszel[1,2], RN, MScN; Joan van den Hoek[1], BNSc; Tram Nguyen[1,3], PhD; Gayatri Aravind[4], PhD; Mark T Bayley[5,6], MD; Marie-Louise Bird[7,8], PhD; Kate Edwards[7], BSc, MSc; Janice J Eng[8], PhD; Jennifer L Moore[9,10], PT, DHS; Michelle L A Nelson[4,11,12], PhD; Michelle Ploughman[13], PhD; Julie Richardson[14], PhD; Nancy M Salbach[6,15], PhD; Ada Tang[14], PhD; Ian D Graham[1,2,3], PhD

[1]Ottawa Hospital Research Institute, Ottawa, ON, Canada
[2]School of Nursing, University of Ottawa, Ottawa, ON, Canada
[3]School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada
[4]March of Dimes Canada, Toronto, ON, Canada
[5]Division of Physical Medicine and Rehabilitation, University of Toronto, Toronto, ON, Canada
[6]The KITE Research Institute, University Health Network, Toronto, ON, Canada
[7]College of Health and Medicine, University of Tasmania, Tasmania, Australia
[8]Department of Physical Therapy, University of British Columbia, Vancouver, BC, Canada
[9]South Eastern Norway Regional Knowledge Translation Center, Sunnaas Rehabilitation Hospital, Oslo, Norway
[10]Institute for Knowledge Translation, Carmel, IN, United States
[11]Lunenfeld-Tanenbaum Research Institute, Sinai Health, Toronto, ON, Canada
[12]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[13]Faculty of Medicine, Memorial University of Newfoundland, St John's, NL, Canada
[14]School of Rehabilitation Science, McMaster University, Hamilton, ON, Canada
[15]Department of Physical Therapy, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Ian D Graham, PhD
Ottawa Hospital Research Institute
501 Smyth Road
Ottawa, ON, K1H 8L6
Canada
Phone: 1 613 737 8899
Email: igraham@ohri.ca

Abstract

Background: As more people are surviving stroke, there is a growing need for services and programs that support the long-term needs of people living with the effects of stroke. Exercise has many benefits; however, most people with stroke do not have access to specialized exercise programs that meet their needs in their communities. To catalyze the implementation of these programs, our team developed the Stroke Recovery in Motion Implementation Planner, an evidence-informed implementation guide for teams planning a community-based exercise program for people with stroke.

Objective: This study aimed to conduct a user evaluation to elicit user perceptions of the usefulness and acceptability of the Planner to inform revisions.

Methods: This mixed methods study used a concurrent triangulation design. We used purposive sampling to enroll a diverse sample of end users (program managers and coordinators, rehabilitation health partners, and fitness professionals) from three main groups: those who are currently planning a program, those who intend to plan a program in the future, and those who had previously planned a program. Participants reviewed the Planner and completed a questionnaire and interviews to identify positive features, areas of improvement, value, and feasibility. We used descriptive statistics for quantitative data and content analysis for qualitative data. We triangulated the data sources to identify Planner modifications.

Results: A total of 39 people participated in this study. Overall, the feedback was positive, highlighting the value of the Planner's comprehensiveness, tools and templates, and real-world examples. The identified areas for improvement included clarifying the

need for specific steps, refining navigation, and creating more action-oriented content. Most participants reported an increase in knowledge and confidence after reading the Planner and reported that using the resource would improve their planning approach.

Conclusions: We used a rigorous and user-centered process to develop and evaluate the Planner. End users indicated that it is a valuable resource and identified specific changes for improvement. The Planner was subsequently updated and is now publicly available for community planning teams to use in the planning and delivery of evidence-informed, sustainable, community-based exercise programs for people with stroke.

(JMIR Form Res 2022;6(7):e37189) doi:10.2196/37189

KEYWORDS

stroke; rehabilitation; community-based exercise programs; knowledge translation; knowledge mobilization; implementation science

Introduction

Community-Based Exercise Programs for People With Stroke

There are >13 million new cases of stroke per year worldwide [1], and 1 in 4 adults aged >25 years will experience a stroke in their lifetime [2]. Advances in acute stroke treatment have significantly reduced mortality; however, this increased survival rate has led to more people living with chronic stroke-related disabilities. With stroke now being a leading cause of long-term disability [3,4], rehabilitation researchers have identified enhancing brain recovery and promoting long-term healthy behaviors as a priority; this research has generated a wealth of new evidence-based stroke recovery practices [5,6]. However, there is a need to move this evidence into practice and close the gap between best and current stroke rehabilitation practices [7].

Although many individuals see improvements during the acute rehabilitation phase after the stroke, many lose their initial gains when they return to the community, and their disability progresses over time [8]. Evidence suggests that exercise improves motor function [9,10], health-related quality of life [11,12], cognitive function [13,14], and cardiovascular risk factors [15,16] in those with stroke. The implementation of community-based exercise programs, which are defined as "structured, instructional programs of exercise for groups or individuals delivered outside the public health care setting and available in community settings," are avenues for engaging in lifelong physical activity [17]. Community programs that are focused primarily on walking, such as outdoor walking or mall walking programs, are often difficult to follow for individuals with mobility impairment from stroke, and many traditional fitness facilities and health clubs have accessibility barriers [18] that present additional challenges for people with stroke. Thus, people living with stroke may require adaptations to meet their unique needs and abilities. Stroke exercise programs should incorporate functional tasks that mimic daily activities, be guided by trained personnel knowledgeable in stroke and stroke-related impairments, and be delivered with an appropriate instructor-to-participant ratio [17]. Moreover, pre-exercise medical clearance by a health care provider and further eligibility screening by exercise providers are recommended to ensure the safety and appropriateness of adapted programs [17]. Despite the established benefits of ongoing exercise and evidence-based recommendations in the design and delivery of community-based exercise programs for stroke [17], most people with stroke do not have access to such community-based exercise programs that provide the specialized support to meet their long-term needs.

In 2016, the Heart and Stroke Foundation's Canadian Partnership for Stroke Recovery (CPSR) [19] convened its Knowledge Translation Advisory Committee to identify priority areas for knowledge translation. The committee, comprising people with stroke, caregivers, stroke recovery experts, health care providers, policy makers, and knowledge translation and mobilization experts, identified poststroke exercise as a high priority and specifically identified the need to develop sustainable evidence-based and community-based exercise programs for people with stroke. Within Canada and internationally, researchers and clinicians have developed various community-based exercise programs for people with stroke [20-22]; there is now a need to catalyze the implementation of these evidence-based approaches to optimize the health and social benefits for people with stroke.

Development of the Stroke Recovery in Motion Implementation Planner

Building on this momentum, our team aimed to develop an evidence-informed resource [23] to guide community program planners (eg, program managers and coordinators, rehabilitation health partners, fitness professionals, people with stroke, and caregivers) through the process of planning for the successful implementation of community-based exercise programs for people with stroke. We used a multistep process over 3 years to develop the Stroke Recovery in Motion Implementation Planner (hereafter referred to as the "Planner"; Figure 1).

The well-established Knowledge-to-Action (KTA) cycle, which helps bring the results of health care research into effective changes in practice, provided the overarching framework for the planning process [24-26]. It also builds on the CAN-IMPLEMENT guideline adaptation process, which divides the KTA cycle into 3 substantive planning phases [27,28]. Furthermore, the planning model incorporates elements of the Implementation Roadmap, which is based on the KTA cycle, and further breaks the 3 planning phases into practical steps and activities to facilitate implementation planning and execution [29]. The planning process is underpinned by 6 guiding principles (Textbox 1).

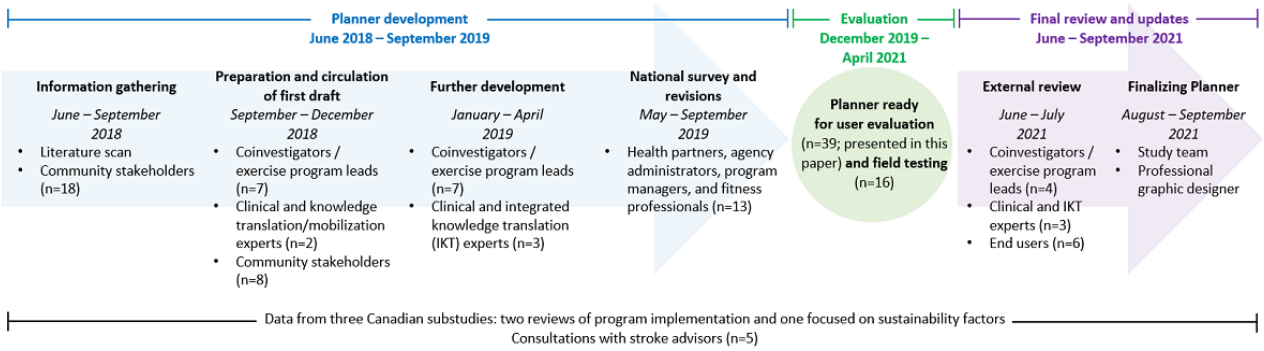
The "Planner development" phase (Figure 1) informed early decisions related to the content and organization of the Planner.

For example, we identified the type of information and tools that would be most helpful to end users, decided on the use of a generic approach that could be applied to any stroke-specific community-based exercise program, worked to include real-world examples, and reduced technical language. As part of this development process, we conducted a national survey of potential end users across Canada (health partners, agency administrators, program managers, and fitness professionals). Of the 21 invited people, 13 (62%) reviewed the initial Planner prototype and completed a web-based questionnaire on what they liked and what could be improved.

As the prototype was developed, the research team engaged stroke advisors, including members from an existing group of Stroke Community Advisors through the CPSR. These stroke advisors (3 people with stroke and 2 caregivers of an individual with stroke) reviewed the Planner and provided feedback on the content and format to ensure that people with stroke, their families, and caregivers were reflected appropriately.

On completing this rigorous initial Planner development process, the resource was deemed *ready for evaluation (user evaluation and field testing)*.

Figure 1. Summary of the Planner development process and stakeholders involved in the process.



Textbox 1. Implementation Planning Roadmap guiding principles underpinning the Stroke Recovery in Motion Implementation Planner.

Characteristics of the planning approach:

- Intended for exercise programs *situated within the community* and provided by organizations with a mandate for community service (vs provision of individual therapeutic care)
- *Participant centered* (putting people with stroke and caregivers at the center of decisions and seeing them as experts working with service providers to achieve the best outcomes) [30]
- *Participatory and inclusive* (people with stroke and other relevant stakeholders, including health care partners, involved in cocreating the implementation plan)
- *Evidence-informed* (uses effective approaches to planning and implementation and incorporates the use of local data in making decisions)
- Aimed at *strengthening participant health outcomes*
- Focused on *sustaining successful programs*

Study Objectives

The purpose of the study was to conduct a dual-component evaluation of the Planner comprising a user evaluation and field test. The objective of the user evaluation was to elicit user perceptions of the usefulness and acceptability of the Planner and revise the Planner based on the data. The objective of the field test was to describe how teams used the Planner in real-world conditions; describe the effects of using the Planner on participants' implementation planning knowledge, attitudes, and activities; and identify factors influencing Planner use [31].

The goal of this paper (part 1) is to describe the results of the user evaluation and the revisions we subsequently made to the Planner. The results of the field test are reported separately in part 2 [31].

To guide the reporting of this study, we used guidelines for mixed methods studies (GRAMMS [Good Reporting of a Mixed Methods Study]) [32], qualitative studies (COREQ [Consolidated Criteria for Reporting Qualitative Studies]) [33], and survey studies (CROSS [Checklist for Reporting of Survey Studies]) [34]. These checklists are provided in Multimedia Appendix 1 [32-34].

Methods

Design

The user evaluation was a mixed methods study that used a concurrent triangulation design [35]. We used a convergent model where cross-sectional questionnaire data (closed and open-ended questions) and interview and focus group data were collected and analyzed separately, with the quantitative and qualitative findings merged during the interpretive phase [35]. The quantitative and qualitative data were assigned equal weight. The quantitative data facilitated identifying the "what"; specifically, it helped us identify what aspects of the Planner were most frequently identified as needing modification or removal, which enabled prioritization of essential changes. The

qualitative data helped us understand the "why" of the participant feedback and provided insight into how we could best address participant concerns as we revised the Planner. The qualitative data also provided additional ideas to strengthen the Planner, which may not have been captured in the structured questionnaire. The settings for this study were Canada and Australia.

Sampling and Recruitment

We used purposeful sampling. All staff involved or interested in the planning and delivery of community-based exercise programs for people with stroke were eligible to participate in the user evaluation, including community and municipal program directors, managers and coordinators, regional health authority staff, fitness or exercise professionals, physiotherapists, and other consulting health partners. We identified potential participants through the professional networks of (1) study coinvestigators, many of whom have developed exercise programs; (2) individuals who previously participated in the Planner development process; and (3) participants enrolled in the study (ie, snowball sampling).

In Canada, we aimed to identify participants from different geographical regions with various population densities. Using definitions from Statistics Canada [36], we created a matrix based on geographical region and population size category. As recruitment progressed, we aimed to identify individuals from the remaining undersampled cells (region × urban or rural) in our recruitment matrix. As the Canadian study progressed, we identified an emergent opportunity to include participants from Australia. One of the coinvestigators returned to Tasmania, Australia, where there was a state goal to increase community-based exercise opportunities for people with stroke. Unlike Canada, Tasmania had low rates of COVID-19 and fewer pandemic restrictions during the study period, and community-based program planning and implementation proceeded as usual. This gave us a unique opportunity to use the professional network of the coinvestigator, who was situated in an urban area of Tasmania, to identify participants who were currently actively engaged in the planning and delivery of community-based exercise programs for people with stroke. We consulted and received approval from our funding partner (CPSR) for the addition of participants from Australia, as there was perceived value in having an international perspective on the Planner.

We started recruiting individuals to prospectively field test the Planner in December 2019. In March 2020, the COVID-19 pandemic caused significant challenges for recruitment as the pandemic resulted in many teams stopping community-based program planning. In May 2020, we amended our protocol to add 2 additional groups (Table 1) to our study with different levels of experience in planning community-based exercise programs for people with stroke, allowing data collection on the Planner within pandemic restrictions. Study enrollment was completed in February 2021. The eligibility criteria for each participant group are presented in Table 1.

All 3 groups participated in the user evaluation component. The *current* program planners engaged in additional study activities as part of the field test component of the evaluation [31].

The study staff contacted potential participants via email with study information. If there was no response, we sent 2 follow-up reminders. Interested participants connected with the study staff to review the requirements, confirm eligibility, obtain informed written consent, and schedule a time for the study activities.

Table 1. Eligibility criteria and data collection methods for the 3 groups of participants in the study.

Participant group	Inclusion criteria	Data collection methods
Current program planners	• Interested in implementing a community-based exercise program for people with stroke in the next 6 to 12 months • Are willing to use the Planner to guide their planning process	• Questionnaire • Baseline interview or focus group • Monitoring interviews • End-of-study interview or focus group
Future program planners[a]	• Have a vested interest in community-based exercise programs for people with stroke and the development of a useful resource for program planning • Have not previously launched a community-based exercise program for people with stroke and are not currently considering planning a program	• Questionnaire • Follow-up interview or focus group
Past program planners[a]	• Have previously implemented a community-based exercise program for people with stroke in the past 1 to 5 years	• Interview or focus group about past experience • Questionnaire • Follow-up interview or focus group

[a]New participant group added in May 2020.

Data Collection

The study participants completed a web-based questionnaire and a minimum of 1 interview. The order of the activities and the content of the questionnaire and interviews were tailored to each group (Table 1). Participants received an honorarium at a rate of CAD $25 (US $19.5) per hour to compensate for their time.

Questionnaire

Each participant completed a web-based questionnaire created in LimeSurvey (LimeSurvey GmbH) [37]. This was a "restricted" questionnaire, meaning that only participants enrolled in the study could complete the questionnaire through a single-use unique URL generated by the research team. There was no validated instrument that met our needs; therefore, the

core study team created a questionnaire that was specific to the Planner content during the development phase (Figure 1). It was tested internally with 2 other research team members for functionality and clarity and then administered to 13 end users who represented people who were currently running programs, had run programs in the past, or were interested stakeholders. We then modified the questionnaire to tailor the content to the 3 participant groups and optimize functionality (eg, adding branching logic). Participants were required to read the Planner before starting the questionnaire. Depending on the group, the questionnaire had between 79 and 89 closed-ended questions, most of which also included open textboxes to expand on their selected answers, and between 11 and 22 open-ended questions. Questions were spread across 7 sections and focused on participants' impressions of the Planner sections and tools (keep, modify, or remove), format and presentation, value of the Planner for community program planners, impact of the Planner on knowledge and confidence, and questions on respondents' demographics and experience. The estimated time to read the Planner and complete the questionnaire was 4 hours. Multimedia Appendix 2 presents the questionnaires administered to the 3 study groups.

Interviews and Focus Groups

Each participant completed at least one interview or focus group. Some participants took part in the study as a team and therefore chose to complete a focus group together. The semistructured interview and focus group guides were developed by the research team and informed through discussions with end users in the Planner development phase. Question topics included overall impressions of the Planner, likes and dislikes regarding the Planner, in-depth discussion about the questionnaire responses, and perceived feasibility of the planning process. In addition, *past* program planners were asked to compare their current experience using the Planner with their previous experience. Multimedia Appendix 3 presents the semistructured interview and focus group guides.

Current program planners completed brief "monitoring" interviews that were scheduled every 1 to 2 months with the research staff. Discussion points during these interviews included how they had been using the Planner, what they liked about using it, and what was missing.

Interviews and focus groups were conducted either in person, through video calls, or by phone, depending on participant preference. Sessions were facilitated by 1 of 4 female researchers (1 master's degree–prepared nurse researcher [J Reszel], 1 PhD-prepared rehabilitation researcher [TN], 1 master's degree–prepared nutrition researcher [KE], and 1 PhD-prepared physiotherapy researcher [MLB]) experienced in qualitative data collection. None of them had pre-existing relationships with the participants. The interviewers worked to create a nonjudgmental environment and welcomed both positive and negative feedback. On average, the interviews and focus groups lasted 44 (range 21-106) minutes, and the monitoring interviews lasted 24 (range 13-39) minutes. The interviews were audio recorded and transcribed verbatim. Altogether, there were 42.9 hours of audio recordings, yielding 912 pages of transcripts. Field notes were made after the interviews and focus groups to document researcher reflections and observations, including participant interactions.

Data Analysis

Questionnaire

We analyzed the data using descriptive statistics in SPSS (version 27; IBM Corp) [38]. For nominal data (eg, gender and geographical location) and ordinal data (eg, confidence and knowledge), we calculated frequencies and percentages to illustrate the distribution of the responses across categories. For continuous data (eg, years of experience), we calculated measures of central tendency (eg, mean and median) and measures of variability (eg, SD and range). The text in the open-ended questions was analyzed by grouping content into categories to identify what participants liked and disliked about the Planner steps, tools, format, and planning approach. Given the small sample size, no comparisons were made based on role, organization, or geography.

Interviews and Focus Groups

We used conventional content analysis, an inductive approach through which the codes and categories emerged from the data [39]. Each transcript was verified against the audio recording, read as a whole, and then segments of the text were labeled with codes in Microsoft Word. As the analysis progressed, we continued to develop our coding scheme while adding emerging codes. To enhance the trustworthiness of our findings, 40% of the transcripts were coded independently by 2 researchers [40]. The 2 researchers met regularly to compare their coding, resolve discrepancies, and update the coding scheme. All the transcripts were revisited to apply the final version of the coding scheme. We reached inductive thematic saturation [41] as no new codes were added after the 31st transcript (out of 67 transcripts). Data analysis occurred concurrently with data collection, with interview probes evolving to explore emerging themes from the analysis. The team reviewed and discussed field notes, and no notable group dynamics were identified. Although transcripts were not returned to participants, in the final stages of the study, 4 interview participants reviewed and shared their impressions of the revised Planner. This participant check allowed the study team to assess the extent to which we successfully applied the study findings to the Planner.

Triangulation of Quantitative and Qualitative Data to Inform Planner Revisions

On completion of the descriptive analysis of the questionnaire data and coding of the qualitative data, we created a master data summary document that included both data sets to facilitate comparing and contrasting. The core research team met regularly during this triangulation phase (>40 hours of meetings) and discussed the similarities and differences between the qualitative and quantitative data.

The team focused on changes to be made in the Planner, grounded in the study data. First, to identify "essential" changes, we reviewed the questionnaire data by using a threshold of 75%, a commonly used figure to define consensus [42]. Where <75% of respondents (in any of the 3 participant groups) selected the most positive response option for a question (ie, keep as is;

strongly agree or agree), we carefully reviewed the open-ended questionnaire responses and qualitative interview data for that Planner section to determine the required changes. Next, even when consensus was reached (ie, ≥75%), we still reviewed every comment provided in the questionnaire and interviews to identify other opportunities to enhance the Planner. Decisions regarding whether to address the suggestions were based on (1) alignment with the guiding principles underpinning the Planner (Textbox 1), (2) alignment with evidence from implementation science (eg, best and promising practices for implementation) and pedagogical science (eg, value of repetition for learning), (3) the significance of the suggestions, and (4) the feasibility of the changes related to formatting and design constraints.

Ethics Approval

We obtained ethics approval from the Ottawa Health Science Network Research Ethics Board (protocol 20190594-01H) and the Tasmania Health and Medical Human Research Ethics Committee (project ID 23559) for the initial and amended protocols. All participants signed an informed consent form before starting any study activities, and all study procedures were conducted in accordance with privacy and confidentiality requirements.

Results

Overview

We enrolled 39 participants between December 2019 and February 2021. We contacted 27 *current* program planners, of whom 16 (59%) enrolled; 14 *future* program planners, of whom 9 (64%) enrolled; and 43 *past* program planners, of whom 14 (33%) enrolled. The participant demographic data are presented in Table 2, participant roles based on employment setting are presented in Table 3, and the types of data collected during the study are presented in Table 4.

Table 2. Participant demographics (N=36)[a].

Variable	Current program planners (n=15)	Future program planners (n=9)	Past program planners (n=12)	All participants
Gender, n (%)[b]				
Female	4 (57)	8 (89)	10 (83)	22 (79)
Male	3 (43)	1 (11)	2 (17)	6 (21)
Gender fluid	0 (0)	0 (0)	0 (0)	0 (0)
Location of community, n (%)				
Alberta	1 (7)	3 (33)	0 (0)	4 (11)
British Columbia	4 (27)	0 (0)	2 (17)	6 (17)
Manitoba	0 (0)	2 (22)	0 (0)	2 (6)
Newfoundland and Labrador	1 (7)	0 (0)	0 (0)	1 (3)
Nova Scotia	0 (0)	0 (0)	1 (8)	1 (3)
Ontario	5 (33)	3 (33)	8 (67)	16 (44)
Prince Edward Island	0 (0)	1 (11)	1 (8)	2 (6)
Tasmania, Australia	4 (27)	0 (0)	0 (0)	4 (11)
Population density, n (%)				
Rural or mostly rural	6 (40)	1 (11)	4 (33)	11 (31)
Urban or mostly urban	7 (47)	7 (78)	5 (42)	19 (53)
Combination of rural and urban	2 (13)	1 (11)	3 (25)	6 (17)
Size of community, n (%)				
<5000	0 (0)	0 (0)	1 (8)	1 (3)
5000-9999	5 (33)	0 (0)	0 (0)	5 (14)
10,000-24,999	4 (27)	1 (11)	1 (8)	6 (17)
25,000-50,000	1 (7)	0 (0)	0 (0)	1 (3)
>50,000	5 (33)	8 (89)	10 (83)	23 (64)
Type of organization where the program is or will be offered, n (%)[c]				
Community recreation center (public and municipal)	5 (33)	6 (67)	5 (42)	16 (42)
YMCA[d]	1 (7)	0 (0)	5 (42)	6 (17)
Community health center	1 (7)	2 (22)	1 (8)	4 (11)
Recreation center for older adults	0 (0)	1 (11)	1 (8)	2 (6)
Physiotherapy clinic	3 (20)	1 (11)	1 (8)	5 (14)
Nursing home	0 (0)	0 (0)	1 (8)	1 (3)
Retirement residence	0 (0)	0 (0)	2 (17)	2 (6)
Private gym or facility	0 (0)	3 (11)	2 (17)	3 (8)
Family health team	4 (27)	0 (0)	0 (0)	4 (11)
Nonprofit community space	0 (0)	0 (0)	3 (25)	3 (8)
Web-based program	2 (13)	1 (11)	1 (8)	4 (11)
Number of years of experience in community program planning and/or delivery[b]				
Values, median (range)	2 (0-21)	5 (0-20)	10 (1-20)	7 (0-21)
Values, mean (SD)	5 (8)	8 (6)	10 (6)	8 (7)
Individual role in planning or delivering this exercise program, n (%)[e]				
Provider agency administration	0 (0)	1 (11)	1 (8)	2 (6)

Variable	Current program planners (n=15)	Future program planners (n=9)	Past program planners (n=12)	All participants
Program manager or coordinator	7 (47)	2 (22)	4 (33)	13 (36)
Fitness or exercise professional	2 (13)	2 (22)	3 (25)	7 (19)
Rehabilitation health professional	6 (40)	4 (44)	4 (33)	14 (39)
Experience in *planning* this type of program, n (%)				
Previous experience planning adapted or specialized fitness programs	10 (67)	6 (67)	12 (100)	28 (78)
Experience in *delivering* this type of program, n (%)				
Previous experience delivering adapted or specialized fitness programs	12 (80)	6 (67)	8 (67)	26 (72)
Current confidence in planning an adapted or specialized fitness program, n (%)				
Extremely confident	0 (0)	1 (11)	1 (8)	2 (6)
Very confident	5 (33)	4 (44)	6 (50)	15 (42)
Moderately confident	8 (53)	3 (33)	5 (42)	16 (44)
Slightly confident	2 (13)	1 (1)	0 (0)	3 (8)
Not at all confident	0 (0)	0 (0)	0 (0)	0 (0)
Knowledge of how to use evidence to inform decision-making in program planning, n (%)				
Extremely knowledgeable	3 (20)	0 (0)	0 (0)	3 (8)
Very knowledgeable	4 (27)	8 (89)	3 (25)	15 (42)
Moderately knowledgeable	2 (13)	1 (11)	8 (67)	11 (31)
Slightly knowledgeable	4 (27)	0 (0)	1 (8)	5 (14)
Not at all knowledgeable	2 (13)	0 (0)	0 (0)	2 (6)

[a]In this study, of the 39 participants, 36 (92%) completed the questionnaire. Of the 36 participants who completed the questionnaire, 35 (97%) questionnaires were complete, with only 1 (3%) participant skipping 2 out of 86 questions (the skipped questions are not reported in this table).

[b]The first version of the current program planner questionnaire did not include this question; therefore, 8 responses are missing.

[c]Respondents could select >1 response option.

[d]YMCA: Young Men's Christian Association.

[e]Some participants may actually represent >1 group (eg, a rehabilitation professional who is working as a program coordinator); however, these data reflect how participants self-identified their primary role in planning as per the questionnaire responses.

Table 3. Role of participants based on their employment setting (N=39).

Employment setting	Participant role, n (%)				Total, N
	Provider agency administration	Program manager or coordinator	Fitness or exercise professional	Rehabilitation health professional	
Community-based nonprofit[a]	2 (20)	4 (40)	4 (40)	0 (0)	10
Municipality[b]	0 (0)	3 (60)	2 (40)	0 (0)	5
Health authority[c]	0 (0)	2 (40)	0 (0)	3 (60)	5
Private practice[d]	0 (0)	1 (17)	2 (33)	3 (50)	6
Primary care[e]	0 (0)	1 (25)	0 (0)	3 (75)	4
University	0 (0)	0 (0)	0 (0)	4 (100)	4
Hospital	0 (0)	1 (50)	0 (0)	1 (50)	2
Stroke network[f]	0 (0)	2 (67)	0 (0)	1 (33)	3
Total	2 (5)	14 (36)	8 (21)	15 (38)	39

[a]For example, the Young Men's Christian Association (YMCA).

[b]For example, a city.

[c]For example, a provincial, state, or regional health authority.

[d]For example, a physiotherapy clinic or gym.

[e]For example, a family health team.

[f]For example, a provincial or regional network.

Table 4. Summary of quantitative and qualitative data collected by participant groups (N=39).

Data collection method	Current program planners (n=16[a])	Future program planners (n=9[b])	Past program planners (n=14[c])	All participants
Questionnaire	15 responses	9 responses	12 responses	36 responses
Interviews and focus groups	15 interviews and focus groups with 16 participants	9 interviews with 9 participants	25 interviews and focus groups with 14 participants	49 interviews and focus groups with 39 participants
Monitoring interviews	18 interviews with 10 participants	N/A[d]	N/A	18 interviews with 10 participants

[a]From 7 planning teams.

[b]From 9 sites.

[c]From 13 sites.

[d]N/A: not applicable.

Overall Impressions of the Planner

Overall, feedback on the Planner was positive. All questionnaire participants felt that the Planner addressed the key factors to consider when planning a community-based exercise program for people with stroke; almost all (34/36, 94%) felt it would help them make decisions informed by evidence, and most (29/36, 81%) indicated that it would improve their usual approach to planning:

> *The Planner is a very good teaching tool and I think you could easily work through that with a team. It would put your team all on the same page and certainly enhance some people's background...I think it would keep the group on the same page and have a good idea of what we're doing, why we're doing it, who we're doing it for, and the benefits.* [Fitness or exercise professional, ID38]

Most participants (26/28, 93%) reported that reading the Planner increased their knowledge of how to use evidence to inform decision-making in program planning. Nearly as many participants (24/28, 86%) also reported that reading the Planner increased their confidence in their ability to plan an adapted or specialized fitness program.

Several *past* program planners who had successfully launched and sustained programs commented on the alignment between the Planner and their own experiences and the added benefit of the Planner:

> *Initially it [the Planner] was overwhelming and I went "oh my gosh I missed every single step"...But then I looked back and I saw that I did most of it, but just not necessarily with as much intention or thought. A lot of it [Planner content] is just kind of intuitive and I did it but it wasn't as thoughtful.* [Rehabilitation health professional, ID37]

Nearly all *future* and *past* program planners (19/21, 91%) indicated that they would likely use the Planner for future program planning. Similarly, most questionnaire participants (33/36, 92%) said they would likely recommend it to colleagues to support program planning, described by one participant as follows:

> *If I knew somebody was thinking of [planning a program], I think a resource such as this would be the exact one they should be using. If I were in a position where I was implementing or supporting or advising on the implementation of a program, I would definitely recommend the use of this resource at the planning stage in order to make sure that everything has been thought about.* [Rehabilitation health professional, ID21]

Most questionnaire participants (34/36, 94%) agreed that the Planner could be applied to planning programs that were not stroke specific, including exercise programs for people with other health conditions and nonexercise programs:

> *I'm working on a different program and some of the concepts in the Planner have helped develop that program. It's not an exercise program but the Planner has helped me think about other things when it comes to programming.* [Program manager or coordinator, ID12]

Although the most commonly reported first impression of the Planner was that it was long and possibly overwhelming, over three-quarters (28/36, 78%) of the questionnaire participants agreed that the Planner presented the right amount of information. Most participants indicated that after reading and digesting the information, they saw great value in all the presented materials and could not identify materials that could be removed. Participants recognized the need to strike a balance between making the Planner a comprehensive information source for people with various roles and experience levels while also ensuring that the material was not overly long or onerous to read:

> *As far as usability, it's trying to walk that fine line between providing too much information and not enough information. Because you get a variety of people, from those who have never implemented a community-based program, to those who are very used to that. Trying to make it work for both those groups, how do you do that as best as possible? That's not an easy answer.* [Program manager or coordinator, ID13]

All questionnaire participants (36/36, 100%) agreed that the Planner was well organized, and nearly all agreed that it was easy to read and understand (35/36, 97%) and clearly presented the planning process (33/36, 92%). Participants were split on the format of the Planner: 56% (20/36) of participants would have preferred a web-based version over a paper-based version, and 28% (10/36) did not have a preference. Participants frequently desired access to both a web-based version for easy navigation and tool completion and a paper-based version for hard copy use.

Feedback That Prompted Changes to the Planner

Planner Content

There were several recurring comments related to the Planner content, which cut across sections and tools that resulted in edits. For example, we observed a discordance in some participant settings between the Implementation Planning Roadmap and the "usual way things are done," reflecting a difference in philosophy. To address this, we made the guiding principles and assumptions explicit at the beginning of the Planner to make the approach and values transparent and explain the rationale for the planning process (Textbox 1). Furthermore, some participants expressed concerns about the theoretical language used in the Planner. We carefully reviewed the Planner and simplified the technical terms wherever possible.

For questions assessing specific Planner sections and tools, 70% (30/43) of the items were deemed "necessary—keep as is" by at least 75% of the questionnaire participants (Multimedia Appendix 4). Key content changes were made to the Planner and tools based on the participant feedback (Table 5).

Table 5. Examples of key content changes made to the Planner and tools.

Identified area of improvement	Changes made to the Planner and tools
Include more information to clarify why specific steps and activities are important to complete during implementation planning (eg, forming planning partnership, decision-making methods, terms of reference, celebrating the launch, and preparing an evaluation plan)	• "Why is this important" statements were emphasized throughout the Planner to provide the rationale and potential benefits of completing the step or activity
Include more examples of the real-world solutions used by other teams to address planning challenges; include examples of completed tools from planning teams	• Addition of the "Tips and Potholes" section at the end of each planning phase to highlight the success factors and challenges encountered by teams involved in the development and evaluation of the Planner • Added samples of completed tools created by study sites (with permission)
Wherever possible, make content action oriented	• Implementation Planning Roadmap revised from 13 steps to 8 steps and Planner guidance edited to provide greater clarity and focus on specific activities and tasks to complete • All tools reviewed and edited to ensure templates provide concrete guidance • Creation of standardized cover sheets for each tool, which include "Why is this important?" and "How to use this tool" statements
Include information on how to consider the specific needs of people with stroke or caregivers as planning partners	• New section and tool with specific guidance on factors to consider and questions to ask when engaging people with stroke and caregiver partners in the team • Voices of people with stroke and caregivers were brought to the forefront by inserting verbatim quotes collected during our evaluation throughout the Planner
Include more exercise program–specific information to facilitate program comparisons	• Creation of a "program comparison template" with guiding questions for planning teams to assess the history, attributes, and requirements of programs under consideration
Emphasize the importance of considering and addressing program sustainability factors early and often	• Sustainability information was included in all 3 phases of the Planner • Creation of a new section on sustainability capacity • Key sustainability factors identified in the end-of-phase checklists and throughout tools
Make tools concise (eg, implementation work plan and assessment of barriers) and avoid duplication between tools (eg, community assessments)	• Tool content reorganized, simplified, and relabeled to align more clearly with road map steps • Repetitive content merged and the number of tools reduced • Longer tools split into easy-to-manage sections (eg, identifying barriers to program, program users, and program setting became 3 short worksheets)

Planner Format and Organization

The participants offered constructive comments on how to optimize the format and organization of the Planner, which resulted in several key changes (Table 6).

Table 6. Examples of key format and organization changes made to the Planner and tools.

Identified area of improvement	Changes made to the Planner and tools
Simplify structure, balance the workload across the 3 phases, and reorder the sequence of activities and steps	• Implementation Planning Roadmap reduced and simplified from 13 steps to 8 steps • Implementation planning process reorganized to better balance planning activities within and across the 3 phases • Phase 2 and 3 steps reordered to make the planning sequence more logical (eg, developing an evaluation plan before launching the exercise program)
Improve navigation; clearly align Planner content with the phases and steps of the road map	• Road map figure moved to the start of the Planner as a key navigation element • Planner redesigned to better link content to road map phase and step and orient the reader to the location on the map • Professional graphic design concept developed to facilitate navigation
Facilitate different "starting points" in the Planner to help situate readers from different contexts and starting places in their planning journey	• Developed a new "Where do we start?" section in the Planner introduction to outline different planning scenarios and potential starting points and how to use the Planner accordingly • Directed readers to the progress checklists at the end of each phase to assess what work still needs to be completed
Keep the body of the Planner concise for easy reading	• Selected content (eg, additional resources and program samples) moved from the body of the Planner to the appendix as "Read more" sections for interested readers
Provide easy access to tools and appendices (additional resources); ensure tools are fillable and editable	• Tools summarized at end of each phase with links • PDF and original, editable files provided for easy download • Design concept to include both hard copy and web-based versions of the Planner

Feedback Considered but Existing Approach Maintained

The study participants made some suggestions for Planner modifications that after careful consideration, we decided not to make. Here, we provide 2 key examples with the rationale.

First, some participants requested that to facilitate use, the Planner should be separated into different sections to assign specific content to different roles on the planning team. However, the desire to distinguish between those "planning" and those "delivering" was not aligned with our guiding principle of using a participatory and inclusive approach. Planning teams are most effective when those delivering the program (eg, fitness or exercise professionals) and using the program (eg, people with stroke) are engaged early in the planning process. In addition, team members' roles may be fluid over time and in different contexts. Therefore, we left the Planner as one document designed for all team members, with the goal of encouraging awareness of and participation in the full planning process. The importance of an integrated planning team was supported by a study participant who regularly delivered community programming:

> *I was really excited about having "boots on the ground" kind of people [on the planning team]. Because often times you don't hear about these things until after others have made the decision and you're like "oh if only you had talked to the people who actually implement these things. It would have been helpful." Because the concept of policy can often vary greatly from the reality.* [Fitness or exercise professional, ID22]

Second, several participants commented on the level of repetition throughout the Planner and recommended making the document shorter and more concise. Although we carefully reviewed the Planner and eliminated unnecessary redundancies, some repetition was left in for pedagogical reasons [43]. Repetition helps readers remember and understand the information. We also recognized that some readers would not read the Planner from start to finish and skip from section to section; therefore, we chose to judiciously repeat essential contextual information throughout the Planner.

Positive Feedback on the Planner Supporting the Existing Approach and Resulting in Enhancements

All the participants provided positive feedback on the Planner and identified features that they found useful. We used this positive feedback to identify content and features to keep in the final version. We also identified ways of expanding upon and further enhancing these features where possible (Table 7).

Table 7. Sections and features of the Planner rated positively with illustrative quotes.

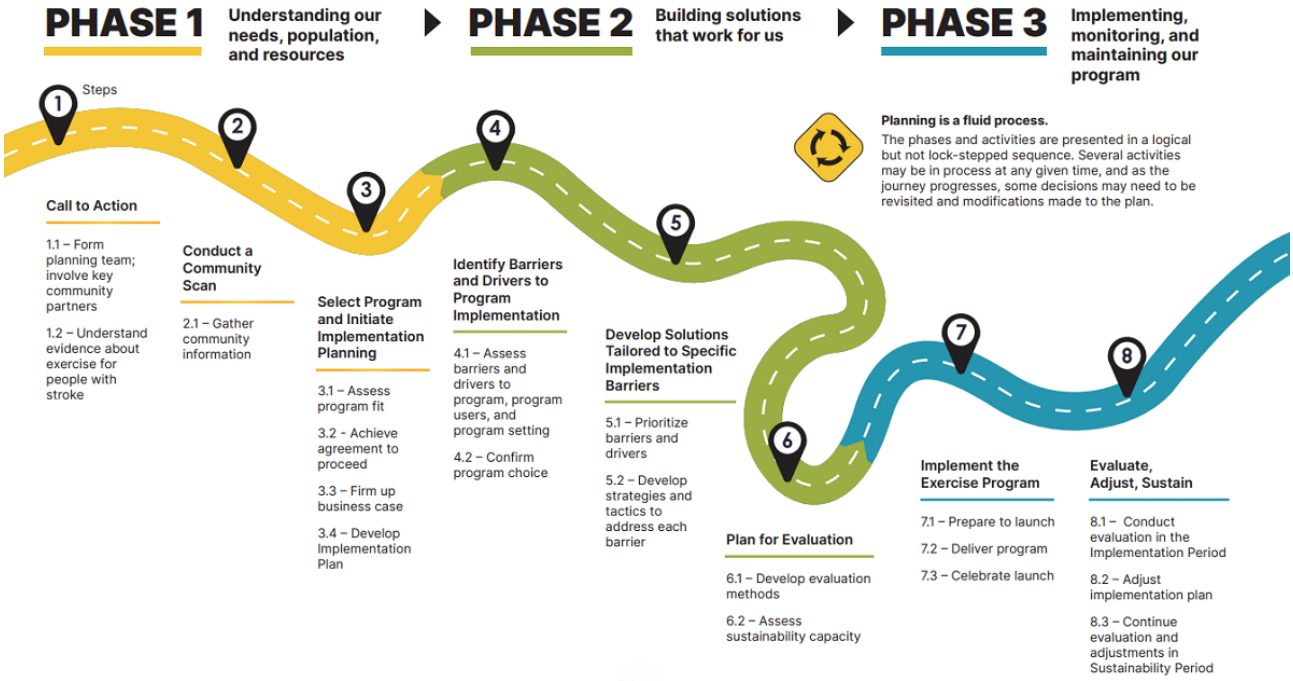
Positive feedback on the Planner	Illustrative quote	Planner decisions made based on positive feedback
Although many participants felt they would not necessarily need to use all Planner tools to implement every exercise program, they generally appreciated the inclusion of various tools, should they be needed.	"Although I feel all [tools] are important to keep, I don't feel I would use them all each time I would start a program. It would depend on the type of organization I was working with and how much detail would be needed, thus having all the tools available is important." [Rehabilitation health professional, ID32]	• Kept a variety of tools to meet the needs and contexts of different planning teams • Created cover pages for each tool, further highlighting who, how, when, and why planning teams can use the various tools
Although many participants felt the Planner was long, most participants appreciated the comprehensiveness of the Planner and the breadth of information presented.	"There's lots of information. You can go lots of places to look at program planning information, but having it all consolidated...is really helpful to me. Because I could get lost and I could go down a significant rabbit hole if I start Googling all this stuff on my own. To forego the Google rabbit hole is very helpful." [Program manager or coordinator, ID35]	• Kept the Planner as a comprehensive document to meet the needs of various planning team members • New content added based on participant feedback to improve comprehensiveness; for example, more details on developing a planning partnership, how to engage people with stroke and caregiver partners, and web-based program information
Nearly all participants commented positively on the summary checklists at the end of each phase as a clear way of assessing progress and the remaining planning tasks.	"I did like the progress checklists. I really liked that at the end of each section. It was a nice way to kind of bring all of that together and in a practical tool that people can use." [Rehabilitation health professional, ID33]	• Kept checklists at the end of each phase • Phase checklists were made into a separate tool for easy access and printing • Content of checklists was integrated into the implementation work plan
Many participants valued the quotes and field notes from other planning teams to learn about real-world successes and challenges and highlight the importance of the various planning steps.	"I liked the field notes about programs—this is what happened and this is the result...It makes it relatable; when you're reading all the info, it pulls you back into the practical side of it, which is good." [Fitness or exercise professional, ID31]	• New quotes from study participants added throughout the Planner • "Tips and Potholes" added to the end of each phase to further highlight study participant experiences and learnings

Planner Revisions

Using the data reported in this study, we revised both the content and format of the Planner. Integrating this feedback involved extensive rewriting and editing by the core study team; contracting the graphic designer; and a final review by 16 coinvestigators, consultants, end users, and stroke advisors. The largest revision of the Planner involved improving the overall structure and navigation (Figure 2).

The Stroke Recovery in Motion Implementation Planner is now hosted through March of Dimes Canada, a Canadian nonprofit organization offering services for people with disabilities, including the After Stroke program focused on stroke recovery in the community [44]. The Planner is free and publicly available here [23].

Figure 2. Implementation Planning Roadmap from the Planner summarizing the phases, steps, and activities.



Discussion

Principal Findings

This paper provides a comprehensive description of how the Planner was developed, informing consumers of the rigorous process used. Our mixed methods user evaluation demonstrated that, overall, end users viewed the Planner positively and indicated that it was a useful and valuable resource. They reported that it would improve the planning process and help them make planning decisions informed by evidence. The participants' constructive feedback on the content and organization was used to revise and strengthen the Planner.

Comparison With Prior Work

Our development process aligns with that used by others to develop implementation guides and toolkits in different settings, including selecting an underpinning theory or framework [45-48], searching the academic and gray literature [46-50], consulting experts and consumers [45-48,50], and refining the guide based on end user feedback [45,46,48-50].

Some participants provided feedback consistent with findings reported in the literature; for example, perceptions that the implementation process could be overwhelming or the guide too long [50,51], the need to reduce the number of steps and technical language [51], and the desire for new content and tools to address team needs [48,49,51]. Engaging a wide variety of stakeholders in the development process and using several user-centered design strategies ensured that the final product was grounded in the needs and experiences of those who will use it in real-world settings [52].

The use of an implementation framework can contribute to more systematic planning, delivery, and evaluation of programs, thereby contributing to improved success and sustainability [53]. However, many implementers lack knowledge and experience in using these frameworks [53], and many third-sector organizations (eg, voluntary and community organizations and social enterprises [54]) face capacity and capability issues when implementing evidence-based interventions [55]. Leeman et al [56] identified "tools" as a strategy for building implementation capacity in community-based practitioners. The Planner is an evidence-informed tool for building the capacity of practitioners (in this context, community program planners, health professionals, fitness professionals, people with stroke, and caregivers) to plan for implementation in an applied and approachable way. Most study participants reported an increase in knowledge and confidence after reading the Planner, including many who reported being experienced program planners.

The Planner is based on the KTA, CAN-IMPLEMENT, and Implementation Roadmap frameworks [24,25,27-29] and is grounded in implementation science and practice. Studies on third-sector organizations in general [55] and on poststroke exercise specifically [57-60] have revealed factors that can influence program implementation, such as equipment, space, time, staffing, training, funding, marketing and recruitment, class capacity, sustainability, program adaptation challenges, organizational culture and priorities, and collaboration between organizations and professionals [55,57-60]. The Planner is specifically designed to help planning teams identify these and other factors unique to their settings, which may impede or support the implementation of community-based exercise programs for people with stroke. The Planner offers strategies, including some recommended by others [55,57], to overcome these barriers. Most importantly, the Planner provides a step-by-step action-oriented road map to plan for successful implementation and sustainability.

Limitations and Strengths

This study has some limitations that should be acknowledged. A key challenge was conducting the study during the COVID-19 pandemic. Although we successfully recruited 39 participants to review the Planner, the data collection period coincided with the pandemic, a time of high personal and professional stress. It is unknown how the stress of the pandemic may have influenced decisions to participate in the evaluation or how it influenced participants' perspectives of the Planner. However, nearly half of the individuals approached to participate did so despite the burden of having to read the comprehensive Planner, complete a lengthy questionnaire, and participate in an interview. This suggests that those who participated were strongly committed to providing input on the Planner and provided thoughtful and detailed feedback.

Although our sample was diverse, we would have liked to have enrolled more fitness professionals, given their role in delivering exercise programs. There may be several reasons for their limited enrollment. For the *current* program planning group, our primary contact was often the manager or coordinator or health partner. These planning leads had sometimes not yet identified the fitness professional team members and were unable to connect the research team with fitness professional study candidates. Furthermore, the pandemic led to the closure of community centers and fitness facilities, and therefore, many fitness professionals were not actively employed during this time. Despite these challenges, approximately 20% (8/39) of the study participants were fitness or exercise professionals who offered rich and thoughtful guidance on how to improve the Planner to meet their needs.

Finally, the user evaluation involved the hard copy version of the Planner, and therefore, the findings may not reflect perceptions of a web-based format. Plans are underway to develop a web-based toolkit based on the hard copy version.

A strength of the study was the mixed methods design, which facilitated assessing perceptions of the Planner through a comprehensive questionnaire with standard questions for all participants, followed by an interview for in-depth discussions on dimensions that were particularly important to each participant. We used multiple types of triangulation, including methods (questionnaires and interviews), sources (participants from diverse settings), and analysts (coding and interpretation by multiple researchers), all of which enhanced the quality and credibility of our findings [61]. The targeted enrollment of stakeholders from 3 groups (*current*, *future*, and *past*) allowed us to collect data (and reach saturation) from a broad range of stakeholders in various geographical areas with differing experiences in community-based exercise programs for people with stroke. Our rigorous analysis process resulted in every comment being reviewed and carefully considered to inform the Planner revisions and facilitated improved relevance and feasibility of the Planner.

Future Directions

With the Planner finalized and freely and publicly available for use by community teams, we now have the opportunity to further evaluate its ongoing use and impact. Access and use will be monitored through website statistics and consumer inquiries. Working with key stakeholders, we also plan to augment the Planner to better address culturally tailored physical activity programs for racialized populations with stroke, as well as issues related to planning web-based programs for people with stroke. Finally, future work will involve developing products that distill the Planner information into alternate formats to meet diverse learning needs, including video vignettes, infographics, a condensed pocket guide, and a presentation slide deck.

Conclusions

Community-based exercise programs are urgently required to address community reintegration and transitions for people living with the effects of stroke. The Stroke Recovery in Motion Implementation Planner [23] was designed to address the limited use of evidence-informed planning practices for community-based exercise programs for people with stroke. Guided by knowledge from the field of implementation science on how to facilitate implementation, we used a rigorous process to develop and evaluate the Planner. The evaluation revealed that the Planner was perceived to be a valuable resource that may be used to guide interdisciplinary teams in the planning and delivery of evidence-informed, sustainable, community-based exercise programs for people with stroke.

Acknowledgments

The authors would like to thank all the study participants, who generously gave their time amidst the challenges of the COVID-19 pandemic to provide them with detailed and insightful input on the Stroke Recovery in Motion Implementation Planner. IDG is a recipient of a Canadian Institutes of Health Research Foundation grant (grant number 143237). NMS was supported by a Heart and Stroke Foundation Mid-Career Investigator award. AT was supported by a Clinician-Scientist Award (phase II) from the Ontario Heart and Stroke Foundation (P-19-TA-1192). The authors would like to acknowledge Dr Robert Teasell, Dr Sharon Straus, Dr Justin Presseau, Dr Elizabeth Inness, and Dr Elizabeth Lynch, who contributed to the design of the Stroke Recovery in Motion Implementation Planner. The authors would also like to acknowledge Patrick Faucher for the final graphic design of the Planner.

This project has been made possible by Brain Canada Foundation through the Canada Brain Research Fund, with the financial support of Health Canada and the Heart and Stroke Foundation Canadian Partnership for Stroke Recovery. The views expressed herein do not necessarily represent the views of the Minister of Health or the Government of Canada.

Authors' Contributions

IDG and MTB co-led all aspects of the study. IDG, MTB, JvdH, GA MLB, JJE, JLM, MLAN, MP, J Richardson, NMS, and AT contributed to the conceptualization and design of the study. J Reszel, TN, KE, and MLB collected the data; J Reszel, TN, JvdH, and IDG analyzed the questionnaire and interview data; and J Reszel drafted the initial manuscript. All authors participated in the writing of this manuscript and reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Reporting checklists.
[PDF File (Adobe PDF File), 221 KB - formative_v6i7e37189_app1.pdf]

Multimedia Appendix 2
Questionnaires for current, future, and past program planners.
[PDF File (Adobe PDF File), 642 KB - formative_v6i7e37189_app2.pdf]

Multimedia Appendix 3
Semistructured interview and focus group guides.
[PDF File (Adobe PDF File), 265 KB - formative_v6i7e37189_app3.pdf]

Multimedia Appendix 4
Questionnaire data on Planner sections and tools according to the study group.
[PDF File (Adobe PDF File), 296 KB - formative_v6i7e37189_app4.pdf]

References

1. GBD 2016 Stroke Collaborators. Global, regional, and national burden of stroke, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 May;18(5):439-458 [FREE Full text] [doi: 10.1016/S1474-4422(19)30034-1] [Medline: 30871944]
2. GBD 2016 Lifetime Risk of Stroke Collaborators, Feigin VL, Nguyen G, Cercy K, Johnson CO, Alam T, et al. Global, Regional, and Country-Specific Lifetime Risks of Stroke, 1990 and 2016. N Engl J Med 2018 Dec 20;379(25):2429-2437 [FREE Full text] [doi: 10.1056/NEJMoa1804492] [Medline: 30575491]
3. Stroke in Canada: Highlights from the Canadian Chronic Disease Surveillance System. Ottawa, ON: Public Health Agency of Canada; 2017. URL: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/stroke-canada-fact-sheet.html
4. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, American Heart Association Council on Epidemiology Prevention Statistics Committee Stroke Statistics Subcommittee. Heart disease and stroke statistics-2020 update: a report from the American Heart Association. Circulation 2020 Mar 03;141(9):e139-e596 [FREE Full text] [doi: 10.1161/CIR.0000000000000757] [Medline: 31992061]
5. Evidence-Based Review of Stroke Rehabilitation. Heart & Stroke Foundation Canadian Partnership for Stroke Recovery. 2018. URL: http://www.ebrsr.com//evidence-review [accessed 2022-02-08]
6. Teasell R, Salbach NM, Foley N, Mountain A, Cameron JI, Jong AD, et al. Canadian stroke best practice recommendations: rehabilitation, recovery, and community participation following stroke. 6th edition update 2019. Int J Stroke 2020 Oct;15(7):763-788. [doi: 10.1177/1747493019897843] [Medline: 31983296]
7. Eng JJ, Bird M, Godecke E, Hoffmann TC, Laurin C, Olaoye OA, et al. Moving stroke rehabilitation research evidence into clinical practice: consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. Int J Stroke 2019 Oct;14(8):766-773. [doi: 10.1177/1747493019873597] [Medline: 31564224]
8. Dhamoon MS, Moon YP, Paik MC, Sacco RL, Elkind MS. Trajectory of functional decline before and after ischemic stroke: the Northern Manhattan Study. Stroke 2012 Aug;43(8):2180-2184 [FREE Full text] [doi: 10.1161/STROKEAHA.112.658922] [Medline: 22649168]
9. Hasan SM, Rancourt SN, Austin MW, Ploughman M. Defining optimal aerobic exercise parameters to affect complex motor and cognitive outcomes after stroke: a systematic review and synthesis. Neural Plast 2016;2016:2961573 [FREE Full text] [doi: 10.1155/2016/2961573] [Medline: 26881101]
10. Mehta S, Pereira S, Janzen S, Mays R, Viana R, Lobo L, et al. Cardiovascular conditioning for comfortable gait speed and total distance walked during the chronic stage of stroke: a meta-analysis. Top Stroke Rehabil 2012;19(6):463-470. [doi: 10.1310/tsr1906-463] [Medline: 23192710]

11. Chen M, Rimmer JH. Effects of exercise on quality of life in stroke survivors: a meta-analysis. Stroke 2011 Mar;42(3):832-837. [doi: 10.1161/strokeaha.110.607747]
12. Rand D, Eng JJ, Tang P, Hung C, Jeng J. Daily physical activity and its contribution to the health-related quality of life of ambulatory individuals with chronic stroke. Health Qual Life Outcomes 2010 Aug 03;8(1):80. [doi: 10.1186/1477-7525-8-80] [Medline: 20682071]
13. Oberlin LE, Waiwood AM, Cumming TB, Marsland AL, Bernhardt J, Erickson KI. Effects of physical activity on poststroke cognitive function. Stroke 2017 Nov;48(11):3093-3100. [doi: 10.1161/strokeaha.117.017319]
14. Ploughman M, McCarthy J, Bossé M, Sullivan HJ, Corbett D. Does treadmill exercise improve performance of cognitive or upper-extremity tasks in people with chronic stroke? A randomized cross-over trial. Arch Phys Med Rehabil 2008 Nov;89(11):2041-2047. [doi: 10.1016/j.apmr.2008.05.017] [Medline: 18996231]
15. D'Isabella NT, Shkredova DA, Richardson JA, Tang A. Effects of exercise on cardiovascular risk factors following stroke or transient ischemic attack: a systematic review and meta-analysis. Clin Rehabil 2017 Dec 19;31(12):1561-1572. [doi: 10.1177/0269215517709051] [Medline: 28523989]
16. Brouwer R, Wondergem R, Otten C, Pisters MF. Effect of aerobic training on vascular and metabolic risk factors for recurrent stroke: a meta-analysis. Disabil Rehabil 2021 Jul;43(15):2084-2091. [doi: 10.1080/09638288.2019.1692251] [Medline: 31794269]
17. Inness EL, Brown G, Tee A, Kelly L, Moller J, Aravind G, et al. Canadian Stroke Community-based Exercise Recommendations 3rd Edition. Canada: Heart and Stroke Foundation Canadian Partnership for Stroke Recovery; 2021. URL: https://kite-uhn.com/can-stroke-community-based-exercise-recommendations
18. Rimmer JH, Riley B, Wang E, Rauworth A. Accessibility of health clubs for people with mobility disabilities and visual impairments. Am J Public Health 2005 Nov;95(11):2022-2028. [doi: 10.2105/AJPH.2004.051870] [Medline: 16254234]
19. Canadian Partnership for Stroke Recovery - About us. Canadian Partnership for Stroke Recovery. URL: http://canadianstroke.ca/index.php/en/about-us [accessed 2022-02-08]
20. Fitness and Mobility Exercise Program. FAME. 2021. URL: http://fameexercise.com/ [accessed 2022-02-08]
21. Fit for Function Stroke Wellness Program. Mobility Research. URL: http://mobilityresearch.ca/fit-for-function-2/ [accessed 2022-02-08]
22. Together In Movement and Exercise (TIMETM) Program. University Health Network. URL: https://www.uhn.ca/TorontoRehab/Clinics/TIME [accessed 2022-02-08]
23. Graham ID, van den Hoek J, Reszel J, Nguyen T, Aravind G, Bayley MT, et al. Stroke Recovery in Motion: Community-based exercise program implementation planner. Ottawa: Ottawa Hospital Research Institute; 2022. URL: http://www.afterstroke.ca/srimp [accessed 2022-06-09]
24. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, et al. Lost in knowledge translation: time for a map? J Contin Educ Health Prof 2006;26(1):13-24. [doi: 10.1002/chp.47] [Medline: 16557505]
25. Knowledge Translation Planner. Ottawa, ON: Government of Canada; 2017. URL: https://www.canada.ca/en/health-canada/corporate/about-health-canada/reports-publications/grants-contributions/knowledge-transfer-planner.html
26. Azer SA. The top-cited articles in medical education: a bibliometric analysis. Acad Med 2015 Aug;90(8):1147-1161. [doi: 10.1097/ACM.0000000000000780] [Medline: 26061861]
27. Harrison MB, van den Hoek J, Graham ID. CAN-Implement©: Planning for Best-Practice Implementation. Philadelphia, PA: Lippincott, Williams and Wilkins; 2014.
28. Harrison MB, Graham ID, van den Hoek J, Dogherty EJ, Carley ME, Angus V. Guideline adaptation and implementation planning: a prospective observational study. Implement Sci 2013 May 08;8(1):49 [FREE Full text] [doi: 10.1186/1748-5908-8-49] [Medline: 23656884]
29. Harrison MB, Graham ID. Knowledge Translation in Nursing and Healthcare: A Roadmap to Evidence-informed Practice. Oxford, UK: Wiley Blackwell Publishing; 2021.
30. What is person-centred care and why is it important? Health Innovation Network. URL: https://healthinnovationnetwork.com/system/ckeditor_assets/attachments/41/what_is_person-centred_care_and_why_is_it_important.pdf [accessed 2022-02-08]
31. Reszel J, van den Hoek J, Nguyen T, Aravind G, Bayley MT, Bird ML, et al. How community-based teams use the Stroke Recovery in Motion Implementation Planner: longitudinal qualitative field test study. JMIR Form Res 2022;6(7):e37243. [doi: 10.2196/37243]
32. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
33. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
34. Sharma A, Minh Duc NT, Luu Lam Thang T, Nam NH, Ng SJ, Abbas KS, Jacqz-Aigrain, et al. A consensus-based Checklist for Reporting of Survey Studies (CROSS). J Gen Intern Med 2021 Oct;36(10):3179-3187. [doi: 10.1007/s11606-021-06737-1] [Medline: 33886027]

35. Creswell J, Plano Clark V, Gutmann M, Hanson W. Advances in mixed methods research designs. In: Tashakkori A A, Teddlie C C, editors. Handbook of Mixed Methods in Social & Behavioral Research. Thousand Oaks, CA: SAGE; 2003:209-240.
36. Population centre and rural area classification 2016. Statistics Canada. URL: https://www.statcan.gc.ca/eng/subjects/standard/pcrac/2016/introduction#s7 [accessed 2022-02-08]
37. LimeSurvey: An Open Source survey tool. LimeSurvey GbmH. URL: http://www.limesurvey.org/ [accessed 2022-02-08]
38. IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM Corp; 2020.
39. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
40. Lincoln YS, Guba E. Naturalistic Inquiry. Thousand Oaks, California: SAGE Publications; 1985.
41. 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. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
42. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol 2014 Apr;67(4):401-409. [doi: 10.1016/j.jclinepi.2013.12.002] [Medline: 24581294]
43. Bruner RF. Repetition is the first principle of all learning. SSRN. 2001. URL: https://papers.ssrn.com/abstract=224340 [accessed 2022-02-08]
44. After Stroke. March of Dimes Canada. URL: https://www.afterstroke.ca/ [accessed 2022-02-08]
45. Dollar KM, Kirchner JE, DePhilippis D, Ritchie MJ, McGee-Vincent P, Burden JL, et al. Steps for implementing measurement-based care: implementation planning guide development and use in quality improvement. Psychol Serv 2020 Aug;17(3):247-261. [doi: 10.1037/ser0000368] [Medline: 31318240]
46. Glowacki K, Arbour-Nicitopoulos K, Burrows M, Chesick L, Heinemann L, Irving S, et al. It's more than just a referral: development of an evidence-informed exercise and depression toolkit. Mental Health Physical Activity 2019 Oct;17:100297. [doi: 10.1016/j.mhpa.2019.100297]
47. Salbach NM, MacKay-Lyons M, Solomon P, Howe J, McDonald A, Bayley MT, et al. The role of theory to develop and evaluate a toolkit to increase clinical measurement and interpretation of walking speed and distance in adults post-stroke. Disabil Rehabil 2021 Jan 16:1-17 (forthcoming). [doi: 10.1080/09638288.2020.1867653] [Medline: 33459080]
48. Murray E, May C, Mair F. Development and formative evaluation of the e-Health Implementation Toolkit (e-HIT). BMC Med Inform Decis Mak 2010 Oct 18;10:61 [FREE Full text] [doi: 10.1186/1472-6947-10-61] [Medline: 20955594]
49. Thoele K, Ferren M, Moffat L, Keen A, Newhouse R. Development and use of a toolkit to facilitate implementation of an evidence-based intervention: a descriptive case study. Implement Sci Commun 2020 Oct 06;1(1):86 [FREE Full text] [doi: 10.1186/s43058-020-00081-x] [Medline: 33043301]
50. O'Toole JK, Starmer AJ, Calaman S, Campos M, Goldstein J, Hepps J, et al. I-PASS mentored implementation handoff curriculum: implementation guide and resources. MedEdPORTAL 2018 Aug 03;14:10736 [FREE Full text] [doi: 10.15766/mep_2374-8265.10736] [Medline: 30800936]
51. Vanderbom K, Eisenberg Y. Implementing inclusive programming tailored to the community context. In: Proceedings of the APHA 2020. 2020 Presented at: APHA 2020; Oct 24-28, 2020; Virtual.
52. Dopp AR, Parisi KE, Munson SA, Lyon AR. A glossary of user-centered design strategies for implementation experts. Transl Behav Med 2019 Nov 25;9(6):1057-1064. [doi: 10.1093/tbm/iby119] [Medline: 30535343]
53. Lynch EA, Mudge A, Knowles S, Kitson AL, Hunter SC, Harvey G. "There is nothing so practical as a good theory": a pragmatic guide for selecting theoretical approaches for implementation projects. BMC Health Serv Res 2018 Nov 14;18(1):857 [FREE Full text] [doi: 10.1186/s12913-018-3671-z] [Medline: 30428882]
54. Nelson ML, Armas A, Thombs R, Singh H, Fulton J, Cunningham HV, et al. Synthesising evidence regarding hospital to home transitions supported by volunteers of third sector organisations: a scoping review protocol. BMJ Open 2021 Jul 05;11(7):e050479 [FREE Full text] [doi: 10.1136/bmjopen-2021-050479] [Medline: 34226235]
55. Bach-Mortensen AM, Lange BC, Montgomery P. Barriers and facilitators to implementing evidence-based interventions among third sector organisations: a systematic review. Implement Sci 2018 Jul 30;13(1):103 [FREE Full text] [doi: 10.1186/s13012-018-0789-7] [Medline: 30060744]
56. Leeman J, Calancie L, Hartman MA, Escoffery CT, Herrmann AK, Tague LE, et al. What strategies are used to build practitioners' capacity to implement community-based interventions and are they effective?: a systematic review. Implement Sci 2015 May 29;10(1):80 [FREE Full text] [doi: 10.1186/s13012-015-0272-7] [Medline: 26018220]
57. Salbach NM, Howe JA, Baldry D, Merali S, Munce SEP. Considerations for expanding community exercise programs incorporating a healthcare-recreation partnership for people with balance and mobility limitations: a mixed methods evaluation. BMC Res Notes 2018 Apr 02;11(1):214 [FREE Full text] [doi: 10.1186/s13104-018-3313-x] [Medline: 29609662]
58. Moncion K, Biasin L, Jagroop D, Bayley M, Danells C, Mansfield A, et al. Barriers and facilitators to aerobic exercise implementation in stroke rehabilitation: a scoping review. J Neurol Phys Ther 2020 Jul;44(3):179-187. [doi: 10.1097/NPT.0000000000000318] [Medline: 32516297]

59. Gaskins NJ, Bray E, Hill JE, Doherty PJ, Harrison A, Connell LA. Factors influencing implementation of aerobic exercise after stroke: a systematic review. Disabil Rehabil 2021 Aug 25;43(17):2382-2396. [doi: 10.1080/09638288.2019.1704075] [Medline: 31875459]
60. Skrastins O, Tsotsos S, Aqeel H, Qiang A, Renton J, Howe J, et al. Fitness coordinators' and fitness instructors' perspectives on implementing a task-oriented community exercise program within a healthcare-recreation partnership for people with balance and mobility limitations: a qualitative study. Disabil Rehabil 2020 Sep;42(19):2687-2695. [doi: 10.1080/09638288.2019.1570357] [Medline: 30739500]
61. Patton MQ. Enhancing the quality and credibility of qualitative analysis. Health Serv Res 1999 Dec;34(5 Pt 2):1189-1208 [FREE Full text] [Medline: 10591279]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Studies
CPSR: Canadian Partnership for Stroke Recovery
CROSS: Checklist for Reporting of Survey Studies
GRAMMS: Good Reporting of a Mixed Methods Study
KTA: Knowledge-to-Action

Edited by A Mavragani; submitted 12.02.22; peer-reviewed by S Castiglione, F Fischer, C Moore; comments to author 25.04.22; revised version received 10.06.22; accepted 11.06.22; published 29.07.22.

Please cite as:
Reszel J, van den Hoek J, Nguyen T, Aravind G, Bayley MT, Bird ML, Edwards K, Eng JJ, Moore JL, Nelson MLA, Ploughman M, Richardson J, Salbach NM, Tang A, Graham ID
The Stroke Recovery in Motion Implementation Planner: Mixed Methods User Evaluation
JMIR Form Res 2022;6(7):e37189
URL: https://formative.jmir.org/2022/7/e37189
doi:10.2196/37189
PMID:35904870

Original Paper

How Community-Based Teams Use the Stroke Recovery in Motion Implementation Planner: Longitudinal Qualitative Field Test Study

Jessica Reszel[1,2], RN, MScN; Joan van den Hoek[1], BNSc; Tram Nguyen[1,3], PhD; Gayatri Aravind[4], PhD; Mark T Bayley[5,6], MD; Marie-Louise Bird[7,8], PhD; Kate Edwards[7], BSc, MSc; Janice J Eng[8], PhD; Jennifer L Moore[9,10], PT, DHS; Michelle L A Nelson[4,11,12], PhD; Michelle Ploughman[13], PhD; Julie Richardson[14], PhD; Nancy M Salbach[6,15], PhD; Ada Tang[14], PhD; Ian D Graham[1,2,3], PhD

[1]Ottawa Hospital Research Institute, Ottawa, ON, Canada
[2]School of Nursing, University of Ottawa, Ottawa, ON, Canada
[3]School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada
[4]March of Dimes Canada, Toronto, ON, Canada
[5]Division of Physical Medicine and Rehabilitation, University of Toronto, Toronto, ON, Canada
[6]The KITE Research Institute, University Health Network, Toronto, ON, Canada
[7]College of Health and Medicine, University of Tasmania, Tasmania, Australia
[8]Department of Physical Therapy, University of British Columbia, Vancouver, ON, Canada
[9]South Eastern Norway Regional Knowledge Translation Center, Sunnaas Rehabilitation Hospital, Oslo, Norway
[10]Institute for Knowledge Translation, Carmel, IN, United States
[11]Lunenfeld-Tanenbaum Research Institute, Sinai Health, Toronto, ON, Canada
[12]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[13]Faculty of Medicine, Memorial University of Newfoundland, St John's, NL, Canada
[14]School of Rehabilitation Science, McMaster University, Hamilton, ON, Canada
[15]Department of Physical Therapy, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Ian D Graham, PhD
Ottawa Hospital Research Institute
501 Smyth Road
Ottawa, ON, K1H 8L6
Canada
Phone: 1 613 737 8899
Email: igraham@ohri.ca

Abstract

Background: The Stroke Recovery in Motion Implementation Planner guides teams through the process of planning for the implementation of community-based exercise programs for people with stroke, in alignment with implementation science frameworks.

Objective: The purpose of this study was to conduct a field test with end users to describe how teams used the Planner in real-world conditions; describe the effects of Planner use on participants' implementation-planning knowledge, attitudes, and activities; and identify factors influencing the use of the Planner.

Methods: This field test study used a longitudinal qualitative design. We recruited teams across Canada who intended to implement a community-based exercise program for people with stroke in the next 6 to 12 months and were willing to use the Planner to guide their work. We completed semistructured interviews at the time of enrollment, monitoring calls every 1 to 2 months, and at the end of the study to learn about implementation-planning work completed and Planner use. The interviews were analyzed using conventional content analysis. Completed Planner steps were plotted onto a timeline for comparison across teams.

Results: We enrolled 12 participants (program managers and coordinators, rehabilitation professionals, and fitness professionals) from 5 planning teams. The teams were enrolled in the study between 4 and 14 months, and we conducted 25 interviews. We observed that the teams worked through the planning process in diverse and nonlinear ways, adapted to their context. All teams

provided examples of how using the Planner changed their implementation-planning knowledge (eg, knowing the steps), attitudes (eg, valuing community engagement), and activities (eg, hosting stakeholder meetings). We identified team, organizational, and broader contextual factors that hindered and facilitated uptake of the Planner. Participants shared valuable *tips from the field* to help future teams optimize use of the Planner.

Conclusions: The Stroke Recovery in Motion Implementation Planner is an adaptable resource that may be used in diverse settings to plan community-based exercise programs for people with stroke. These findings may be informative to others who are developing resources to build the capacity of those working in community-based settings to implement new programs and practices. Future work is needed to monitor the use and understand the effect of using the Planner on exercise program implementation and sustainability.

(JMIR Form Res 2022;6(7):e37243) doi:10.2196/37243

KEYWORDS

knowledge translation; knowledge mobilization; implementation science; implementation planning; stroke; rehabilitation; capacity building; community-based exercise programs

Introduction

Background

With more people surviving stroke [1], community-based exercise programs [2-4] have emerged to provide safe and effective exercise opportunities for people living with the effects of stroke. Although participating in exercise after stroke conveys broad benefits [5], most people with stroke lack access to these specialized exercise programs in their own communities. As a result of this gap, the Canadian Partnership for Stroke Recovery identified increasing implementation and sustainability of these community-based exercise programs as a knowledge translation priority [6,7].

Planning and implementing new evidence-based practices and programs is a complex process requiring specific knowledge and skills. There is a growing body of literature and resources aimed at building the capacity for implementation science, but there is less support for building the capacity of those actually practicing implementation [8-10]. Practitioners may not have the knowledge or skills to apply implementation science theories or frameworks in their work [11,12], nor do they report feeling confident in their ability to complete key steps in implementation planning, such as conducting a barriers and facilitators assessment [12]. Community organizations in particular may face barriers to implementing evidence-based practices, including a lack of implementation expertise [13].

As a means to build implementation-planning capacity for community-based exercise programs for people with stroke, our team developed an implementation-planning guide, the Stroke Recovery in Motion Implementation Planner (hereafter referred to as the Planner) [14]. The Planner was based on established knowledge translation and implementation frameworks, including the Knowledge-to-Action framework [15], CAN-IMPLEMENT [16,17], and the Implementation Planning Roadmap [18]. The focus of the Planner is on implementation *planning*, which is a process that turns strategy into action through three phases: phase 1: establishing a diverse, interdisciplinary planning team and working together to understand the community population, their needs, potential program options, and the available resources; phase 2: conducting a barriers and drivers assessment and developing tailored solutions (implementation strategies), as well as building an evaluation plan; and phase 3: launching, monitoring, and maintaining the exercise program.

Many strategies promoting uptake of evidence-based practices focus on modifying the individual or their environment. Throughout the Planner development and evaluation process [7], we prioritized the design of the product itself by applying a user-centered design approach [19]. User-centered design is an iterative process whereby the needs and context of the intended users are central to informing the content and design of the product [19,20]. An example of a user-centered design strategy is to field test the product. Through prolonged engagement with end users who provide feedback on their experience with the product, field testing facilitates an understanding of how the product is being applied in real-world settings [21]. Such an exploratory approach can help to challenge the prototype and subsequently lead to meaningful changes [22], thereby creating a product that is more functional, acceptable, and effective. Upon completing a prototype of the Planner [7], we conducted a field test study with end users to explore how the Planner was used and adapted in diverse settings.

Objectives

This study was part of a larger research program that developed and evaluated the Planner, which included a cross-sectional user evaluation of the Planner by diverse stakeholders (reported elsewhere [7]) and the field test reported here. The objectives of the field test were to (1) describe how teams used the Planner in real-world conditions; (2) describe the effects of using the Planner on participants' implementation-planning knowledge, attitudes, and activities; and (3) identify factors influencing use of the Planner.

Methods

Design

This field test study used a longitudinal qualitative design [23,24], which facilitated in-depth discussions about use and impressions of the Planner over time. We used the Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist

[25] to inform the reporting of this study (Multimedia Appendix 1).

Participants and Setting

Using the professional networks of the study team, we identified primary contacts (ie, team leads tasked with program planning) at Canadian community organizations that were intending to implement a community-based exercise program for people with stroke in the next 6 to 12 months (*current* planners) and willing to use the Planner to guide their ongoing planning. We welcomed representatives from several occupations, including program managers and coordinators, health partners (eg, physiotherapists and occupational therapists), and fitness professionals. As the purpose of this study was to observe how teams used the Planner to guide planning processes, teams were required to be at an early stage in their planning. A research team member (J Reszel) contacted potential participants by email with study information, with up to two reminders. If the contact was interested in participating, a memorandum of understanding was signed between the participating organization and our research institution. We asked the contact person to nominate other planning team members (ie, snowball sampling) whom we could invite to take part in the study. All participants signed an individual consent form. Participants did not receive any incentives to test the Planner or fund their program initiatives; however, participants were provided an honorarium to compensate them for time spent on study activities.

Initially, we aimed to enroll 9 to 12 teams to field test the Planner; however, the onset of the COVID-19 pandemic resulted in significant disruptions to community-based program planning and the closure of many facilities. This recruitment challenge caused us to lower our enrollment to 5 teams prepared to use the Planner to guide program planning during this period.

Data Collection

The field test teams participated in 3 core data collection points. After reading the Planner and completing a questionnaire [7], all participants completed a *baseline* interview or focus group to discuss their initial impressions of the planning process. Next, research staff conducted *monitoring interviews* (ideally every 1-2 months) with a primary contact for each team to discuss planning work completed, Planner sections and tools used, and feedback on what was helpful and what was not. Each monitoring call started with the researcher summarizing the last call, allowing the participants to make corrections. Finally, primary contacts completed an *end-of-study* interview. At this time (February-March 2021), each team was at a different point in their planning process, but this end-of-study interview provided an opportunity for participants to reflect on their overall experience of using the Planner to date and share final feedback. The primary contact was the team member who was taking the lead in the planning process and had the most direct experience working through all Planner steps and activities. We anticipated that this person could therefore provide the richest updates on the team's work throughout the process.

In keeping with a user-centered design approach, the baseline, monitoring, and end-of-study interview guides (Multimedia Appendix 2) focused on the users' needs and their experiences and contexts as they engaged with the Planner [19,20]. The discussions were all conducted by video call or phone. The interviews were audio recorded, and the baseline and end-of-study interviews were transcribed verbatim. Field notes were written after the discussions to document observations about the setting and participants' (including interactions between participants in focus groups) and interviewer's reflections. All data collection was conducted by 1 cisgender female research coordinator (J Reszel), a master's-prepared registered nurse experienced in qualitative research. At the start of the study, the researcher had no relationship with participants. Extensive notes were taken during each monitoring call and later verified and enhanced using the audio recording to complete the call log (Multimedia Appendix 3). In addition, we asked each team to share completed planning tools to understand how teams used and adapted the Planner material.

Data Analysis

Applying a cross-sectional approach [23], we used conventional content analysis [26] to inductively code all transcripts and monitoring-interview notes as they became available. This allowed us to identify what the teams were currently working on and their perceptions of the Planner at that time. This approach facilitated probing and follow-up in the subsequent interviews with the participant. All interviews were coded in Microsoft Word by 1 research staff member (J Reszel), with 20% coded independently by a second research team member (TN) as a form of analyst triangulation to enhance credibility [27,28]. Any differences in coding were discussed and resolved by the 2 coders. As analysis progressed, coding and findings were discussed at regularly scheduled meetings with the core research team. We also discussed the field notes for additional context on the setting and team dynamics, such as the extent to which different participants contributed to the discussion and which questions they answered versus those they did not answer. As codes emerged from the transcripts and notes, and the coding scheme was developed, we grouped similar codes into broader categories, including contextual information, Planner feedback, the planning process, and how the Planner is used. We also analyzed the data temporally [23] to identify changes in perceptions over time and to create timelines to map if and when teams completed the various Planner activities. The completed Planner tools were reviewed by the research team and documented as completed or not. When reviewing the tools, the researchers assessed whether the teams made any adaptations to the tools and whether they were used as intended. Exemplary completed tools were identified to be included in the Planner, with permission.

Ethics Approval

We received ethics approval for this study from the Ottawa Health Science Research Ethics Board (20190594-01H). Before starting any study procedure, each participant signed a consent form.

Results

Demographic and Contextual Information on Planning Teams

We enrolled 5 planning teams in Canada. The teams were diverse in their geography and composition. There were teams representing the west coast to east coast of Canada in urban and rural settings. The teams ranged from a single person leading all aspects of planning to interdisciplinary teams sharing the planning work (Table 1).

Table 1. Attributes of teams taking part in the field test study.

	Team 1	Team 2	Team 3	Team 4	Team 5
Planning team information					
Geographic area of planning team	Western Canada	Western Canada	Atlantic Canada	Central Canada	Central Canada
Number of people on core planning team identified at the time of study	4	1	6	4	2
Occupations of planning team members	Physiotherapist; fitness coordinator; fitness professionals	Program coordinator	Physiotherapist; program coordinator; fitness professional; person with stroke	Physiotherapist; occupational therapist; rehabilitation manager	Program coordinators
Multiorganization collaboration?	Yes (municipality and private physiotherapy practice)	No (municipality only)	Yes (municipality and health authority)	Initially: no (primary care center only); during study: yes (municipality and primary care center)	No (community-based nonprofit only)
Types of partners participating[a] in planning process	Brain injury group; municipality; physiotherapy clients; stroke club	Local university; health authority; local stroke association	Health authority; inpatient rehabilitation services; outpatient rehabilitation services; municipality	Allied health partners in clinic; clinic clients; municipality	Internal staff; past program participants
Program information					
Planned geographic area for program implementation	City	City	City	City	National (web-based)
Population density of community where program would be offered	Rural or mostly rural	Urban or mostly urban	Urban or mostly urban	Rural or mostly rural	Combination of urban and rural
Size of community where program would be offered	10,000 to 24,999	>50,000	25,000 to 50,000	5000 to 9999	National
Type of organization planning to offer program	Municipality	Municipality	Municipality	Family health team	Web-based

[a]Ranging from consultation to collaboration, as per the International Association for Public Participation (IAP2) Spectrum of Public Participation [29].

Team Composition

To provide context for interpreting the findings, we provide a brief description of each planning team before describing the study participants.

Team 1 included a physiotherapist in private practice and a program coordinator from the municipality who had previously collaborated to plan and implement other adapted fitness programs in the community. On the basis of their previously successful partnership and the perceived need for stroke-specific community programs, they decided to plan a new program together.

Team 2 comprised a single program coordinator from the municipality who was working within its usual organizational model whereby an individual coordinator is largely responsible for all planning activities. The municipality had an existing suite of adapted fitness programs and wished to explore adding a stroke-specific exercise class to its model.

Team 3 was led by a physiotherapist who had previously piloted a stroke-specific exercise program in a long-term care setting. The municipality had expressed interest in collaborating with the health authority on an adapted exercise program. The physiotherapist subsequently formed a new partnership with the municipality to begin planning a stroke-specific exercise program in the community.

Team 4 comprised rehabilitation health professionals and a manager from a primary care clinic affiliated with the local hospital. The team members had experience planning and implementing other group programs tailored to various health conditions within their primary care setting and were considering offering a stroke-specific program in their clinic.

Team 5 included 2 program coordinators from a nonprofit organization planning a web-based adapted exercise program. Before the COVID-19 pandemic, the program was offered in person. The onset of the pandemic provided an opportunity to explore whether the program could transition to a new web-based format reaching a broader geographical area.

Field Test Study Participants and Data Collected

From these 5 teams, we enrolled 12 participants. Between February 2020 and March 2021, we conducted 25 interviews and focus groups with the 12 participants: the 7 baseline sessions lasted an average of 57 (range 40-75) minutes, the 13 monitoring calls lasted an average of 27 (range 18-35) minutes, and the 5 end-of-study sessions lasted an average of 44 (range 31-53) minutes. On average, the planning teams were followed by the research team for 9.6 (range 4-14) months (Table 2).

Table 2. Data collection approaches and description of study participants (N=12).

	Team 1 (n=4)	Team 2 (n=1)	Team 3 (n=1)	Team 4 (n=4)	Team 5 (n=2)
Data collection					
Dates of participation in study	December 2019 to February 2021 (14 months)[a]	April 2020 to March 2021 (11 months)	May 2020 to February 2021 (9 months)	June 2020 to March 2021 (9 months)	October 2020 to February 2021 (4 months)
Types of qualitative data collected	1 baseline interview and 1 baseline FG[b] (with 3 participants); 1 monitoring call; 1 end-of-study interview	1 baseline interview; 3 monitoring calls; 1 end-of-study interview	1 baseline interview; 3 monitoring calls; 1 end-of-study interview	1 baseline interview and 1 baseline FG (with 3 participants); 4 monitoring calls (with 2 participants); 1 end-of-study FG (with 2 participants)	1 baseline FG (with 2 participants); 2 monitoring calls (with 2 participants); 1 end-of-study FG (with 2 participants)
Study participants' role on the planning team, n (%)					
Program manager or coordinator	1 (25)	1 (100)	0	1 (25)	2 (100)
Rehabilitation health professional	1 (25)	0	1 (100)	3 (75)	0
Fitness professional	2 (50)	0	0	0	0
Study participants' previous experience in *planning* adapted or specialized fitness programs, n (%)[c]					
Yes	2 (67)	1 (100)	1 (100)	3 (75)	1 (50)
No	1 (33)	0	0	1 (25)	1 (50)
Study participants' previous experience in *delivering* adapted or specialized fitness programs, n (%)[c]					
Yes	2 (67)	1 (100)	1 (100)	4 (100)	1 (50)
No	1 (33)	0	0	0	1 (50)

[a]We lost contact with team 1 between month 2 and month 11, both inclusive, and no data were collected during this time.

[b]FG: focus group.

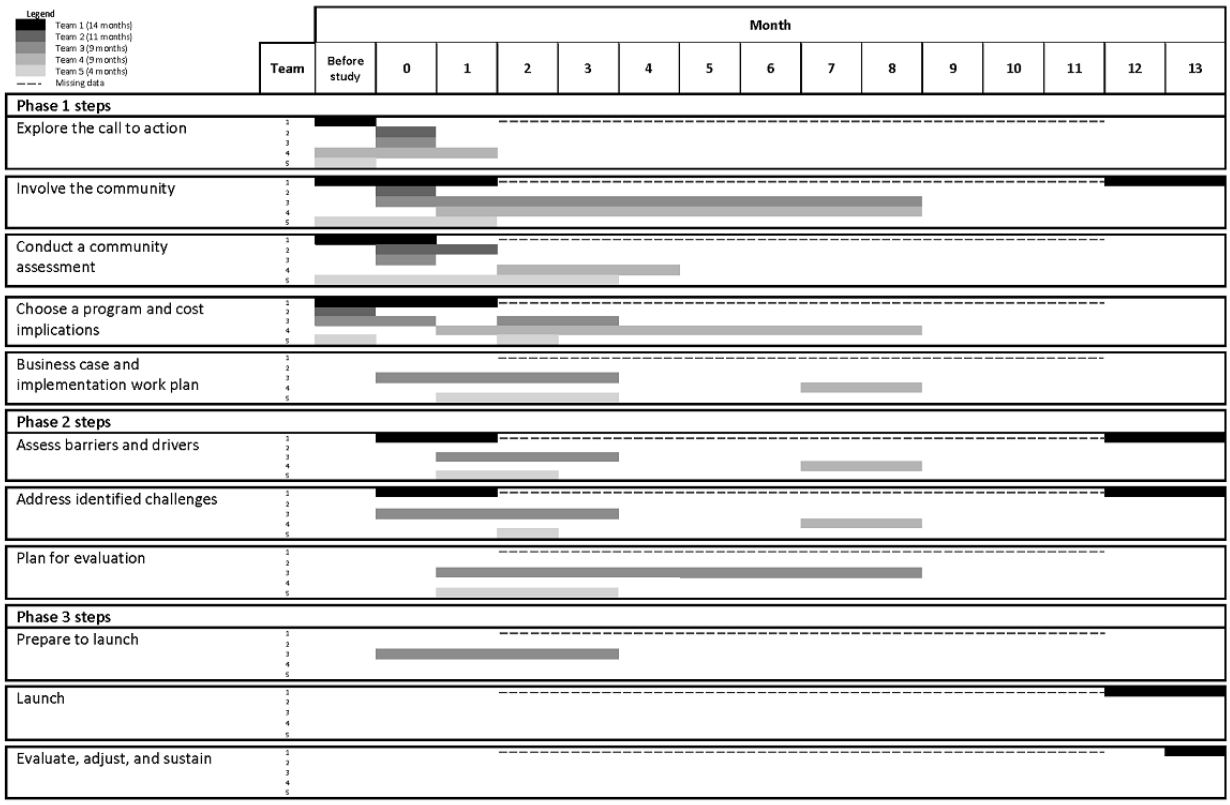
[c]A participant from team 1 did not complete the questionnaire that included these demographic questions.

How Teams Used the Planner in Real-world Conditions

Overview

The 5 teams varied in how they used the Planner, ranging from methodically working through each step to using the Planner more as a reference guide when needed. Figure 1 provides a visual summary of the phases and steps the teams completed during the study period. Most of their planning progress focused on phase 1. Teams took from 1 to ≥8 months to focus on the early planning activities. Because of the impact of the pandemic (eg, facility closures and suspended programs), the teams completed fewer activities beyond phase 1. Despite these challenges, of the 5 teams, 4 (80%) were able to complete at least one step in phase 2 or phase 3. Instead of moving through the steps sequentially, the teams tended to use a nonlinear approach to address many tasks concurrently and revisited some steps multiple times. The results are organized by implementation-planning phase and step.

Figure 1. Summary of Planner phases and steps completed by field test teams during the study period.



Explore the Call to Action

The call to action was one of the first steps completed by each team. Although most (4/5, 80%) of the teams had previously been considering the value of planning a stroke-specific exercise program in their community, these teams described how participating in the field test prompted them to formally start their planning process.

Involve the Community

We observed that involving the community was the step that required the greatest amount of time (Figure 1), reflecting the need to engage community partners throughout the entire planning process. All teams had an intensive focus on involving the community at the front end as they identified key partners to join the planning team and relevant stakeholders to consult. Participants from all teams reported that using the Planner prompted them to consider alternative ways of identifying and engaging people or groups. For example, team 2 reconsidered their usual way of working independently within their own organization and began reaching out to stroke-related organizations in the community. Team 3 decided to add representatives from key referral sources as well as a person living with stroke and their caregiver to the planning team and explored the need for developing a formal partnership agreement. Involving these different partners informed referral pathways and how to accommodate participant needs. This team used the Planner itself as a tool to structure team meetings and engage team members:

> *I think it (the Planner) was quite good actually because it just allowed me to know where to focus and make some plans around how those meetings were going to look. It was really very helpful for meeting planning to focus from a team perspective. I think if it was just myself doing it I might have been able to manage without the Planner, but when you're trying to actually organize your thoughts and articulate them to other people I think the Planner was really helpful in doing that.* [Physiotherapist, team 3, ID6, end of study]

Conduct a Community Assessment

The teams described several methods to conduct their community assessments, which were completed at various time points in the planning process. For example, based on an early assessment of services in the community and the number of people living with stroke, team 1 decided to move forward with planning and implementing a program. However, after launching the program and having lower than anticipated enrollment, this team appreciated the importance of taking the time to complete a detailed community needs assessment upfront:

> *Make sure there's a need for it in your community...do that research before you start so that you're not launching a program that you get only three registrants for, and you don't know if you can run it after all that work.* [Physiotherapist, team 1, ID1, end of study]

At the start of the planning process, team 4 felt confident that they would offer a future program in their clinic setting. However, as they took inventory of the programs currently offered, they realized that their initial idea was similar to an

existing program at their organization. As a result, this team revisited their initial program idea and formed a new partnership with the municipal recreation center to address the service gap they identified. This marked the first time that their organization established a formal partnership with an external organization to plan (and potentially implement) a program in their community.

Team 5 had started their community assessment before enrolling in the study and described administering a survey to potential participants and staff. This allowed them to collect information on community needs, interests, and preferences. They also saw these community surveys as a form of *preadvertising* for their future program and used the survey results in drafting their business case.

Choose a Program and Cost Implications

Of the 5 teams, 4 (80%) had selected an exercise program before joining the study. These decisions were made based on team members' knowledge of, and experiences with, existing programs. Team 4 had not selected a program and undertook a comprehensive process to explore the various options, including reading the content in the Planner, exploring program websites and resources, and connecting with exercise program contacts and other teams who had previously used the programs.

Although many (3/5, 60%) of the teams reported that their organization had a set participant enrollment fee, they indicated that the budget planning tool in the Planner helped them to assess what items and resources they had available to them and what unique start-up and ongoing costs they needed to budget for a program designed specifically for people with stroke.

Business Case and Implementation Work Plan

In total, 40% (2/5) of the teams completed a business case. Team 3 used the business case template in the Planner and adapted it to suit the needs of their team. Team 5 was required to use their organization's template to prepare a business case; however, the team was able to use information from the Planner document and steps completed to produce a comprehensive business case in alignment with their usual organizational processes.

The Planner included an implementation work plan template, although only team 4 used it. The reasons given by other teams for not using the implementation work plan included having another preferred method for tracking planning work or not being able to prepare a detailed work plan at this time because of the uncertainty and limitations imposed by the COVID-19–related restrictions.

Assess Barriers and Drivers and Address Identified Challenges

Of the 5 teams, 4 (80%) assessed barriers and drivers to program implementation and started planning for how challenges could be addressed. Although this step is described in phase 2, exploring potential barriers tended to occur early and simultaneously with phase 1 activities. In total, 60% (3/5) of the teams engaged with their community planning partners to explore potential challenges, which often resulted in the formation of practical solutions. For example, team 1 held a stakeholder meeting with representatives from local stroke and brain injury groups where they identified concerns related to the location, program time, transportation, and costs. Team 3 had a meeting with their stroke and caregiver planning partners and learned about potential barriers related to the accessibility of the building. This led to changes in their program screening form and development of an information sheet for future participants.

Team 1 proceeded to launch a program but revisited their barriers assessment when they encountered enrollment issues. To understand the factors discouraging people from participating, they continued working with their community partners to identify potential issues with referral pathways, reassess barriers to participation, and consider how these could be addressed in future program sessions.

Prepare to Launch and Launch the Program

Because of the pandemic, only 40% (2/5) of the teams worked on this step. Team 3 aimed to complete as much planning work as possible, including launch-readiness activities, with the intention of launching the program when pandemic conditions allowed. However, this site was never able to launch because of ongoing COVID-19–related restrictions. After approximately one year, team 1 was able to launch a session of their exercise program.

Develop Evaluation Plan, Monitor Delivery and Use, and Assess Program and Participant Outcomes

Although unable to launch, 40% (2/5) of the teams did spend time exploring how they could evaluate their program. Team 3 prepared a participant satisfaction survey, selected client-centered before-and-after measures, and collaborated with their internal evaluation team to identify available data that could be pulled to show broader system impact. Team 5 reviewed the program-fidelity templates included in the Planner and subsequently selected and adapted one of the tools for use in their setting.

Although at baseline team 1 indicated that the Planner made them think about the importance of developing a comprehensive monitoring plan that includes before-and-after participant outcome measures, at the end of the study the participants explained that they had overlooked this, and it was never implemented. However, they planned to administer a participant satisfaction survey.

Evaluate, Adjust, and Sustain

Because of pandemic-related restrictions and delays, none of the enrolled field test teams evaluated, made adjustments, or focused on sustainment during the study period.

Effects of Using the Planner

Participants' Changing Perceptions of the Planner Itself

As some participants used the Planner, their impressions of the proposed process changed. For example, upon first reviewing the Planner, at least one member from every team expressed concerns about the length and potential complexity of the planning process and tools. However, after gradually working through the planning process over several months during field testing, the participants' feedback became more positive as they

saw the benefits. The following quotes illustrate how a program coordinator changed their views over time as they used the Planner:

> *There is too much information. It's overwhelming to read, review, use, and implement.* [Program coordinator, team 5, ID12, baseline]

> *Once you get over the size of it, I can't stress enough what a great resource it is...This [Planner] really gave me a great overview of the right way to plan something...When you see something laid out from start to finish it makes a real big difference. The Planner is going to be so helpful for all programming that I am involved in moving forward. I have learned so much.* [Program coordinator, team 5, ID12, end of study]

Building Capacity for Implementation Planning

All teams gave examples of how the Planner led them to undertake steps and activities that contributed positively to their implementation planning, which they would not have undertaken otherwise. All teams identified ways in which the Planner changed their overall approach to program planning. A participant identified themselves as a *ready-set-go* personality and indicated that the Planner helped them to pause and consider other activities to enhance the success of programs in their organization (eg, conducting a thorough community assessment and developing referral pathways from the community). Another participant identified several Planner steps that were not part of their organization's usual practices (eg, partner engagement, decision-making methods, and planning for evaluation at the outset) and acknowledged that although these steps would lengthen the planning process, it would be worthwhile. All teams indicated that they would use the Planner for program planning in the future, both to continue planning their current program (when the pandemic-related restrictions allow) and for other program planning. We have summarized the participants' usual planning process and provided examples of how the Planner changed their implementation-planning knowledge, attitudes, and activities (Table 3).

Table 3. Examples of the effects of the Planner on study participants' implementation-planning knowledge, attitudes, and activities.

Team	Summary of usual approach to planning	Examples of how the Planner influenced planning knowledge, attitudes, and activities	Illustrative quotes from interviews

1	• Participants gave differing views:
	  • Comprehensive, formalized planning process generally in alignment with the SRiM[a] Implementation Planner (fitness coordinator)
	  • No formal planning framework; experience launching adapted exercise programs (private practice physiotherapist)
		• Increased knowledge regarding participant-centered considerations (eg, room location)
		• Decision to host a stakeholder meeting to engage community members in the planning process
			"We had a stakeholder meeting as a result of utilizing the toolkit. Had I been doing this on my own, I probably would have thought I didn't need to do that, but it was really good to have. I looked at the Planner before the meeting to think about who do we invite to this meeting? Who are the key stakeholders? What are the key questions we should be discussing at this planning stage? And when we had the stakeholder meeting, it just brought up some really valid points around who are we targeting? Who are we missing? What are the barriers?" [Physiotherapist, team 1, ID1, monitoring interview 1]

2	• Program coordinators have significant autonomy to propose and launch new programs. • Typically driven by the program of interest and the recreation center, rather than by a formal assessment of needs in the community	• More positive attitudes about the benefits of completing early planning steps (eg, partnerships and community assessment) before launching • Shifting from an individualized to a more inclusive community-centered planning model	"We've been talking about new programs and talking about building relationships with other community partners and the health system, and I'm like, that's that idea within the Planner—doing that full community survey and getting into the actual community." [Program coordinator, team 2, ID5, end of study]
3	• Programs are typically initiated by staff members within the organization, either as an organizational or provincial directive, or by a frontline staff member seeking managerial approval for a specific program.	• Increased ability to use a community-centered approach and successfully engage a diverse team of community stakeholders on the planning team • Increased understanding of program planning by working through the Planner	"Just having that [the Planner] as a reference guide for future planning...I think I have a better understanding of how to go about the planning." [Physiotherapist, team 3, ID6, end of study]
4	• Programs to be offered typically built into the job descriptions of clinic staff and based on needs observed in clinic • Programs typically planned and implemented in the clinic setting by clinic staff	• The Planner process prompted them to shift from a planning team at 1 organization to forming a new partnership with the municipality	"Especially I should say like never working with an outside partner...I'm used to teaching group exercise classes in the hospital, but now we're looking at doing them outside with groups and partners; it's uncharted territory for me...Because we are a hospital, a lot of that stuff that the Planner goes through we didn't have to do because it was already established for us. And now that we've decided we are going to be working with the municipality, we're looking to the Planner even more now for the implementation planning." [Physiotherapist, team 4, ID10, end of study]
5	• Program planning typically driven by an observed community need or through a desire to expand or adapt an existing successful program to other regions	• More positive attitudes toward using a formal framework to structure their process • Increased knowledge about new steps to integrate into their process (eg, planning for evaluation and fidelity assessment upfront)	"The process has been amazing and it has been really refreshing—we were just rushing to [adapt this program], to now having the process to go oh yeah, let's use this Planner to direct our focus...we definitely wouldn't have come to the same place without the Planner." [Program coordinator, team 5, ID11, end of study]

[a]SRiM: Stroke Recovery in Motion.

Conditions That Hindered or Facilitated Uptake of the Planner Process by Community Groups

By following the 5 teams and comparing their engagement with the process defined in the Planner, we identified conditions that made the recommended planning elements easier or more difficult to apply in practice. These factors typically related to organizational context and support, team leadership style, the value placed on community-partner engagement, and the COVID-19 pandemic.

Challenging Conditions

Teams who followed their organization's usual processes for planning programs sometimes prematurely judged recommended activities in the Planner as not applicable. Planner use was challenging for organizations and staff primarily focused on the number of programs developed and launched. The

comprehensive planning process was viewed as potentially too long to meet usual organizational timelines. The *one-person show* model, in which planning activities are undertaken by 1 organization or person without engagement of a diverse team, similarly discouraged consideration and use of the entire planning process. In addition, organizational contextual factors such as restructuring and changing priorities, as well as broader contextual factors such as the COVID-19 pandemic, created barriers and delays to following the planning process. The COVID-19 pandemic presented significant challenges because organizational priorities and planning timelines shifted quickly and often to address changing pandemic conditions, and difficulties were encountered in forming and maintaining external partnerships.

Supportive Conditions

Embracing the Planner process was facilitated by a firm belief in the role of community-based organizations in enhancing the health and well-being of people with stroke. Organizational leaders who valued an evidence-informed planning approach and provided their staff with the dedicated time and resources to work through planning activities created positive conditions supportive of using the Planner. Another positive factor was having a team lead who was open-minded to the Planner process and willing to undertake new steps and activities. Several team leads acknowledged that they personally did not have all the answers or resources and worked early on to identify a diverse group of people to join their core planning team and act as advisers. This openness and engagement led to these teams working through later Planner activities successfully (eg, asking people with stroke and caregivers about their needs and preferences and collaborating with fitness professionals on participant-screening protocols). Finally, for some (2/5, 40%) of the teams, an unanticipated condition that supported Planner use related to the COVID-19 pandemic—with the closure of many services, team members described having *more* time to dedicate to planning.

Lessons Learned on How to Effectively Use the Planner

On the basis of their field test experience, participants offered insightful suggestions about how to apply the Planner (Textbox 1), which may be informative to future users.

Textbox 1. Lessons learned from field test participants on how to most effectively use the Stroke Recovery in Motion Implementation Planner.

Suggestions for how to use the Planner and illustrative quotes

- The Planner is a comprehensive document with a lot of information. Get the *big picture* overview, use strategies to break up the content, and flag priority areas for your team.
  - "I was trying to read it all at once when I wasn't actively doing any planning. Now I am going through it one section at a time and just trying to tackle that section. I'm actively planning as I'm doing it and checking off the tasks, and it definitely feels more manageable." [Physiotherapist, team 4, ID10, monitoring interview 1]
  - "Even though it's long, read the whole thing first...If someone's just starting and they're going to use the Planner, read it first to keep in mind what you've done in the past and then highlight and make notes on the sections that you know are really going to be useful to you. So that way when you do go back you know exactly where to go." [Program coordinator, team 5, ID12, end of study]
- Do not treat the process as completely linear—give yourself permission to jump around the Planner.
  - "Depending on their personalities, some people are very like 'I must follow the steps. I must do the step-by-step-by-step' and I am an example of that person. And so there are some times where it was like well I've gotten to here and I haven't been able to do a proper community survey therefore stop, I cannot go any further because I haven't done this step yet. Well no, you could still flip forward and see what else you can get started on." [Program coordinator, team 2, ID5, end of study]
- Take time and get the early stages of the planning process right.
  - "Even if you think you have this all in the bag and you have a program that's going to start in a couple of months, it's still really important to go through all of the process because you really need to have those evaluation bits in place from the outset...It's the depth of quality. Having the time to put into this process will save you time down the road and avoids situations that you can't really get yourself out of...would that have looked different if we had a different process in place from the beginning?" [Program coordinator, team 5, ID11, end of study]
- Engage with the Planner steps and tools beyond just ticking items off. Take an active approach with the Planner and use it to actually engage with people and organizations.
  - [related to using a Planner tool to assess the facility] "Actually have them go in the building and go through [the checklist] instead of just doing it by memory and what they think it's going to be. Have someone actually go and do a walk-through of the building." [Physiotherapist, team 3, ID6, end of study]
- Following the planning process can be more manageable if the load is shared among the team members. Identify a team lead who has the full view of the Planner and can delegate tasks and activities to other team members.
  - "I would have looked at delegating a little bit more. I did a little bit of 'okay you do this, you do that,' but I think it could have been done more effectively where [other planning team members] could have taken on some of the bigger pieces..." [Physiotherapist, team 3, ID6, end of study]

Discussion

Principal Findings

We followed 5 diverse teams as they used a newly created implementation-planning guide, the Stroke Recovery in Motion Implementation Planner, to plan for the implementation of community-based exercise programs for people with stroke. Because of the COVID-19 pandemic, none of the teams were able to work through the full planning process, with most activities focused on the first of the 3 phases. The findings of this field test study showed that teams took different approaches to applying and adapting the Planner in their settings. All teams indicated that the Planner influenced their approach to program planning and that they intended to continue using the resource in the future. We identified various team, organizational, and broader contextual factors that hindered or facilitated uptake of the Planner by teams. Study participants shared valuable *tips from the field* to help future teams optimize their use of the Planner. Given the paucity of literature reporting implementation toolkit evaluations [30,31], the findings of this study contribute data from our evaluative work on a novel implementation toolkit.

Comparison With Prior Work

A key finding was that the teams did not complete the planning process in a linear manner. All teams conducted multiple steps simultaneously, with some steps being revisited multiple times during the planning process, a pattern reported elsewhere [16]. We also observed that many planning activities (including those across multiple phases) were worked on early in the planning journey. Although we attributed this finding partly to pandemic-related challenges, previous research found a similar pattern, with teams reporting using an implementation guide most frequently in the early stages [32]. Furthermore, despite all having access to the Planner, no 2 teams followed an identical planning path. Although we are unable to speak to the impact of these diverse planning journeys in terms of implementation outcomes, a previous study using an implementation guide found that it facilitated standardization, while allowing flexibility according to the individual context and resources, with all sites successfully implementing the planned program [33]. Among the barriers to implementation, one of the most common relates to challenges adapting evidence-based interventions [13]. The Planner was therefore designed to be a practical approach to planning based on evidence, which means that it is to be used in alignment with the local context [18]. It was not designed as a *recipe* that must be followed in a lock-step manner. The finding that diverse teams could, in fact, adapt and navigate the planning process in various ways suggests that the resource was applied as intended in real-world settings.

Even experienced planners described new things they learned from the Planner and how this improved their processes, suggesting that the Planner positively influenced end users' planning capabilities, similar to other work reporting improvements in practitioner implementation skills [34]. In our study, 60% (3/5) of the planning teams were led by, or included, rehabilitation professionals. A recent study [12] of 384 allied health professionals (nearly half of whom were rehabilitation professionals) indicated that these practitioners reported lower levels of confidence in planning, implementation, and evaluation. Nearly all expressed an interest in learning about knowledge translation, with web-based training and resources (such as the Planner) being the preferred format [12].

Fitness professionals are another group essential to planning and implementing fitness programming; yet, it is largely unknown how fitness instructors engage in knowledge translation and implementation and what barriers and facilitators they encounter [35]. Although it is well acknowledged that fitness professionals have a critical role in delivering physical activity interventions and programs, there is limited information on this diverse group's capacity and training needs [36]. In alignment with the Planner guiding principle of *inclusiveness* [7], we assert that fitness professionals are essential in cocreating the implementation plan. This field test included 12 participants, but only 2 (17%) were fitness professionals, which we attribute to the pandemic-related closure of fitness facilities and layoffs of fitness professionals. The inclusion of only 2 fitness professionals in our study was insufficient to reflect the overall heterogeneity of this group related to education, qualifications, and practice settings. Understanding this diversity and its effect on Planner use by this group may be important because there is evidence that fitness trainers with higher levels of education are more likely to access scientific journals than those with lower education, who prefer mass media and the internet [35]. Future work is needed to understand the implementation capacity and training needs of fitness professionals to enhance full participation in implementation planning.

In this study we used a passive form of implementation; that is, the teams had access to the Planner guide and tools, but no active implementation facilitation or support or any financial incentives were provided. However, there is evidence to suggest that active facilitation can enhance implementation efforts [37,38]. Given the potential value of implementation facilitation, our research team formed a new partnership with March of Dimes Canada, with the goal of having the Planner endorsed and supported by a well-connected and reputable organization in community-based stroke care. March of Dimes Canada has a national After Stroke program [39] that includes regional coordinators, providing the infrastructure to develop its organizational role in implementation facilitation by supporting sites using the Planner. There is an opportunity for future exploratory work that can contribute to the implementation literature on the role of a national organization and its regional coordinators in disseminating and supporting the use of the Planner. Furthermore, although the purpose of this study was not to assess the impact of the Planner on program outcomes, previous literature has reported that implementation toolkits may contribute to improved clinical outcomes [30]; future work is needed to evaluate the influence of the Planner on program outcomes and sustainability.

We identified several cross-cutting barriers and facilitators to following the implementation-planning process proposed in the guide, including organizational context and support, team leadership style, and the value placed on community-partner engagement. These factors were also identified in a review as core capacity–building domains (leadership, organizational

climate and culture, partnerships, workforce development, and financial processes) that can be modified to build capacity for the implementation of evidence-based practices [40]. Knowing that simply creating an implementation guide does not guarantee use, we have carefully considered the barriers and facilitators identified in the evaluation of the Planner and have used this information to make changes to the Planner itself. For instance, related to community-partnership engagement, we added new content to the Planner on partnership agreements and a sample invitation letter for stroke and caregiver partners [7]. Furthermore, to address the potential barrier of perceiving that some Planner activities were not applicable, we added "Why is this important" statements throughout to clearly show the potential benefits of completing the various steps, activities, and tools [7].

Finally, it is noteworthy that 80% (4/5) of the teams had already selected the specific exercise program they wanted to implement before enrolling in the study. The finding that most teams had already selected a program speaks to the reality that for many teams, identifying a program is, in fact, the starting point that launches their implementation-planning journey. Observing that not every team starts with a *blank slate* at step 1, we revised the Planner to illustrate various starting points, with directions on how to use the Planner accordingly [7].

Limitations and Strengths

The greatest limitation of this study was conducting a prospective field test during a pandemic. The pandemic created significant recruitment challenges, and we were unable to recruit our originally planned sample size. Furthermore, for the teams we did enroll in the field test study, the closure of community facilities and the public health restrictions led to significant planning delays. The enrolled teams were therefore unable to progress through the entire planning process, with only 20% (1/5) of the teams able to launch a program during the study period. The purpose of this study was not to ascertain the effectiveness of the Planner related to program implementation and outcomes; rather, we sought to understand if, how, and why the Planner was used in practice. Despite the pandemic-related planning challenges, many teams were still able to work through many steps and provided valuable insight on how they used the Planner, allowing us to meet the field test study objectives. However, it is important to acknowledge that because the teams engaged much more heavily in the early phases and activities of the Planner, we do not have the same depth of understanding of how teams would use the Planner in the later stages (ie, program launch, evaluation, and sustainability).

Furthermore, most (4/5, 80%) of the teams indicated that it was the study itself that prompted them to officially launch their planning activities. The study data collection procedures may have caused participants to engage more with the Planner (eg, completing Planner activities in anticipation of an upcoming monitoring interview with the study team). However, to model more typical program-planning scenarios, the research team did not provide any program funding and provided only minimal assistance with connecting study participants with resources (eg, people and websites). The interviewer (J Reszel) did not contribute to making decisions in the program-planning process. Finally, it is important to note that the monitoring calls and end-of-study interview were completed with 1 primary contact at each site, and their perceptions and experience of working through the Planner may not represent the experiences of their team members.

A strength of the study was the diversity of the sample, which included planning teams with differing compositions based in both urban and rural settings across the country. This allowed us to observe the use of the Planner in a variety of settings, which enhances the transferability of the findings. In addition, despite the challenging pandemic conditions, we were able to maintain contact with 80% (4/5) of the teams throughout the study period, allowing us to collect comprehensive data on Planner use and experiences over time.

Conclusions

Catalyzing the expansion of safe, effective, and sustainable community-based exercise programs is important to the long-term health of people with stroke. The Stroke Recovery in Motion Implementation Planner [14] is an adaptable resource that may be used in diverse settings to plan and implement community-based exercise programs for people with stroke. The results of this study contribute to the implementation science literature by describing how end users made use of an implementation guide and the influence of the guide on implementation-planning knowledge, attitudes, and activities. These findings may be informative to others who are developing resources to build the capacity of those working in community-based settings to implement new programs and practices. Future work is needed to understand how teams use the Planner to launch, evaluate, and sustain programs; to monitor ongoing use; and to understand the effect and outcomes of using the Planner. The Planner is now hosted by March of Dimes Canada and can be accessed on the organization's website [14].

Acknowledgments

The authors thank all the study participants who generously gave their time amid the challenges of the COVID-19 pandemic to provide them with detailed and insightful input on the Planner. IDG is a recipient of a Canadian Institutes of Health Research Foundation grant (FDN 143237). NMS was supported by a Heart and Stroke Foundation Mid-Career Investigator award. AT is supported by a Clinician-Scientist Award (Phase II) from the Ontario Heart and Stroke Foundation (P-19-TA-1192). The authors acknowledge Patrick Faucher for the final graphic design of the Planner. This project has been made possible by Brain Canada Foundation through the Canada Brain Research Fun, with the financial support of Health Canada and the Heart and Stroke Foundation Canadian Partnership for Stroke Recovery. The views expressed herein do not necessarily represent the views of the Minister of Health or the Government of Canada.

Authors' Contributions

IDG and MTB coled all aspects of the study. IDG, MTB, JvdH, GA, MLB, KE, JJE, JLM, MLAN, MP, J Richardson, NMS, and AT contributed to the conceptualization and the design of the study. J Reszel collected the data. J Reszel, TN, JvdH, and IDG analyzed the data. J Reszel drafted the initial manuscript. All authors participated in the writing of this manuscript and reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Reporting checklist.
[PDF File (Adobe PDF File), 179 KB - formative_v6i7e37243_app1.pdf]

Multimedia Appendix 2
Semistructured interview and focus group guides.
[PDF File (Adobe PDF File), 201 KB - formative_v6i7e37243_app2.pdf]

Multimedia Appendix 3
Monitoring call documentation form.
[PDF File (Adobe PDF File), 178 KB - formative_v6i7e37243_app3.pdf]

References

1. Stroke in Canada: Highlights from the Canadian Chronic Disease Surveillance System. Public Health Agency of Canada. Ottawa, ON: Public Health Agency of Canada; 2017. URL: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/stroke-canada-fact-sheet.html [accessed 2022-02-08]
2. Together In Movement and Exercise (TIME™). University Health Network. 2022. URL: https://www.uhn.ca/TorontoRehab/Clinics/TIME [accessed 2022-02-08]
3. Fit for Function Community Stroke Wellness Program. Mobility Research. 2021. URL: http://mobilityresearch.ca/fit-for-function-2/ [accessed 2022-02-08]
4. FAME - Fitness and Mobility Exercise Program. 2021. URL: http://fameexercise.com/ [accessed 2022-02-08]
5. Saunders DH, Sanderson M, Hayes S, Johnson L, Kramer S, Carter DD, et al. Physical fitness training for stroke patients. Cochrane Database Syst Rev 2020 Mar 20;3(3):CD003316 [FREE Full text] [doi: 10.1002/14651858.CD003316.pub7] [Medline: 32196635]
6. Annual Report 2017-18: Making an IMPACT. Canadian Partnership for Stroke Recovery. Ottawa, ON: Canadian Partnership for Stroke Recovery; 2018. URL: https://canadianstroke.ca/wp-content/themes/canadian-parternship-for-stroke-recovery/annualreports/2017-2018annualreport.html [accessed 2022-02-08]
7. Reszel J, van den Hoek J, Nguyen T, Aravind G, Bayley MT, Bird ML, et al. The Stroke Recovery in Motion Implementation Planner: Mixed methods user evaluation. JMIR Form Res 2022 Jul 28;6(7):e37189. [doi: 10.2196/37189]
8. Davis R, D'Lima D. Building capacity in dissemination and implementation science: a systematic review of the academic literature on teaching and training initiatives. Implement Sci 2020 Oct 30;15(1):97 [FREE Full text] [doi: 10.1186/s13012-020-01051-6] [Medline: 33126909]
9. Kislov R, Waterman H, Harvey G, Boaden R. Rethinking capacity building for knowledge mobilisation: developing multilevel capabilities in healthcare organisations. Implement Sci 2014 Nov 15;9:166 [FREE Full text] [doi: 10.1186/s13012-014-0166-0] [Medline: 25398428]
10. Proctor EK, Chambers DA. Training in dissemination and implementation research: a field-wide perspective. Transl Behav Med 2017 Sep;7(3):624-635 [FREE Full text] [doi: 10.1007/s13142-016-0406-8] [Medline: 27142266]
11. Lynch EA, Mudge A, Knowles S, Kitson AL, Hunter SC, Harvey G. "There is nothing so practical as a good theory": a pragmatic guide for selecting theoretical approaches for implementation projects. BMC Health Serv Res 2018 Nov 14;18(1):857 [FREE Full text] [doi: 10.1186/s12913-018-3671-z] [Medline: 30428882]
12. Barrimore SE, Cameron AE, Young AM, Hickman IJ, Campbell KL. Translating research into practice: how confident are allied health clinicians? J Allied Health 2020;49(4):258-262. [Medline: 33259570]
13. Bach-Mortensen AM, Lange BC, Montgomery P. Barriers and facilitators to implementing evidence-based interventions among third sector organisations: a systematic review. Implement Sci 2018 Jul 30;13(1):103 [FREE Full text] [doi: 10.1186/s13012-018-0789-7] [Medline: 30060744]
14. Graham ID, van den Hoek J, Reszel J, Nguyen T, Aravind G, Bayley MT, et al. Stroke Recovery in Motion: Community-based exercise program implementation planner. Ottawa Hospital Research Institute. Ottawa, ON: Ottawa Hospital Research Institute; 2022. URL: http://www.afterstroke.ca/srimp [accessed 2022-06-09]

15. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, et al. Lost in knowledge translation: time for a map? J Contin Educ Health Prof 2006;26(1):13-24. [doi: 10.1002/chp.47] [Medline: 16557505]
16. Harrison MB, Graham ID, van den Hoek J, Dogherty EJ, Carley ME, Angus V. Guideline adaptation and implementation planning: a prospective observational study. Implement Sci 2013 May 08;8:49 [FREE Full text] [doi: 10.1186/1748-5908-8-49] [Medline: 23656884]
17. Harrison MB, van den Hoek J, Graham ID. CAN-Implement©: Planning for Best-Practice Implementation. Philadelphia, PA, USA: Lippincott, Williams and Wilkins; 2014.
18. Harrison MB, Graham ID. Knowledge Translation in Nursing and Healthcare: A Roadmap to Evidence-informed Practice. Oxford, UK: Wiley Blackwell Publishing; Mar 2021.
19. Dopp AR, Parisi KE, Munson SA, Lyon AR. A glossary of user-centered design strategies for implementation experts. Transl Behav Med 2019 Nov 25;9(6):1057-1064. [doi: 10.1093/tbm/iby119] [Medline: 30535343]
20. User-Centered Design Basics. Usability.gov. 2022. URL: https://www.usability.gov/what-and-why/user-centered-design.html [accessed 2022-02-08]
21. Georges A, Schuurman D, Vervoort K. Factors affecting the attrition of test users during living lab field trials. Technol Innov Manag Rev 2016;6(1):35-44 [FREE Full text] [doi: 10.22215/timreview/959]
22. Lyon AR, Koerner K. User-centered design for psychosocial intervention development and implementation. Clin Psychol (New York) 2016 Jun;23(2):180-200 [FREE Full text] [doi: 10.1111/cpsp.12154] [Medline: 29456295]
23. Thomson R, Holland J. Hindsight, foresight and insight: the challenges of longitudinal qualitative research. Int J Soc Res Methodol 2003 Jan;6(3):233-244. [doi: 10.1080/1364557032000091833]
24. Calman L, Brunton L, Molassiotis A. Developing longitudinal qualitative designs: lessons learned and recommendations for health services research. BMC Med Res Methodol 2013 Feb 06;13:14 [FREE Full text] [doi: 10.1186/1471-2288-13-14] [Medline: 23388075]
25. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
26. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
27. Lincoln YS, Guba EG. Naturalistic Inquiry. Newbury Park, CA, USA: Sage Publications; 1985.
28. Patton MQ. Enhancing the quality and credibility of qualitative analysis. Health Serv Res 1999 Dec;34(5 Pt 2):1189-1208 [FREE Full text] [Medline: 10591279]
29. IAP2 Spectrum of Public Participation. International Association for Public Participation. 2018. URL: https://cdn.ymaws.com/www.iap2.org/resource/resmgr/pillars/Spectrum_8.5x11_Print.pdf [accessed 2022-02-08]
30. Yamada J, Shorkey A, Barwick M, Widger K, Stevens BJ. The effectiveness of toolkits as knowledge translation strategies for integrating evidence into clinical care: a systematic review. BMJ Open 2015 Apr 13;5(4):e006808 [FREE Full text] [doi: 10.1136/bmjopen-2014-006808] [Medline: 25869686]
31. Barac R, Stein S, Bruce B, Barwick M. Scoping review of toolkits as a knowledge translation strategy in health. BMC Med Inform Decis Mak 2014 Dec 24;14:121 [FREE Full text] [doi: 10.1186/s12911-014-0121-7] [Medline: 25539950]
32. O'Toole JK, Starmer AJ, Calaman S, Campos ML, Goldstein J, Hepps J, et al. I-PASS mentored implementation handoff curriculum: implementation guide and resources. MedEdPORTAL 2018 Aug 03;14:10736 [FREE Full text] [doi: 10.15766/mep_2374-8265.10736] [Medline: 30800936]
33. Carter M, Karwalajtys T, Chambers L, Kaczorowski J, Dolovich L, Gierman T, CHAP Working Group. Implementing a standardized community-based cardiovascular risk assessment program in 20 Ontario communities. Health Promot Int 2009 Dec;24(4):325-333. [doi: 10.1093/heapro/dap030] [Medline: 19819896]
34. Eisenman DP, Adams RM, Lang CM, Prelip M, Dorian A, Acosta J, et al. A program for local health departments to adapt and implement evidence-based emergency preparedness programs. Am J Public Health 2018 Nov;108(S5):S396-S398. [doi: 10.2105/AJPH.2018.304535] [Medline: 30260692]
35. Stacey D, Hopkins M, Adamo KB, Shorr R, Prud'homme D. Knowledge translation to fitness trainers: a systematic review. Implement Sci 2010 Apr 15;5:28 [FREE Full text] [doi: 10.1186/1748-5908-5-28] [Medline: 20398317]
36. De Lyon AT, Neville RD, Armour KM. The role of fitness professionals in public health: a review of the literature. Quest 2017;69(3):313-330. [doi: 10.1080/00336297.2016.1224193]
37. Harvey G, Llewellyn S, Maniatopoulos G, Boyd A, Procter R. Facilitating the implementation of clinical technology in healthcare: what role does a national agency play? BMC Health Serv Res 2018 May 10;18(1):347 [FREE Full text] [doi: 10.1186/s12913-018-3176-9] [Medline: 29743068]
38. Sridhar A, Drahota A, Walsworth K. Facilitators and barriers to the utilization of the ACT SMART Implementation Toolkit in community-based organizations: a qualitative study. Implement Sci Commun 2021 May 26;2(1):55 [FREE Full text] [doi: 10.1186/s43058-021-00158-1] [Medline: 34039434]
39. After Stroke. March of Dimes Canada. 2020. URL: https://www.afterstroke.ca/ [accessed 2022-06-09]

40. Brownson RC, Fielding JE, Green LW. Building capacity for evidence-based public health: reconciling the pulls of practice and the push of research. Annu Rev Public Health 2018 Apr 01;39:27-53 [FREE Full text] [doi: 10.1146/annurev-publhealth-040617-014746] [Medline: 29166243]

Edited by A Mavragani; submitted 12.02.22; peer-reviewed by S Castiglione, E Linkewich; comments to author 07.04.22; revised version received 10.06.22; accepted 21.06.22; published 29.07.22.

Please cite as:
Reszel J, van den Hoek J, Nguyen T, Aravind G, Bayley MT, Bird ML, Edwards K, Eng JJ, Moore JL, Nelson MLA, Ploughman M, Richardson J, Salbach NM, Tang A, Graham ID
How Community-Based Teams Use the Stroke Recovery in Motion Implementation Planner: Longitudinal Qualitative Field Test Study
JMIR Form Res 2022;6(7):e37243
URL: https://formative.jmir.org/2022/7/e37243
doi:10.2196/37243
PMID:35904855

Original Paper

Participant Experiences of a COVID-19 Virtual Clinical Study Using the Current Health Remote Monitoring Platform: Case Study and Qualitative Analysis

Juliana Pugmire[1], MPH, DrPH; Jessie Lever Taylor[2], BA, MSc; Matt Wilkes[1], MBChB, PhD; Adam Wolfberg[2], MPH, MD; Nicole Zahradka[2], PhD

[1]Current Health Ltd, Edinburgh, United Kingdom

[2]Current Health Inc, Boston, MA, United States

Corresponding Author:
Juliana Pugmire, MPH, DrPH
Current Health Ltd
Suite 2, Ground Floor, Orchard Brae House
30 Queensferry Road
Edinburgh, EH4 2HS
United Kingdom
Phone: 44 131 285 8101
Email: juliana.pugmire@currenthealth.com

Abstract

Background: During the COVID-19 pandemic, individuals with a positive viral test were enrolled in a study, within 48 hours, to remotely monitor their vital signs to characterize disease progression and recovery. A virtual trial design was adopted to reduce risks to participants and the research community in a study titled Risk Stratification and Early Alerting Regarding COVID-19 Hospitalization (RiskSEARCH). The Food and Drug Administration–cleared Current Health platform with a wearable device is a continuous remote patient monitoring technology that supports hospital-at-home care and is used as a data collection tool. Enrolled participants wore the Current Health wearable device continuously for up to 30 days and took a daily symptom survey via a tablet that was provided. A qualitative substudy was conducted in parallel to better understand virtual trial implementation, including barriers and facilitators for participants.

Objective: This study aimed to understand the barriers and facilitators of the user experience of interacting with a virtual care platform and research team, while participating in a fully virtual study using qualitative and quantitative data.

Methods: Semistructured interviews were conducted to understand participants' experience of participating in a virtual study during a global pandemic. The schedule included their experience of enrollment and their interactions with equipment and study staff. A total of 3 RiskSEARCH participants were interviewed over telephone, and transcriptions were inductively coded and analyzed using thematic analysis. Themes were mapped onto the Theoretical Domains Framework (TDF) to identify and describe the factors that influenced study adherence. Quantitative metrics, including adherence to wearable and scheduled tasks collected as part of the RiskSEARCH main study, were paired with the interviews to present an overall picture of participation.

Results: All participants exceeded our definition of a fully adherent participant and reported that participation was feasible and had a low burden. The symptoms progressively resolved during the trial. Inductive thematic analysis identified 13 main themes from the interview data, which were deductively mapped onto 11 of the 14 TDF domains, highlighting barriers and facilitators for each.

Conclusions: Participants in the RiskSEARCH substudy showed high levels of adherence and engagement throughout participation. Although participants experienced some challenges in setting up and maintaining the Current Health kit (eg, charging devices), they reported feeling that the requirements of participation were both reasonable and realistic. We demonstrated that the TDF can be used for inductive thematic analysis. We anticipate expanding this work in future virtual studies and trials to identify barriers and enabling factors for implementation.

(JMIR Form Res 2022;6(7):e37567) doi:10.2196/37567

KEYWORDS

virtual trial designs; virtual enrollment; digitalized health; theoretical domains framework; thematic analysis; remote patient monitoring

Introduction

Background

With the onset of the COVID-19 pandemic in 2020, we have seen rapid shifts in the way people work, engage with education and health care, and conduct their activities of daily life [1]. Many traditional clinical trials have slowed to a halt because of health care shortages and fear of increasing viral transmission [2]. Studies involving human participants have adapted to better use digitalized, decentralized, or virtual trial designs by the end of 2020 (though perhaps not as drastically as expected) [3]. Similar to remote working, virtual trial designs were a possibility that existed before the pandemic but have become a necessity for many researchers wanting to reduce the risk of transmission in human participants and the research community alike, while still conducting research [4,5].

Virtual clinical trials (VCTs) are site-less and rely on technologies such as apps, web-based platforms, wearable devices, and remote monitoring [6]. Digitized clinical trials also use technology to recruit and retain participants and for data collection and analysis [7]. Digitized clinical trials or VCTs leverage digital health technologies to improve participant access and engagement [7-9]. These trial designs have the potential to lower the cost of these studies and expand participation by making trials more accessible to participants [9,10].

With the shift to virtual, digitalized clinical trial designs, it may be helpful for study participants to understand specific implementation issues, including barriers and facilitators. Recruitment and retention in clinical trials are persistent challenges, whether traditional or virtual [7,11,12]. In VCTs, the study participant will likely have to interact with technology they may not have previous experience with, such as a remote continuous monitor, new apps for e-consenting and tracking, or daily surveys delivered by tablets [13,14]. There will almost certainly be a learning curve, with any instruction or assistance available also delivered remotely. Besides technical barriers, there may also be concerns about participant privacy when it comes to sharing sensitive health information [9].

Current Health and Risk Stratification and Early Alerting Regarding COVID-19 Hospitalization

Current Health (Current Health Ltd, Edinburgh, United Kingdom) is a medical technology company that creates a platform that enables continuous remote patient monitoring to support hospital at home programs and care [15]. The Food and Drug Administration–cleared Current Health kit includes a wearable device, which is a small, round disk that is attached to a band and worn on the upper arm. It monitors respiration rate, heart rate, oxygen saturation, skin temperature, and activity [15]. It can be integrated with peripheral devices, including those measuring blood pressure, axillary temperature, spirometry, weight, and continuous glucose. It also incorporates a tablet that can deliver surveys, reminders to take measurements (eg, blood pressure, weight), or a video connection to a health care provider or investigator. It requires approximately 5 minutes for a participant to set up the Current Health kit, including measuring and selecting the correct arm band size, and begin transmitting vital sign data via the secure wireless home hub. The home hub allows the Current Health platform to operate without an in-home Wi-Fi connection, thereby making the technology more inclusive.

The Current Health platform was used in the study, Risk Stratification and Early Alerting Regarding COVID-19 Hospitalization (RiskSEARCH; NCT04709068) [16], funded by the US Department of Health and Human Services branch, the Biomedical Advanced Research and Development Authority. Its purpose was to remotely monitor individuals who tested positive for COVID-19 infection, within the previous 48 hours, to learn more about disease progression and recovery. The enrolled participants wore the Current Health wearable device continuously for up to 30 days. Health data were collected to develop predictive models for the risks of hospitalization and death.

As part of the main study, the research team designed a qualitative substudy run in parallel to gain an in-depth understanding of the participant's experience of taking part in a virtual study. Participants first had to show that they were eligible for the study by answering a web-based eligibility questionnaire, chose a time to connect with a study coordinator to be consented and enrolled, and finally had to set up and use the Current Health kit, which was shipped to their home address, all without meeting the study personnel in person. Once enrolled, participants were asked to answer a daily symptom survey delivered via a tablet and wear the Current Health wearable device 24 hours a day, except when charging the device or showering, bathing, or swimming. For the substudy, participants also agreed to conduct an interview of up to 40 minutes about the experience of participating in the RiskSEARCH study and using the Current Health kit.

The RiskSEARCH study did not progress beyond the pilot phase because of the changing landscape of the COVID-19 pandemic, including vaccine development and receding waves of infection, which negatively affected recruitment [17]. However, the substudy collected in-depth data on 3 participants, presented here as a case series, and qualitative analysis applying the Theoretical Domains Framework (TDF) to better understand the participant experience.

Theoretical Domains Framework

Virtual studies such as RiskSEARCH have many components that demand behavioral adaptation to adhere to the study intervention (eg, engaging in specific ways with the Current Health platform). The TDF synthesizes 128 theoretical constructs from 33 theories into a combined theoretical framework comprising 14 domains [18]. The TDF has been used to evaluate implementation problems, understand the mechanisms of change, and design interventions. The TDF helps

researchers identify and describe the factors that influence a set of behaviors (eg, study adherence). More specifically, it can help investigate implementation issues, including barriers and facilitators, to participating in studies such as the RiskSEARCH study and adopting the behavior changes necessary for adherence.

This is an exploratory piece of research based on a virtually delivered study run during the global COVID-19 pandemic from March 2021 to May 2021. The study team conducted this research to explore the participant experience for improving (1) recruitment and retention in future studies, (2) user experience with the Current Health platform, and (3) the ease with which the platform can be harnessed in other clinical studies, and in particular, virtual studies. We hope that these findings will aid other investigators to successfully conduct virtual studies and VCTs.

Methods

The RiskSEARCH Main Study

The RiskSEARCH study was a virtual, time-sensitive trial for individuals, aged >21 years, who tested positive for COVID-19 infection. The primary purpose of this study was to develop a machine learning–based algorithm to predict the likelihood of requiring a hospital stay of at least 24 hours using data collected from a remote patient monitoring wearable device and symptom surveys. This study used the Current Health platform for hospital-grade remote patient monitoring of vital signs and daily symptom surveys. Participants were recruited through advertisements on social media (Facebook, LinkedIn, etc) and word of mouth from March 2021 to May 2021. If an individual met the inclusion and exclusion criteria (Multimedia Appendix 1) and were interested in participating, they had 48 hours to enroll in the study. They were then consented and shipped a Current Health kit. The details of the main study will be published in a separate paper.

Each day, the participants were sent a 21-question survey to complete on the Current Health tablet. The survey asked if participants experienced 8 specific symptoms (chills, fever, nausea, diarrhea, sore throat, dry cough, muscle ache, and loss of smell or taste) and whether they were better, worse, or the same as the previous day. In addition, there was a free-text response in which participants could add any other symptoms they were experiencing. Questions were also included about whether participants were likely to contact a health care provider or attend a hospital based on how they felt that day. This symptom survey was developed and piloted internally before it was shared with the RiskSEARCH study participants. Its purpose was to capture the symptoms and symptom severity associated with COVID-19 infections, to help drive the prediction model of the main study. In parallel, participants were asked to wear the Current Health wearable device for up to 30 days, taking it off only to charge (up to 30 minutes every 24 hours), shower, bathe, or swim.

Qualitative Substudy

We used semistructured interviews and reported the results following the consolidated criteria for reporting qualitative research checklist [19]. Specifically, we wanted to understand what it was like to use the Current Health kit and participate in a fully remote virtual study during a global pandemic. We collected in-depth data on the acceptability of the RiskSEARCH study and Current Health kit. Focused qualitative and quantitative research provided insights into the user experience of interacting with the Current Health kit, the Current Health research team, enrollment process, and participation in a fully virtual study.

Research Objectives

The research objectives of this study are presented in Textbox 1.

Textbox 1. Research objectives.

Research objectives

- To explore recruitment and retention for the Risk Stratification and Early Alerting Regarding COVID-19 Hospitalization (COVID-19) study
- To explore the feasibility, acceptability, and usability of the intervention, that is, the Current Health wearable device and tablet
- To explore barriers and facilitators of study compliance

Topic Guide and Interviewing

On the basis of the literature, our research objectives, and previous experience in developing interviews to understand engagement with digital technology, the study team designed an interview schedule (Multimedia Appendix 2) to explore barriers and facilitators around different aspects of the study and intervention (web-based enrollment, answering the daily survey, charging the wearable device, etc). One-to-one interviews were conducted by JP via telephone at a prearranged, mutually convenient time. JP was a senior clinical research scientist at Current Health at the time of this study, has >10 years of experience conducting interviews for qualitative and mixed methods research, holds a Doctor of Public Health and Master of Public Health in epidemiology, and is a woman.

Recruitment and Procedure

Although we planned to use a purposive sampling strategy, we changed to convenience sampling when the main study recruitment remained low. A total of 7 participants were offered the opportunity to participate in a one-to-one interview with a research team member (JP). Participants were approached by the study coordinator (JLT) through text messaging or telephone conversations after building rapport through the study enrollment process. A total of 4 participants agreed to participate in the study, and 3 interviews were conducted. A participant could not be contacted to set up the interview. No relationship was

established between the interviewer and the interviewee before the commencement of the study. The participants knew that JP was a research scientist at Current Health and was interested in understanding their experience of participating in RiskSEARCH and using the Current Health kit.

Participants who agreed to participate in the interview were sent a PDF version of the informed consent form (ICF). Participants were sent the ICF via DocuSign (DocuSign, Inc) 24 hours before the interview. Participants could sign ahead of the call with the researcher or wait until the call to complete the ICF and ask any questions before signing. The researcher (JP) ensured that the participant questions were answered and that the participants understood the risks of study participation. Participants could opt out of recording the interviews, but none chose this option.

Intervention

Once enrolled in the main study, participants were required to wear the Current Health device at all times, except when charging the device (20-30 minutes every 24 hours) or when showering, bathing, or swimming. They were also required to keep the tablet charged and answer the daily symptom surveys delivered by the wearable device. Textbox 2 shows the components of the intervention.

Textbox 2. Components of the intervention.

Components

- Be home to receive the Current Health kit delivered by FedEX
- Open Current Health box
- Set up home hub which includes plugging hub into the wall
- Select correct armband size using included sizing guide (out of 3 sizes)
- Charge wearable device on included dock until fully charged, indicated by green lights, and charge daily thereafter
- Insert wearable device into armband and wear next to skin under clothing
- Remove wearable device for showering, bathing, or swimming
- Charge tablet daily
- Answer daily symptom surveys delivered on the tablet
- At the completion of the main study (up to 30 days), repackage the Current Health kit back into the box and use the return label provided to arrange return
- For substudy, arrange a mutually convenient time to be interviewed
- Participate in an over-the-phone interview lasting up to 40 minutes about using the Current Health kit and participating in the Risk Stratification and Early Alerting Regarding COVID-19 Hospitalization study

Data Collection

Participant interviews were conducted over telephone and audio recorded using a laptop application (Windows Voice Recorder, Microsoft Corporation) and a handheld digital recorder as a backup. Interviews were anonymized and transcribed using Trint software (Trint Ltd) and checked, corrected, and edited for accuracy by the researcher who conducted the interviews (JP). Familiarization with the data began at this early stage. Participants were also asked to take a modified Telehealth Usability Questionnaire (TUQ) sent to them via an email link. The TUQ is a validated survey tool that quantifies the usability of telehealth implementations and services [20]. No repeat interviews were carried out, no field notes were made, transcripts were not returned to participants for correction, and participants did not provide feedback on the findings.

Metrics

As part of the main study, interview participants also contributed quantitative data, such as daily symptom surveys submitted via tablets. The data collected relevant to the substudy included the following variables as shown in Textbox 3.

The participants' symptoms and vital sign alarms were presented alongside the qualitative results, as their clinical course may have influenced their experiences.

Textbox 3. Data collected.

Variables

1. Wearable adherence: the time the wearable device was worn compared with the study duration.
2. Daily survey adherence: the number of daily surveys completed compared with the number of daily surveys assigned.
3. Fully adherent, determined using 3 criteria: wearables worn for at least 20 hours a day and at least 6 days a week up to 30 days, daily survey responses at least 6 days a week up to 30 days, and a returned Current Health kit at the end of study participation.
4. Vital signs alarms: alarm thresholds were set for vital sign data going out of range, which could only be seen by the study team.

Analysis

A researcher (JP) conducted the interviews, transcribed the audio recordings using Trint transcription software, and coded the data using NVivo Qualitative Data Analysis Software (version 12; QSR International) [21]. We used reflexive thematic analysis [22]. Data were analyzed inductively following the steps of Braun and Clarke [22,23], specifically (1) familiarization of data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report. Initial codes were inductively generated from the interview transcripts, iteratively condensed, and expanded into themes. The themes were then deductively mapped onto the domains of the TDF.

Ethics Approval

Ethics approval was obtained from the Institutional Review Board, Advarra (Columbia, Maryland, ethics approval number Pro00047371). The collected data were stored in compliance with the European Union General Data Protection Regulation, Current Health Research Data Management Policy, US Health Insurance Portability and Accountability Act, and Current Health Research Data Management Policy. Data were anonymized, and all personal identifiers were removed.

Results

Participant Characteristics and Quantitative Results

Overview

Participant details are provided in Table 1 and discussed in further sections.

Table 1. Participant characteristics.

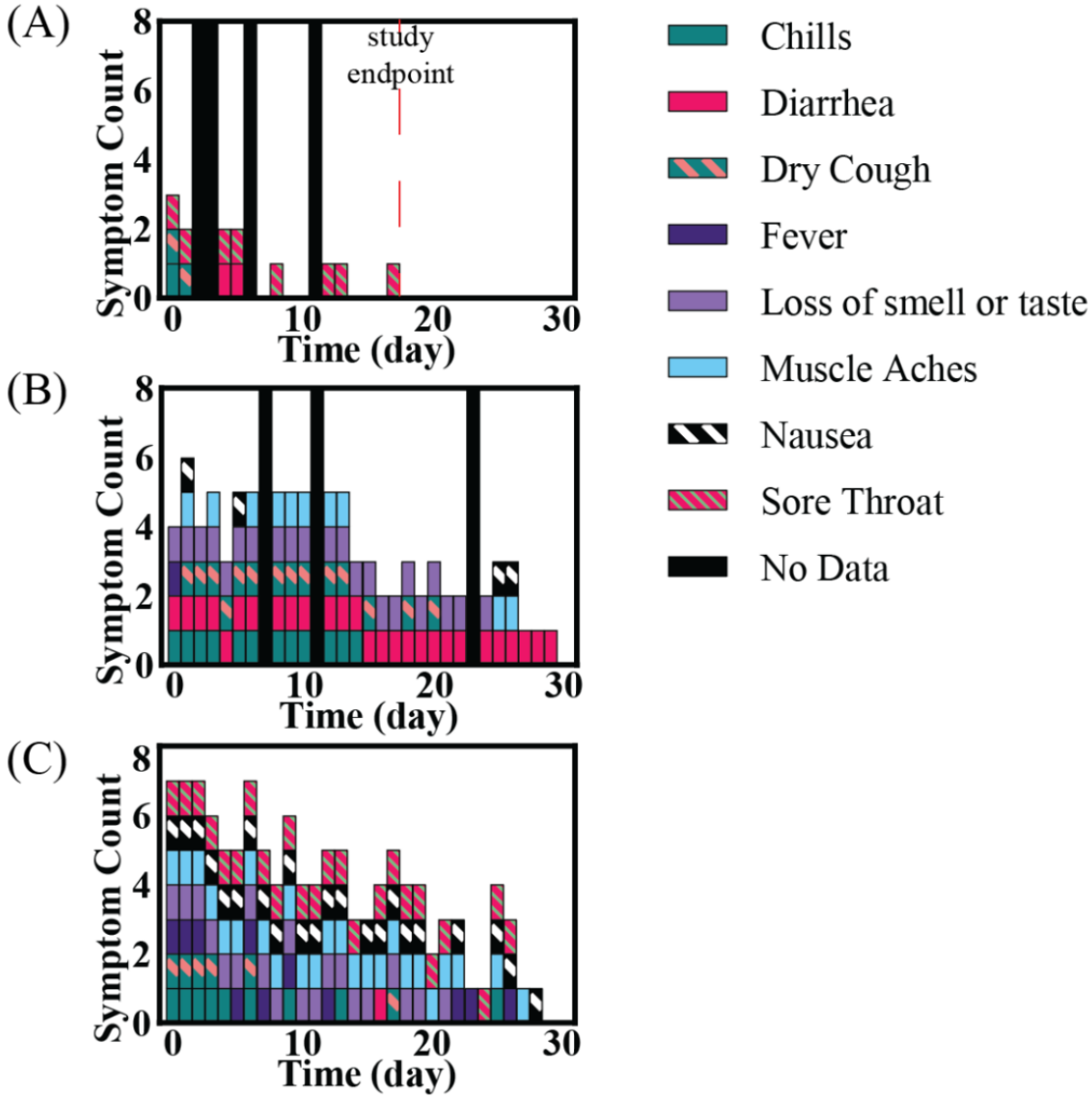
Participant ID	Gender	Age (years), range	Wearable adherence (%)	Daily survey adherence (%)	Telehealth Usability Questionnaire score
RS001	Female	30 to 35	83	76	7
RS006	Female	40 to 45	63	90	—[a]
RS008	Female	35 to 40	92	100	—

[a]Participants did not complete the Telehealth Usability Questionnaire.

Case 1

RS001 initially reported experiencing chills, dry cough, and a sore throat. She did not report experiencing any other symptoms for the duration of her study. By day 6, RS001's chills and dry cough resolved and did not reoccur. However, she reported a sore throat periodically throughout her 17 days in the study. (Figure 1A) Over the course of the study, the only vital sign that triggered an alarm on the Current Health dashboard was a high respiration rate, which occurred on days 0, 3, and 9.

Figure 1. Symptom survey data for participants, daily symptoms reported by participants: (A) RS001, (B) RS006, and (C) RS008. Reported symptoms varied by participants. White gaps between days indicate the participant reported feeling no symptoms. Black bars indicate the days that the participant did not complete the daily symptom survey. Red hatched line indicates the study duration ended before 30 days.



Case 2

RS006 experienced all 8 symptoms, specifically asked about in the daily symptom survey over the course of the study. Her fever resolved on day 1 and did not reoccur. Conversely, her diarrhea did not resolve until day 30. Her symptoms decreased over the 30 days of study participation (Figure 1B). This trend in self-reported symptoms aligned with her vital signs data. RS006 triggered 34 alarms, but no alarms were triggered from day 23 onward. Nearly half of the alarms triggered were for a high pulse rate with a low amount of motion detected (ie, the participant's pulse rate was high while not exerting themselves physically). Other alarms triggered were for low oxygen saturation and a high respiration rate with a low amount of motion detected.

Case 3

RS008 experienced all 8 symptoms included in the daily symptom survey, although diarrhea was reported only once on day 16. Nausea was the most persistent symptom, continuing until day 28. No symptoms were reported on days 29 or 30. As with the other participants, RS008's symptoms improved over the course of the study (Figure 1C), which was also reflected in her vital signs data. RS008 triggered 10 alarms during the 30-day study period. No alarms were triggered after day 19.

Although all participants were asked to follow a link provided via email to complete the TUQ survey, only one participant completed the survey. This participant scored *strongly agree* (7 on the 7-point Likert scale) to all 21 questions of the TUQ, indicating high levels of usefulness, ease of use, effectiveness, reliability, and satisfaction with the Current Health kit.

Qualitative Results

Overview

In all, 3 interviews were conducted toward the end of study participation when participants were feeling better. Interviews ranged from 18 to 35 minutes. Inductive thematic analysis helped identify 13 main themes and subthemes associated with the participant experience of using the Current Health kit and being part of the RiskSEARCH study and included (1) Participant Situations, (2) Getting Started, (3) Study Support, (4) Study Communication, (5) Protecting and Contributing, (6) Determination, (7) Study Pros & Cons, (8) Optimism, (9) Uncertainty, (10) Payment, (11) Accessing Data, (12) Memory & Reminders, and (13) Making Habits.

These themes were deductively mapped to the TDF domains. These domains are described below and presented in Table 2. There were 3 domains of the TDF in which we did not match any data to: *Intentions*, *Goals*, and *Emotion*.

The main domains of the TDF, which we were able to map our themes onto, were *Environmental Context and Resources*; *Knowledge* combined with *Skill*; *Social/Professional Role and Identity*; *Beliefs about Capabilities*; *Optimism*; *Beliefs about Consequences*; *Reinforcement*; *Memory, Attention, and Decision Processes*; *Social Influences*; and *Behavioral Regulation*.

Table 2. Theoretical Domains Framework (TDF) constructs, Risk Stratification and Early Alerting Regarding COVID-19 Hospitalization (RiskSEARCH) themes, and description.

TDF constructs	RiskSEARCH themes	Description
Environmental context and resources	Participant Situations: being sick with COVID-19 infection; caretaking responsibilities	• Study participants were recruited and went through the study after testing positive for COVID-19 infection • Study required steps that may have been more challenging for participants who had many caretaking responsibilities
Environmental context and resources	Getting Started: enrollment; kit components; for example, wearables; unknowns; suggestions	• Participants had to self-navigate through a web-based enrollment system and website • Current Health kit required setup by the participant themselves (though they did have access to technology and study team support) • Communication between tablet or wearable and participant • Wearable needed charging 30 minutes every 24 hours. Participant did not know the battery level of the wearable, but green lights on charger indicated that it was fully charged • Suggestions for improving any aspect of the Current Health kit
Knowledge combined with skill	Study Support[a]: personnel; materials	• Technology support was available 24/7 to help with any aspect of setting up or using the wearable • The study team was available on demand to answer any questions relating to the study or Current Health kit • The Quick Start Guide was available in a hard-copy booklet in the Current Health kit or digitally accessible via the tablet • "Study Support' double coded with Social Influences"
Knowledge combined with skill	Study Communication[a]: passive; active	• Website as a source for information on the study and COVID-19 pandemic • "'Study Communication' double coded with Social Influences"
Social or professional role and identity	Protecting and Contributing	• Help or protect others; feeling a sense of community responsibility; wanting to help in a difficult time; being someone who helps • Motivation to help others • A sense of contributing toward the management of the COVID-19 pandemic
Beliefs about capabilities	Determination	• The level of commitment while engaging with the Current Health kit—survey or tablet or wearable
Optimism	Study Pros and Cons	• Positive and negative aspects of the study
Optimism	Optimism	• Seeing the positive in the bad situation of being tested positive for COVID-19 infection
Beliefs about consequences	Uncertainty	• Feedback regarding user's "performance" or whether kit was working properly
Reinforcement	Payment	• Study participants expressing their views on the US $100 offered for their time and effort
Reinforcement	Accessing Data	• Having access to own data • The wearable device does not transmit data to the participant
Memory, attention, and decision processes	Memory and Reminders	• Remembering to charge and wear the wearable • Reminders to take the survey every day • Reminders to charge the tablet and take the survey
Social influences	Study Support[a]: personnel	• Possibility to contact technology support or study team • "Double coded with Knowledge combined with Skill"
Social influences	Study Communication[a]: active	• Via email, text, or telephone call • "'Study Communication' double coded with Knowledge combined with Skill"

TDF constructs	RiskSEARCH themes	Description
Behavioral regulation	Making Habits	• Habit formation around charging the wearable and tablet and taking the survey • Ability to support routine or habit formation

[a]Mapped to 2 different domains from the TDF.

Environmental Context and Resources

Many contextual factors impacted the participants' ability to successfully participate in the RiskSEARCH virtual study, including the Current Health kit functionality and design, being sick with COVID-19 infection, caretaking responsibilities, and comfort of the wearable device.

First, they had to navigate through a web-based enrollment process that included clicking on an advertisement that took the participant to a brief eligibility screening questionnaire and onto an appointment booker to connect with a member of the study team for consent. Although 2 participants said that the enrollment process was smooth and easy, another participant reported some minor problems that required contact with the study team:

> *I think it was pretty easy.* [RS008]

> *So when I went to enroll, it didn't give me a time slot, like it said that there was no one available. I guess you have to like, talk to someone at first and I remember it led me - it led me all the way to the end. But then I said, like, there was no - no one available for my initial call... So I emailed and then they were available. But it was like, still on the same day. I feel like it was a glitch or something.* [RS001]

Second, participants were required to set up the Current Health kit that was delivered to their home, select the correct-sized armband, charge the device and tablet, insert the device into the armband, and begin wearing it. In all, 2 participants described the setup as an easy process with a participant providing negative feedback on the tablet stand, which they described as nonessential and fussy:

> *Because like, one of the first things in the instructions is...take out the stand, put the...um...put the tablet on the stand. Like I just said, you could take all of that out because at first I thought I had to do it for it to get started and I didn't, and it wasn't like standing up and it just seemed like a waste.* [RS001]

> *It was super easy in the box set up that, you know, getting the tablet and everything and then getting the little charging dock. And I mean, it was easy and then I got it connected to my wifi and started wearing it that day.* [RS006]

The third participant experienced problems during setup. The Current Health kit shipping was delayed, and when she began setting it up, there was a problem with connectivity, which she reported took her a few hours to work out:

> *And then once I got the stuff here, yeah I started setting it up and then either the mobile or the wifi wouldn't work...it just didn't want to connect the wifi or wearable device.* [RS008]

Through the process of enrollment and setting up the wearables, participants were sick with COVID-19 infection, which meant that the usual barriers to joining a study may have been more difficult than usual. As a study team, we attempted to make the process as easy as possible for the participants:

> *It [enrollment] was super easy to me. I mean, even while, you know, sick as a dog with COVID I was still able to navigate and do it. So if I was able to, then you know, anybody could as long as you read and understand what you read, you can do it.* [RS006]

Furthermore, study adherence may have been more difficult for participants who had caretaking responsibilities:

> *It's hard when you're a caretaker and you've got, you know, your mom with breast cancer. So you have to keep her schedule up plus her meds. Plus your schedule and your meds and then hubby and his meds and his schedule. It can get overwhelming, I guess, but it was just because I probably was out of my routine.* [RS006]

Adherence could also be influenced by the comfort of the wearable device, which we asked participants to wear as close to 24/7 as possible, only removing the wearable device to shower, bathe, swim, or charge. They could switch arms but needed to wear it next to their skin, under their clothes. In all, 2 participants said that it was comfortable and did not give them any problems, even during sleep. A third participant provided suggestions for improvement:

> *Maybe if the band was...it could get...more air towards it, so you won't get so much sweat under it...It's you know, really gross with activity sometimes...so maybe just more airing.* [RS008]

Knowledge Combined with Skill

The domains of Knowledge and Skills were combined. Participants had to acquire an understanding of how to use the wearable device and tablet for adherence. Participants were provided with a printed and digital version of the QuickStart Guide (QSG) in the Current Health box and on the tablet. All participants reported using the hard copy QSG, which was positioned to be very visible upon opening the box, and none of the participants were aware that the QSG was also available on the tablet:

> *I remember getting everything [Current Health kit] and then I just - as soon as you open the box, I mean, like literally step by step, as long as you follow that booklet inside. That's what I did. I read it first. And then I started looking at stuff and I went back and I*

> *was like, OK, step 1 is this, step 2 is this. I mean, it hooked up in like literally ten minutes.* [RS006]

When a participant encountered problems with setup or connection, they had access to the study team and Current Health's 24/7 technical support to get things working:

> *I mean, I couldn't connect, of course. Then finally I was like, OK...then I realized I could reach out to the email person or the person that was like head of tech things like and say, OK, it's not working.* [RS008]

An important aspect of technical support is the speed at which they respond, so that the participants do not become frustrated and no data are lost:

> *...they contacted me pretty fast, so...I didn't think I was going to have a response like that.* [RS008]

Participants had to develop skills to interact with the tablet to take daily symptom surveys and to ensure that the tablet and wearable device were charged and working. They also had to experiment with the device fit to ensure comfort:

> *It's [survey] very, very easy to understand straight to the point, like you ask exactly what you need to know. And I love how it gets to all the symptoms, you can hit yes or no. And then it even asks you okay is it better than the day before this time or worse. I have loved that because some of my symptoms are a lot better and some of them are staying worse or getting worse and it varies daily. And so, yeah, I love that. That's really cool.* [RS006]

> *Now, if you've got the band too tight on your arm, your arm will hurt. That's a learning process whenever you're starting. I did that...Then I got it loosened up and it's like perfect now.* [RS006]

Social or Professional Role and Identity

The participants talked about their reasons for joining the study and contributing their time and effort. They were motivated to help in what few ways they felt they could, especially when it was difficult to help beyond isolation at home. Having a sense of contribution to efforts around the COVID-19 pandemic is important:

> *Look, I'm trying to be a good...I'm trying to be a good human. We're trying to quarantine and stay away from people.* [RS006]

> *...well, I'll apply and see and help out the community and help out the hospital or where all this data was going to go to help you guys. See if it would do anything good for covid.* [RS008]

Beliefs About Capabilities

The participant who had problems with the setup of the Current Health kit showed particular determination in working through the issues and troubleshooting until she could get it working. Although she had access to technical support, she was determined to troubleshoot initial connection problems independently:

> *And then once I got the stuff here [Current Health kit], I started setting it up and then either the mobile or the wi-fi wouldn't work... I tried doing stuff on my own... it just didn't want to connect to the wi-fi or wearable device…I was like OK, I'm not going to play with this anymore. And then stayed up, like all night cause I was like OK, I am not letting this thing beat me. I was just determined to like…I'm going to figure this out somehow and then yeah... I don't even know what time it was, I don't know, it was like twelve or one-o'clock in the morning when I finally got it to work.* [RS008]

Optimism

A participant showed tremendous optimism in the face of the COVID-19 pandemic and her own personal trouble in being sick with the disease. This participant focused on what it was teaching her and helping her focus on gratitude for aspects of life that were going well:

> *It's [COVID] definitely teaching us and it's making me learn and making me more aware and I'm thankful for it if you wanna know the truth.* [RS006]

> *Yeah, that's...that's a positive way to look at things, huh?* [Interviewer]

> *Yes, ma'am. I mean, well, you got to be positive. I was already down to the bottom. You know, I was already at the bottom of the COVID barrel.* [RS006]

Although we did not hear that our participants expressed pessimism in relation to the COVID-19 pandemic or their involvement in the study, we asked them if there were any negative aspects to participating in the study. The participants mentioned that charging and remembering to charge the device and tablet were inconvenient tasks. A participant said that you could become tired of wearing the device, but she did not mind wearing it. Wearing the device could also pique the curiosity of strangers:

> *The only negative thing is people ask, what is that on your arm?* [RS006]

> *Oh, interesting.* [Interviewer]

> *Yes, that would be I would say the only thing, just getting questions about what is it? What are you wearing? And so I just tell them, I'm in a medical study since I had COVID and I'm just giving all my data and vitals and having stuff recorded. That's what I tell them.* [RS006]

Beliefs About Consequences

The Current Health platform at the time of this study did not relay any information to the participants. It took some time for participants to adjust to wearing and charging the device and trusting that the data were being transmitted appropriately. Participants did not always know if data would be lost if they left their homes. The biggest issue that came up for participants around charging the device was not knowing how much battery it had left, making it difficult to plan charging the device. The tablet was less of an issue because they could leave it on the charger and only use it once per day to take and submit their daily symptom survey. Several suggestions have been made to

make the battery life more apparent to the user, which are now being integrated into the product:

> *I don't know if it's possible, but like if it told you if you needed to charge your device. Like I know it tells you, please recharge it but if there's like a way it told you that it had a low battery the actual like...I have no idea if it's possible but like, if it somehow would like, oh, your little arm band has low battery, charge it.* [RS001]

> *So I guess, you know, I wasn't sure if like, it would work if I wasn't in my house.* [RS001]

After a power outage, a participant expressed concern over the lost data and whether the device was still transmitting the data:

> *I was having an issue about the connection and I emailed [the Study Coordinator] and I was like, are you getting this [vital signs data] and she said, Yes, you're fine, you're good, it's okay. Cause we had like a power outage so our Internet went out and everything. And I was like, Are you still getting this? Yes, it holds data for eight to ten hours. I said, Okay, just making sure because I thought I'd done messed it up [laughs].* [RS006]

We found this domain to be linked with a concern that participants had about their own study adherence (ie, transmitting vital signs data) and desire to participate in the study to the best of their ability.

Reinforcement

A participant felt that the study would be more engaging if she could see some of her own data, and plans are underway to allow participants and patients to receive feedback on their data from wearing the device.

Participants were asked what they thought about the US $100 they were offered for the time and effort they took to participate in the study. Participants received this payment after the study ended and the equipment was returned to Current Health. All participants thought it was a nice gesture but said it was not the thing that motivated them to participate:

> *I think it's fine, like it didn't...it didn't sway me to participate or not participate...It's just a nice added bonus.* [RS001]

> *I think it's more than fair...I mean, all you're doing is just wearing this device. It's not like you're having to drive anywhere. You're not having to go out of your way. All you have to do is hook up some equipment, wear the device and save the box. And then when you're done, pack it all up and send it in. Woo hoo, I mean. It is not that hard. So I mean, I think it's very fair.* [RS006]

> *I think it's like a nice thing...I didn't do it for the money, I mean, I'm still waiting on that.* [RS008]

Memory, Attention, and Decision Processes

The participants in the RiskSEARCH study had high levels of adherence to wearing the device and taking the daily survey (see the aforementioned results). For the few times they did not answer the symptom survey or wear the device, we asked them to think about the reasons. A major contributor to not wearing the device was forgetting to put it back on after removing it for charging, bathing, or showering. For the tablet, it was forgetting to take the survey:

> *Yes, I think I forgot to put it on...And then other than that, I don't think I...I did stop it early because I went on vacation and I didn't bring it with me [participant asked to stop participating before going on vacation as symptoms had resolved].* [RS001]

> *I think I pretty much wore it all the time. Sometimes...like whatever going to the shower or whatever, then maybe I...might've like left it on the charger a little bit longer.* [RS008]

Behavioral Regulation

Participants spoke about the importance of routine and habits for remaining adherent by wearing a charged device and having a charged tablet to take daily symptom surveys. On days when a participant was out of routine or when normal habits could not be completed, there was an increased risk of forgetting to do these things:

> *I was literally running all day long from home like 9:00 that morning until about 5:30 yesterday afternoon is when I finally stopped and got home. And when I came in and made something to eat, I didn't even...I got, I was out of my routine that day. And I didn't even think about it. I didn't even think about it until three o'clock this morning.* [RS006]

Social Influences

Participants had access to the Current Health technical support 24/7 and the study team close to 24/7. It was critical that we not lose participants or otherwise good adherence because of problems or questions that could not be responded to quickly. The participants could access technical support over telephone and the study team over telephone, email, or text. Participants spoke highly about the study's clinical research associate who was primarily responsible for enrolling participants and helping them get set up.

> *The lady that I was coordinating with was...she was super sweet, she was super informative. Anytime I had a question all I had to do with text or message, and she literally got back to me that same day.* [RS006]

We also looked for feedback about study aspects, such as communication, frequency of messaging, and the Current Health kit itself. We asked in-depth questions about their experience using the study website, study-related emails, texts, and telephone calls. Participants found the methods of communication acceptable and unobtrusive and said that they liked having several avenues available to them for communication with the study team.

Discussion

Principal Findings

To gain a more thorough understanding of participant experience in a fully remote virtual trial during a global pandemic, semistructured interviews were conducted with 3 of the 7 participants in the RiskSEARCH study. All participants met our criteria for being fully adherent to the study and reported through interviews or the TUQ that participation requirements in both the main study (up to 30 days of wearing a wearable device on the upper arm, responding to daily surveys, and communicating with the study team when necessary) and the substudy were feasible and low burden.

Having quick and easy access to support from the study team and Current Health technical support was a critical enabling factor for staying engaged [24]. Future virtual studies should ensure that this resource, a dedicated and responsive study team or technical support, is accessible to participants (Textbox 4). This may necessitate staffing across time zones or during out-of-office hours. Participants reported some barriers around setting up the Current Health kit, keeping devices charged, and remembering to take surveys but described these as minor challenges and showed high adherence while wearing the device and submitting responses to daily symptom surveys. Combining subjective (qualitative and self-report) and objective (quantitative like the number of surveys submitted and hours of vital signs data transmitted) data, the researchers assumed that high adherence was at least partly tied to ease of participation. A participant reported high levels of adherence (ie, reported not wearing the device on only 2 to 3 occasions that were a few hours long), while objectively showing 63% wearable device adherence. In reviewing the data, we believe that some data may have been lost when the participant was away from the home hub for >8 hours. Several factors may contribute to the differences in objective and subjective measures when collecting remote data such as perception, unknown technical issues, or improper use of digital technology. We found that the overall motivation for enrolling was a wish to contribute positively to the COVID-19 effort. In this small sample, we found adherence to be the easy part, whereas the key challenge for the research team was recruitment to the main study amidst the rapidly shifting landscape of the pandemic.

Textbox 4. Recommendations.

Recommendations

- Provide participants quick and easy access to support from the study team or technical support for any digital health technologies used in the study.
- Collect in-depth information around factors that impact on study enrollment and adherence; for example, processes, communications with study team, advertisements, and trouble setting up or using technology.
- Identify discrepancies in subjective and objective measures of study adherence and try to understand why those might exist; for example, participant perception, unknown technical issues, or improper use of digital technology.
- Identify agreement in subjective and objective measures of study adherence to understand what is working well.
- Consider using the Theoretical Domains Framework (or similar framework) for assessing potential implementation problems in virtual trials.

The interview schedule (Multimedia Appendix 2) was developed with the purpose of understanding the participant's experience of interacting with the Current Health kit and the ability and motivation to adhere to the study intervention. This interview schedule can be used in other qualitative studies looking to identify components of the study, such as digital technology and study materials, that facilitate or prevent adherence. We explored the barriers and facilitators to the web-based study enrollment process, communication with the study team, study advertisements and recruitment messaging, troubleshooting, burden of participating in the study, ease of use of the Current Health kit, and benefits and disadvantages of participating in the RiskSEARCH study. The data were then inductively coded into themes related to the TDF domains. The TDF is frequently used to develop interview schedules, with interviews coded into 14 domains of the TDF. The authors could not find research conducted as it was for this study, with an interview schedule developed independently of the TDF with themes from analyzed interview data and then mapped onto the TDF.

The TDF worked well for our qualitative data and the process of mapping inductive themes onto domains from the TDF was relatively straightforward. We chose the TDF because it was developed to make behavior change theories more accessible to implementation researchers [18]. It was revised and validated in 2005 [25] to help researchers identify and describe the barriers and facilitators that could influence behavior and thereby impact implementation. Evidence suggests that theoretically based health interventions are more successful than interventions with no such base [26]. We needed a method for theoretically assessing any potential implementation problems encountered while running the RiskSEARCH study; the TDF provided this method. We are unaware of the use of TDF in other virtual studies or VCTs.

The interview schedule, created independently of the TDF, produced themes that mapped most heavily onto the domains *Environmental Context and Resources* and *Knowledge* combined with *Skills*. We believe this is an indication of the critical aspects of the product itself, the built environment of the participants, and the knowledge and skill acquisition that are essential for setting up and using the Current Health kit.

There were 3 domains that we did not map any themes to: *Intentions*, *Emotions*, and *Goals*. Though we did not map any themes to the TDF domain of *Intentions* it was clear throughout each interview that participants made a conscious effort to be fully adherent by wearing the device as long and as often as

possible, answering the daily symptom survey, and returning the Current Health kit once the study was over. As for *Emotions*, we found that participants were content to wear the device and be adherent. For the one participant who had trouble setting up her Current Health kit, she did not describe feeling frustrated or annoyed but simply determined to get her Current Health kit working without assistance, although knowing that technical support was available. This was not a study designed to get participants to create and follow goals, which is why we likely did not have any themes that could be mapped to the domain *Goals*. In future interview schedules, we could consider targeting these domains to obtain the most complete picture of implementation using the TDF.

Limitations

The biggest limiting factor in this study was the sample size. Although "data saturation" is a term with some conceptual and methodological issues and is not a necessity in every type of qualitative research [27,28], this study would have benefited from more interviews and in particular from interviewees who were different ages, male, and not adherent or those who experienced technical challenges, troubleshooting, and had opportunities to develop strategies around device charging and remembering to complete the daily symptom survey. There is also a possibility that more interviews would have led to more themes that could have been mapped onto the 3 domains of the TDF, for which we did not have data. However, reaching saturation does not necessarily invalidate our findings [29]. Despite the low number of interviews, we believe these exploratory findings add value to identifying barriers and facilitators to adherence in virtual studies and specifically, studies that require using wearables and submitting daily digitally delivered surveys.

Future Work

We hope to expand these preliminary findings to future virtual studies and VCTs that will surely outlive the current COVID-19 pandemic [30]. As a study team, we are highly motivated to reduce the burden placed on study participants to make study adherence as easy and enjoyable as possible and to encourage a more diverse study population by removing logistical barriers to participation [31]. The findings from this exploratory research will contribute to the best-practices literature for VCTs and help improve the Current Health kit and study delivery for future research participants. We believe that there is also more to learn about motivating factors for participants willing to enroll and participate in infectious disease research. We are also developing a means of providing the participant's own data to them to help with engagement and memory around charging and wearing the device and believe that this will affect adherence metrics and overall levels of study engagement.

As the COVID-19 pandemic stretches on and the need for VCTs continues to grow, there is also a need for continuing research that helps us understand the participants' experience of engaging in these studies, including the barriers and enabling factors that influence adherence. The RiskSEARCH study did not progress beyond the pilot study because of limited recruitment, highlighting an ongoing need to improve recruitment for clinical studies, whether virtual or in person. Despite the problems with recruiting, we believe we have learned some critical lessons about the conduct of virtual study or trial presented in this paper and have produced tools to continue collecting this type of data.

Conclusions

Participants in the RiskSEARCH substudy showed high levels of adherence and engagement throughout their participation. Although participants experienced some challenges in setting up and maintaining the Current Health kit (eg, charging devices), they reported feeling that the requirements of participation were both reasonable and realistic. We have shown that the TDF can be used for inductive thematic analysis. We anticipate expanding this work in future virtual studies and trials to identify barriers and enabling factors for implementation.

Acknowledgments

This work was supported by Current Health, Ltd. Additional time on the Risk Stratification and Early Alerting Regarding COVID-19 Hospitalization main study was supported in part by federal funds from the Department of Health and Human Services, Office of the Assistant Secretary for Preparedness and Response; Biomedical Advanced Research and Development Authority, Division of Research Innovation and Ventures under contract 75A50120C00190.

Authors' Contributions

JP was involved in finalizing the study protocol, implementing the study processes, conducting interviews with participants, analyzing the qualitative data, and drafting the manuscript. JLT was the RiskSEARCH study coordinator responsible for implementing the study processes and drafting the quantitative metrics sections. NZ was involved in implementing the study processes, developing the graphic display for symptoms in the quantitative section, and reviewing the manuscript. AW and MW were involved in finalizing the study protocol and reviewing the manuscript.

Conflicts of Interest

JP, JLT, MW, AW, and NZ are employees of Current Health, a Best Buy company, and receive compensation that supports the conduct of this research.

Multimedia Appendix 1
Inclusion and exclusion criteria to the Risk Stratification and Early Alerting Regarding COVID-19 Hospitalization main study.

[PDF File (Adobe PDF File), 11 KB - formative_v6i7e37567_app1.pdf]

Multimedia Appendix 2
Risk Stratification and Early Alerting Regarding COVID-19 Hospitalization interview schedule.
[PDF File (Adobe PDF File), 65 KB - formative_v6i7e37567_app2.pdf]

References

1. Onyeaka H, Anumudu CK, Al-Sharify ZT, Egele-Godswill E, Mbaegbu P. COVID-19 pandemic: a review of the global lockdown and its far-reaching effects. Sci Prog 2021;104(2):368504211019854 [FREE Full text] [doi: 10.1177/00368504211019854] [Medline: 34061685]
2. Ledford H. Coronavirus shuts down trials of drugs for multiple other diseases. Nature 2020 Apr;580(7801):15-16. [doi: 10.1038/d41586-020-00889-6] [Medline: 32214240]
3. Marra C, Gordon WJ, Stern AD. Use of connected digital products in clinical research following the COVID-19 pandemic: a comprehensive analysis of clinical trials. BMJ Open 2021 Jun 22;11(6):e047341 [FREE Full text] [doi: 10.1136/bmjopen-2020-047341] [Medline: 34158302]
4. GlobalData Healthcare. Significant increase in virtual trial deals since 2020 as a result of Covid-19. Pharmaceutical Technology. 2021 Aug 16. URL: https://www.pharmaceutical-technology.com/comment/significant-increase-virtual-trial-2020-covid-19/ [accessed 2021-09-17]
5. The rise of real-world data and digital tools: Conducting remote clinical trials in a pandemic. Open Access Government. 2020 Dec 18. URL: https://www.openaccessgovernment.org/remote-clinical-trials/100566/ [accessed 2021-09-17]
6. Ali Z, Zibert JR, Thomsen SF. Virtual clinical trials: perspectives in dermatology. Dermatology 2020;236(4):375-382 [FREE Full text] [doi: 10.1159/000506418] [Medline: 32126560]
7. Inan OT, Tenaerts P, Prindiville SA, Reynolds HR, Dizon DS, Cooper-Arnold K, et al. Digitizing clinical trials. NPJ Digit Med 2020 Jul 31;3:101 [FREE Full text] [doi: 10.1038/s41746-020-0302-y] [Medline: 32821856]
8. Pfammatter AF, Mitsos A, Wang S, Hood SH, Spring B. Evaluating and improving recruitment and retention in an mHealth clinical trial: an example of iterating methods during a trial. Mhealth 2017 Nov 1;3:49 [FREE Full text] [doi: 10.21037/mhealth.2017.09.02] [Medline: 29184901]
9. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Sciences Policy, Forum on Drug Discovery, Development, and Translation. In: Shore C, Khandekar E, Alper J, editors. Virtual Clinical Trials: Challenges and Opportunities: Proceedings of a Workshop. Washington, DC, USA: National Academies Press (US); Jul 23, 2019.
10. Dorsey ER, Venuto C, Venkataraman V, Harris DA, Kieburtz K. Novel methods and technologies for 21st-century clinical trials: a review. JAMA Neurol 2015 May;72(5):582-588 [FREE Full text] [doi: 10.1001/jamaneurol.2014.4524] [Medline: 25730665]
11. Gul RB, Ali PA. Clinical trials: the challenge of recruitment and retention of participants. J Clin Nurs 2010 Jan;19(1-2):227-233. [doi: 10.1111/j.1365-2702.2009.03041.x] [Medline: 20500260]
12. Brøgger-Mikkelsen M, Ali Z, Zibert JR, Andersen AD, Thomsen SF. Online patient recruitment in clinical trials: systematic review and meta-analysis. J Med Internet Res 2020 Nov 04;22(11):e22179 [FREE Full text] [doi: 10.2196/22179] [Medline: 33146627]
13. Skelton E, Drey N, Rutherford M, Ayers S, Malamateniou C. Electronic consenting for conducting research remotely: a review of current practice and key recommendations for using e-consenting. Int J Med Inform 2020 Nov;143:104271 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104271] [Medline: 32979650]
14. 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]
15. Remote Patient Monitoring. Current Health. URL: https://currenthealth.com/platform [accessed 2022-01-06]
16. Risk Stratification and Early Alerting Regarding COVID-19 Hospitalization (RiskSEARCH). Clinical Trials. 2021 Aug 19. URL: https://clinicaltrials.gov/ct2/show/study/NCT04709068 [accessed 2021-09-17]
17. Lever Taylor J, Wilkes M, Pugmire J, Zahradka N. The challenges of decentralized recruitment during the COVID-19 pandemic. In: Society of Clinical Research Associates (SOCRA) Annual Conference 2021. 2021 Sep 23 Presented at: SOCRA '21; September 22-25, 2021; Virtual.
18. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implement Sci 2017 Jun 21;12(1):77 [FREE Full text] [doi: 10.1186/s13012-017-0605-9] [Medline: 28637486]
19. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
20. Parmanto B, Lewis Jr AN, Graham KM, Bertolet MH. Development of the telehealth usability questionnaire (TUQ). Int J Telerehabil 2016 Jul 1;8(1):3-10 [FREE Full text] [doi: 10.5195/ijt.2016.6196] [Medline: 27563386]

21. NVivo. QSR International. 2020. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home [accessed 2020-11-08]
22. Braun V, Clarke V. Thematic Analysis: A Practical Guide. Thousand Oaks, CA, USA: Sage Publications; 2021.
23. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
24. Zweben A, Fucito LM, O'Malley SS. Effective strategies for maintaining research participation in clinical trials. Drug Inf J 2009 Jul;43(4):10.1177/009286150904300411 [FREE Full text] [doi: 10.1177/009286150904300411] [Medline: 24311825]
25. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
26. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31:399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
27. O'Reilly M, Parker N. 'Unsatisfactory Saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. Qual Res 2012 May 17;13(2):190-197. [doi: 10.1177/1468794112446106]
28. 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]
29. Morse JM, Field PA. Qualitative Research Methods for Health Professionals. Thousand Oaks, CA, USA: Sage Publications; 1995.
30. Erridge S, Majeed A, Sodergren M. Virtual trials: looking beyond covid-19. The BMJ Opinion. 2020 Jul 6. URL: https://blogs.bmj.com/bmj/2020/07/06/virtual-trials-looking-beyond-covid-19/ [accessed 2022-01-18]
31. Bodicoat DH, Routen AC, Willis A, Ekezie W, Gillies C, Lawson C, et al. Promoting inclusion in clinical trials-a rapid review of the literature and recommendations for action. Trials 2021 Dec 04;22(1):880 [FREE Full text] [doi: 10.1186/s13063-021-05849-7] [Medline: 34863265]

Abbreviations

ICF: informed consent form
QSG: QuickStart Guide
RiskSEARCH: Risk Stratification and Early Alerting Regarding COVID-19 Hospitalization
TDF: Theoretical Domains Framework
TUQ: Telehealth Usability Questionnaire
VCT: virtual clinical trial

Edited by A Mavragani; submitted 08.03.22; peer-reviewed by C Carlin, H Karim, V Block, R Ciorap; comments to author 03.05.22; revised version received 19.05.22; accepted 19.05.22; published 05.07.22.

Please cite as:
Pugmire J, Lever Taylor J, Wilkes M, Wolfberg A, Zahradka N
Participant Experiences of a COVID-19 Virtual Clinical Study Using the Current Health Remote Monitoring Platform: Case Study and Qualitative Analysis
JMIR Form Res 2022;6(7):e37567
URL: https://formative.jmir.org/2022/7/e37567
doi:10.2196/37567
PMID:35671408

Original Paper

Measures of Daily Activities Associated With Mental Health (Things You Do Questionnaire): Development of a Preliminary Psychometric Study and Replication Study

Nickolai Titov[1], PhD; Blake F Dear[1], PhD; Madelyne A Bisby[1], PhD; Olav Nielssen[1], PhD; Lauren G Staples[1], PhD; Rony Kayrouz[1], PhD; Shane Cross[1], PhD; Eyal Karin[1], PhD

eCentreClinic, School of Psychological Sciences, Macquarie University, Sydney, Australia

Corresponding Author:
Nickolai Titov, PhD
eCentreClinic
School of Psychological Sciences
Macquarie University
73 Talavera Road
Sydney, 2109
Australia
Phone: 61 0488 991 122
Email: nick.titov@mq.edu.au

Abstract

Background: A large body of research has identified modifiable cognitions and behaviors (actions) associated with psychological health. However, little is known regarding the actions that are most strongly associated with psychological health or the frequency with which they should be performed.

Objective: This paper described 2 studies that used survey methodology to create the Things You Do Questionnaire (TYDQ), which aims to identify and rank actions (items) and domains of actions (factors) most strongly associated with psychological health.

Methods: We used digital marketing strategies to recruit Australian adult participants, who were asked to complete 2 web-based surveys comprising versions of the TYDQ; validated measures of depression, anxiety, and satisfaction with life; and demographic questions. In study 1, a total of 3040 participants rated how often they performed each of the 96 items comprising the TYDQ. This design was replicated in study 2, in which a 59-item version of the TYDQ was completed by 3160 participants. In both studies, the factor structure and validity were examined, as were the associations between individual TYDQ items and 3 mental health outcomes: depression, anxiety, and satisfaction with life.

Results: In study 1, factor analyses revealed that a 5-factor model comprising 27 items achieved an optimum balance between brevity and variance and accounted for 38.1%, 31.4%, and 33.2% of the variance in scores on measures of depression, anxiety, and satisfaction with life, respectively. The factors were interpreted as realistic thinking, meaningful activities, goals and plans, healthy habits, and social connections. These 5 factors were more strongly associated with psychological health than those such as practicing kindness, exercising gratitude, and practicing spirituality. This pattern of results was replicated across gender, age groups, and depression severity. The 5-factor solution found in study 1 was replicated in study 2. Analyses revealed that a 21-item version accounted for 46.8%, 38.2%, and 38.1% of the variance in scores on measures of depression, anxiety, and satisfaction with life, respectively.

Conclusions: These findings indicate that some actions are more strongly associated with psychological health than others and that these activities fall within 5 broad domains, which represent skills often taught in psychological treatments. Subsequent studies are planned to explore the reliability of these items and results in other samples and to examine patterns of change in scores during treatment for anxiety and depression. If replicated, these efforts will assist in the development of new psychological interventions and provide an evidence base for public mental health campaigns designed to promote good mental health and prevent the emergence of common mental disorders.

(JMIR Form Res 2022;6(7):e38837) doi:10.2196/38837

KEYWORDS

anxiety; depression; satisfaction with life; COVID-19; behavior; habits; cognitions; survey; mechanisms; psychological well-being

Introduction

Depressive and anxiety disorders are highly prevalent conditions associated with significant distress, disability, and reduced quality of life [1,2]. These conditions exist along a continuum of severity, and it is now recognized that subclinical or subsyndromal variants can also cause functional disability [3] and considerable economic costs [4]. These disorders are characterized by maladaptive patterns of cognition, behavior, habits, and interpersonal and social engagement, which form the basis for accepted diagnostic criteria [5,6]. Maladaptive patterns of thinking and behavior trigger symptoms and impede recovery. Although the specific mechanisms of change during psychotherapy are not well understood [7-9], the premise that these maladaptive patterns can be modified or changed underlies the teaching of psychological skills within several treatment approaches, including cognitive therapy [10], behavior therapy [11,12], mindfulness [13], and interpersonal therapy [14].

Therefore, identifying adaptive and maladaptive patterns is a priority during treatment, and several approaches have been used to deconstruct these patterns into measurable and modifiable cognitions and behaviors (*actions*). One approach involves creating questionnaires to measure the quality and frequency of use of cognitive and behavioral skills (cognitive behavioral therapy) taught in psychological treatments [15-21]. Examples of items in these questionnaires include *I made an effort to evaluate my negative thoughts by considering just the facts* [19] and *I worked on a project that was meaningful to me* [21]. Consistent with psychological models, these and other studies have found small to moderate relationships between the frequency and quality of cognitive behavioral therapy–related actions and symptom severity over treatment. Although useful in clinical settings, disadvantages with many of these questionnaires include their tendency to focus solely on skills taught in therapy, their frequent use of highly technical language to describe cognitions and behaviors, and the exclusion of a large range of other activities that might be linked to mental health outcomes and that might represent important targets for intervention.

Other researchers surveyed consumers and mental health professionals to generate lists of actions believed to affect the symptoms of anxiety and depression [22-26]. Examples of actions identified in these studies include *focusing on the positive* [25] and *doing more things you enjoy* [22]. Many of the actions identified in this research are consistent with the skills taught and promoted in psychological treatments. Unsurprisingly, studies have also shown that prompting these types of actions is associated with symptom reduction [27]. However, a weakness of many of these measures is that they fail to inquire about the frequency of performing such actions. Such information can be important for developing interventions to improve mental health or prevent mental ill health.

An additional and parallel body of work has sought to examine the relationship between mental health and quality of life. This has led to the development of numerous measures, such as the Quality of Life Inventory [28] and the World Health Organization Quality of Life Scale [29], which have been demonstrated to reliably and validly measure satisfaction across multiple life domains (eg, physical health, psychological health, and relationships) while being reliable across both general and clinical populations. More recent quality of life measures, such as the Brunnsviken Brief Quality of Life Scale [30], have aimed to create freely accessible and brief versions that maintain strong validity and reliability profiles. Although helpful in determining satisfaction levels across different life domains, quality of life measures have not been designed to inquire about the frequency with which people engage in modifiable activities, which limits the utility of these tools.

Several research domains have identified actions that are associated with psychological health; however, several important questions remain unanswered. First, we do not know how often actions need to be performed each day or week for a person to experience an improvement in psychological health. Second, we do not know which actions or domains of actions are most strongly associated with improved mental health. Third, research has not yet explored whether particular actions are more helpful for different people according to their clinical presentation, personality, or demographic characteristics.

The studies presented in this paper are part of a research program, the Things You Do Project, designed to examine these questions with the overall aim of identifying and confirming the everyday actions that are most strongly associated with good mental health. The potential value of this information is not only in the treatment of mental disorder but also in the evidence-based prevention of mental disorder and promotion of mental health at the population level. Consistent with the World Health Organization's definition of health as not only the absence of symptoms but also the presence of a positive state of well-being [31], we used the term *psychological health* in this paper to describe not only the absence or reduction in symptoms of anxiety or depression but also an improvement in well-being, measured based on self-reported satisfaction with life.

In this paper, we have described 2 studies that used survey methodology to create the Things You Do Questionnaire (TYDQ), which aims to identify and rank actions and domains of actions most strongly associated with psychological health. Recognizing the complexity of this objective and the range of potential methods, we opted for a broad and systematic approach to scale development [32].

In the first study, participants rated the frequency with which they completed a range of actions previously identified to be associated with psychological health. Responses were analyzed to assess the factor structure, reliability, validity, and strength of association of TYDQ items with psychological health. In the second study, participants rated a shorter version of the TYDQ. In both studies, participants also completed validated measures of depression, anxiety, and satisfaction with life. Formal

hypotheses were not proposed, but we expected that the process of ranking actions and domains of actions would generate a parsimonious list of psychological health items.

Study 1

The aims of study 1 were to (1) generate a large list of actions previously associated with psychological health and explore the relationship between the weekly frequency of performing these actions and symptoms of anxiety, depression, and life satisfaction; (2) explore the underlying factor structures of the TYDQ and rank the association between the actions and factors and psychological health; (3) explore the importance of different actions across different subgroups; and (4) develop a parsimonious list of items.

Methods

Ethics Approval

A web-based survey design was used for data collection. The study was reviewed and approved by the Macquarie University Human Research Ethics Committee (MQ HREC: 5201700988), and informed consent was obtained from all participants. All methods were performed in accordance with the relevant guidelines and regulations.

The data sets used and analyzed during this study are available from the corresponding author (NT) upon reasonable request.

Questionnaire and Item Generation

An initial list of items was derived from a review of the literature on actions associated with psychological skills identified through multiple sources. These included reviews of (1) skills typically taught in psychological therapies (eg, cognitive skills, behavioral activation, exposure, goal setting, problem solving, social skills training, applied relaxation, mindfulness, acceptance, gratitude, and kindness); (2) actions associated with psychological health, such as healthy daily routines (eg, sleep hygiene, nutrition, and moderated use of electronic devices), physical health (eg, exercise, yoga, stretching, and walking), social activity (eg, talking to friends and spending time with people), meaningful and satisfying activities (eg, doing fun activities), and spiritual and religious activities; (3) symptoms of high prevalence mental disorders as described in psychiatric diagnostic systems [5,6]; (4) previous questionnaires developed by the authors [20], other questionnaires, or similar lists [17-19,22,23]; and (5) consultations with colleagues.

Our research team, consisting of psychiatrists, clinical psychologists, and data analysts with experience in questionnaire development, created a preliminary list comprising >500 actions. Each action was phrased as a verb, and references to psychological techniques or orientations, such as practicing mindfulness, were avoided. During the planning phase, we met regularly to identify, review, edit, and categorize the list, which was reduced to 272 items after duplicates and overlapping items were removed. The items were categorized according to the mechanisms or processes associated with each item; that is, a priori attempts were made to identify theoretical clusters. For example, some items were categorized as primarily concerned with cognitive actions (primary cluster), and within this cluster, they were subdivided into actions associated with challenging unhelpful thoughts, focusing on the future (secondary clusters). In instances where items belonged to multiple clusters, categorization was based on consensus. To capture a broad range of items, the list was shared with colleagues outside the research group, who made further recommendations. The final list of 96 items was based on the following criteria: (1) actions that can be performed daily; (2) measurable, that is, a person can be expected to reliably estimate the frequency of actions over a week; (3) can be completed by most adults; (4) described as a positive action, that is, an action that can be considered a strength; (5) modifiable, that is, can be changed by the person; and (6) did not duplicate another item.

A 5-point Likert rating scale was used to ask participants how often they performed these behaviors over the past week by using the following scoring system: *0=Not at all (0 days per week)*; *1=1 or 2 days per week*; *2=3 to 4 days per week*; *3=5 to 6 days per week*; *and 4=every day (7 days per week)*. None of the items were reverse scored. After completing the TYDQ, participants were also invited to nominate other actions that were not included in the 96 items they used to improve their mental health.

Participants and Procedure

The survey was promoted to adults across Australia via websites and web-based newsletters from Australian mental health services and Facebook posts and advertisements. Advertisements invited people to participate in a study that aimed to *identify activities and habits that affect mental health*. Consenting participants provided demographic details and completed web-based questionnaires. Participants were required to be aged ≥18 years. No other exclusion criteria were applied. Details of the sample are included in the *Results* section.

Measures

Three standardized outcome measures were used to evaluate and rank the association of the items with psychological well-being.

Patient Health Questionnaire 9-Item Scale

The Patient Health Questionnaire 9-item (PHQ-9) scale measures the occurrence of Diagnostic and Statistical Manual IV–congruent depressive symptoms over the past 2 weeks by using a 4-point Likert scale ranging from 0 (*not at all*) to 3 (*nearly every day*). Higher scores indicate greater symptom severity [33]. The PHQ-9 has good internal consistency and is sensitive to change [34], with a total score of ≥10 usually but not always associated with a diagnosis of major depressive disorder [35,36]. The Cronbach α in this study was .90.

Generalized Anxiety Disorder 7-Item Scale

The Generalized Anxiety Disorder 7-item (GAD-7) scale measures the occurrence of general anxiety symptoms over the past 2 weeks by using a 4-point Likert scale ranging from 0 (*not at all*) to 3 (*nearly every day*). The GAD-7 is sensitive to Diagnostic and Statistical Manual IV–congruent GAD, social phobia, and panic disorder, with increasing scores indicating greater severity of symptoms [34,37]. The GAD-7 has been shown to demonstrate sound psychometric properties, with a

total score ≥10 usually indicating a likely diagnosis of an anxiety disorder [38]. The Cronbach α in this study was .92.

Satisfaction With Life Scale

The Satisfaction With Life Scale (SWLS) is a 5-item scale that measures attitudes toward life satisfaction by using a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*), with higher scores indicating greater life satisfaction. The SWLS has demonstrated good psychometric properties [39,40] and has been used extensively as a measure of life satisfaction in mental health research [41]. The Cronbach α of the SWLS in this study was .90.

Statistical Analyses

The examination of the relationship between the TYDQ items and the standardized scales was operationalized in three detailed and comprehensive steps: (1) item-level analysis, (2) factor analysis, and (3) generalizability of factor solutions.

Item-Level Analysis

First, a series of univariate regression models (ANOVA) was used to quantify and statistically test the association between each of the TYDQ item scores and the PHQ-9, GAD-7, and SWLS outcome scores. The strength of the association between items and outcomes was summarized with an R^2 metric, noting the proportion of the variance explained by each TYDQ item and each outcome. The R^2 values were used to weigh and rank the 96 items.

To further visualize the relative importance of different TYDQ items, each possible TYDQ item score, ranging from 0 (*not at all*) to 4 (daily), was tabulated with its corresponding PHQ-9, GAD-7, and SWLS outcomes to form a heat map [42]. Comparing the corresponding outcome scores with the range of TYDQ values enabled the identification of the minimal weekly behavioral frequency score needed to achieve optimal PHQ-9, GAD-7, and SWLS outcome scores for each TYDQ item. The optimal TYDQ score was defined as the minimum weekly behavioral frequency beyond which no additional statistically significant improvement can be observed.

Factor Formation and Analysis

Second, a series of exploratory factor analyses (EFAs) was used to examine how different TYDQ items formed composite latent factors of behavior. An initial EFA was used to identify the factors based on a full list of 96 items. An additional 3 EFAs were conducted to identify latent factors among the subsets of TYDQ items that were associated more closely with the PHQ-9, GAD-7, and SWLS. These EFAs analyzed the subsets of items based on a minimum of 5%, 10%, and 15% R^2 relationship to any one mental health outcome with the aim of identifying more parsimonious lists of factors closely related to psychological health. A factor was considered viable when the item set comprised a composite eigenvalue greater than 1. Additional diagnostic analyses were used to assess each of the identified factors, including item reliability analysis, mean item intercorrelation, and the ability of the identified factors to replicate each item (ie, commonalities).

All identified factors were tested and ranked for their strength of association with the 3 outcomes by using the same analytics and visualization approach used to evaluate TYDQ items. To further evaluate the importance of different behavioral TYDQ factors for mental health, a series of binary logistic regressions was conducted to test the association between the factor scores and the probability of an individual presenting with PHQ-9 or GAD-7 scores that would be considered within the clinical range.

Each factor was assigned a proposed label, which aimed to represent the overall theme of those items. Consensus regarding the factor label was achieved through discussion and debate, considering other labels could have been used.

Generalizability of Factors Solutions Across Different Subgroups

A series of confirmatory factor analyses (CFAs) were conducted to examine the reliability of the EFAs within a sample cross-validation analysis (5 randomized subsamples) and the validity of the factor structure along with key subgroups that differed based on age (<30, 30-44, 45-60, >60 years), gender (male, female, and other), the severity of depressive or anxiety symptoms (minimal-mild, moderate, and severe), an indication of employment, and tertiary education. These CFA tests followed the methodology and reporting conventions outlined by Putnick and Bornstein [43] for the examination of measurement invariance, including the examination of configural invariance, metric invariance, scalar invariance, and strict invariance. From the viewpoint of scale development, this step sought to evaluate the reliability and generalizability of the identified factor solutions along relevant subgroups in clinical research (dimensionality).

All analyses were performed using R statistical software version 4.1.1 [44] and the Lavaan package [45]. A conservative *P* value of .01 was considered the threshold for statistical significance, reflecting the large sample and the multiple number of tests conducted. In all analyses, participants with missing data were not included in the analyses owing to difficulty imputing unbiased missing outcomes in a high-dimensional analysis and cross-sectional settings. All EFAs and CFAs were based on weighted least-squares estimators to account for the ordinal categorical TYDQ item score range. For all EFAs, with the aim of identifying unique subsets of items, a varimax factor rotation was adopted to maximize the number of factors identified and minimize item-factor cross-loadings [46].

Items with factor loadings <0.50, a threshold that aimed to achieve a conservative balance between recommendations in the literature [32,47], were suppressed as were items with suboptimal association with their corresponding factor (accounting for <25% of the factor variance) [48].

Results

Item and Sample Description

Participants in study 1 were recruited from March to June 2018, during which time 3755 people consented to participate and 3040 (80.96%) completed the questionnaire. Sample characteristics are presented in Table 1. Participants had a broad

range of both demographics (age, gender, location, education, and employment) and mental health symptoms. Approximately half of the sample scored above the clinical cutoffs on the PHQ-9 (n=1710, 48.1%), GAD-7 (n=2020, 57.9%), and SWLS (n=2188, 64%), allowing the testing of the association of everyday behaviors for those above and below recognized clinical cutoffs on the PHQ-9 and GAD-7.

Multimedia Appendix 1 includes each of the 96 items ordered based on the a priori primary and secondary theoretical clusters created during item development and based on the observed average weekly score. Multimedia Appendix 1 also uses a heatmap to indicate the frequency with which participants reported performing each item each week. For example, the first item in Multimedia Appendix 1 was the 13th item presented in the survey and read I read, listened, or watched something I enjoyed. A priori, this item was categorized as an Activity/Meaning (primary and secondary clusters), and the average weekly score was 2.72 (mean TYDQ weekly score), indicating that the mean frequency reported was between 4 and 6 days each week. The final column indicates that 37.4% (n=1404) of the sample reported doing this action daily, whereas 3.7% (n=139) reported not doing this action in the previous 7 days.

As shown in Multimedia Appendix 1, the frequency of actions (TYDQ mean scores) ranged across and within clusters. For example, 2 of the highest frequency scores were observed within the *Healthy Routine* cluster, with daily showers or baths reported by 69.3% (n=2602) of the sample, and avoidance of illicit drugs or misuse of medication reported by 83.3% (n=3128).

Table 1. Sample characteristics of studies 1 and 2.

Variable and subgroup	Study 1 (year 2018; N=3755)	Study 2 (year 2020; N=3756)	Test of sample differences
Survey completion, n (%)			$\chi^2_{1,7510}$=13.1, *P*<.001
Completed survey	3040 (81)	3160 (84.1)	
Incomplete survey	715 (19)	596 (15.9)	
Sex, n (%)			$\chi^2_{2,7509}$=303.8, *P*<.001
Male	1806 (48.1)	1072 (28.5)	
Female	1924 (51.2)	2651 (70.6)	
Other	25 (0.7)	33 (0.9)	
Age (years)			t_{7509}=16.13, *P*<.001
Value, mean (SD)	41.9 (13.9)	47.7 (17.1)	
<30, n (%)	976 (26)	944 (25.1)	
30-45, n (%)	1416 (37.7)	1430 (38.1)	
45-60, n (%)	1137 (30.3)	925 (24.6)	
>60, n (%)	226 (6)	456 (12.1)	
PHQ-9[a] category, n (%)			$\chi^2_{1,7095}$=159.1, *P*<.001
Below cutoff (≥10)	1710 (48.1)	2227 (62.9)	
Above cutoff (<10)	1848 (51.9)	1311 (37.1)	
GAD-7[b] category, n (%)			$\chi^2_{1,6961}$=151.3, *P*<.001
Below cutoff (≥10)	1471 (42.1)	2497 (71.9)	
Above cutoff (<10)	2020 (57.9)	974 (28.1)	
SWLS[c] category, n (%)			$\chi^2_{1,6858}$=98.4, *P*<.001
Below cutoff (≥16)	2188 (64)	2578 (75)	
Above cutoff (<16)	1233 (36)	860 (25)	
Depression symptoms (PHQ-9), mean (SD)	10.7 (6.8)	8.5 (6.5)	t_{7095}=14.3, *P*<.001
Anxiety symptoms (GAD-7), mean (SD)	8.8 (5.8)	6.8 (5.6)	t_{6961}=14.17, *P*<.001
Satisfaction with life (SWLS), mean (SD)	18.4 (7.7)	20.8 (7.8)	t_{6858}=13.42, *P*<.001
Australian born, n (%)			$\chi^2_{1,6049}$=32.9, *P*<.001
No	632 (16.8)	829 (22.1)	
Yes	3123 (83.2)	2927 (77.9)	
Location, n (%)			$\chi^2_{2,7392}$=4.9, *P*=.09
Capital city	1921 (51.2)	2015 (54.5)	
Other urban region	934 (24.9)	898 (24.3)	
Rural or remote	841 (22.4)	785 (21.2)	
Education attained, n (%)			$\chi^2_{3,7391}$=75.9, *P*<.001
High school or less	744 (19.8)	589 (15.9)	
Other tertiary qualification or certificate	1200 (32)	1100 (29.7)	
Postgraduate degree	735 (19.6)	1040 (28.1)	
Undergraduate degree	1017 (27.1)	969 (26.2)	
Employment status, n (%)			$\chi^2_{5,7389}$=245.9, *P*<.001

Variable and subgroup	Study 1 (year 2018; N=3755)	Study 2 (year 2020; N=3756)	Test of sample differences
Paid employment	2335 (62.2)	2031 (54.9)	
Student	383 (10.2)	285 (7.7)	
Home duties or parenting	177 (4.7)	154 (4.2)	
Disability support	196 (5.2)	188 (5.1)	
Unemployed or seeking employment	318 (8.5)	289 (7.8)	
Retired	287 (7.6)	751 (20.3)	
Marital status, n (%)			$\chi^2_{3,7391}=119.4$, $P<.001$
Single or ever married	996 (26.5)	1125 (30.4)	
Widowed	55 (1.5)	145 (3.9)	
Separated or divorced	378 (10.1)	555 (15)	
Married or defacto	2267 (60.4)	1873 (50.6)	
Seeing mental health professional about anxiety or depression? n (%)			$\chi^2_{3,7394}=60.0$, $P<.001$
Never	747 (19.9)	999 (27)	
Previously	1737 (46.3)	1714 (46.3)	
Currently	1212 (32.3)	985 (26.6)	
Taking medication for anxiety or depression? n (%)			$\chi^2_{1,7451}=18.5$, $P<.001$
No	2356 (62.7)	2496 (67.5)	
Yes	1399 (37.3)	1202 (32.5)	

[a]PHQ-9: Patient Health Questionnaire 9-item.

[b]GAD-7: Generalized Anxiety Disorder 7 item.

[c]SWLS: Satisfaction With Life Scale.

Item-Level Analysis

The 96 items were then subjected to tests of the strength of association with each of the PHQ-9, GAD-7, and SWLS outcomes. The results from this first series of analyses are presented in Multimedia Appendix 2 for each item, with items grouped around the a priori clusters. The columns of the table report the mean PHQ-9, GAD-7, and SWLS scores observed with differing rates of weekly behavioral frequency. The PHQ-9, GAD-7, and SWLS scores were further incorporated into a heat map visualization to highlight the TYDQ items associated with the highest symptom scores among the 96 TYDQ items, where red hues indicate deterioration of psychological health (ie, symptom scores and satisfaction with life) and blue hues indicate improved psychological health.

The proportion of symptom variance explained (R^2) was reported as a summary measure of the association between each item and outcome. In Multimedia Appendix 2, the presentation order of the 96 items illustrates the relative ability of TYDQ items to account for the cumulative variance of the 3 outcomes, with the strongest items reported at the top of each a priori cluster. Together, the 96 items were significantly associated with each of the 3 outcomes (TYD outcome regression β test resulting in $P<.001$), except item TYDQ 57 (*I fulfilled my responsibilities even though I didn't want to*) and its association with GAD-7 outcomes. Most items showed a linear relationship between higher frequency and lower symptom scores, with a minority of items showing a flattened, convexed, or positive relationship with PHQ-9 and GAD-7 outcomes. However, the strength of the association between each of the items and each outcome varied substantially, both within and across clusters. For example, TYDQ 16, *I treated myself with respect*, illustrated the strongest association with depressive symptoms (accounting for 22.5% of all PHQ-9 scores); TYDQ 70, *I kept a realistic perspective on things*, was the strongest correlate of anxiety symptoms (accounting for 21.1% of all GAD-7 scores); and TYDQ 54, *I had something to look forward to*, was the strongest correlate of satisfaction with life (accounting for 25.9% of all SWLS scores). Across the items, the threshold of activity associated with optimal well-being varied from half a week to daily, with the majority of items requiring frequent (*almost every day*) but not necessarily daily activity.

Factor Formation and Analysis

EFAs were conducted to identify underlying groupings among the items and to explore a brief, latent factor structure of behaviors associated with psychological health. The results from the EFA of the complete 96-item list are presented in Multimedia Appendix 3. In total, 16 factors were identified, with 56 of the items demonstrating factor loadings above 0.5 for any 1 factor. Factors with loadings of ≤0.5 were suppressed. Across the TYDQ list, the factors appeared to form along their a priori clusters, with most items grouped within their clusters. No item-factor cross-loadings were identified above the loading of 0.5.

The results of the tests of association between the resulting factor scores and each of the PHQ-9, GAD-7, and SWLS outcomes are further reported in Multimedia Appendix 3, which describes the outcomes of a series of 4 EFAs, each testing a different number of items (see Multimedia Appendix 3 for items). This shows that the complete list of 16 factors, based on 56 items, accounted for 41%, 37%, and 32% of the outcome R^2, respectively. The resulting 16 factors can also be seen as varying in their strength of association with the 3 outcomes, with estimates of R^2 ranging from <0.1% to 24%.

The EFA was repeated with another subset of items (68 items), selected based on the individual item's ability to account for at least 5% of any of the 3 symptom outcomes (R^2>5%), which provided an opportunity to examine patterns among TYDQ items more closely related to psychological health. The analysis identified 11 factors, with only 45 items demonstrating factor loadings >0.5 for any one factor. The items were clustered within their respective a priori clusters, and no item-factor cross-loadings were identified. A test of the association between the resulting factor scores and each of the PHQ-9, GAD-7, and SWLS outcomes illustrated that a complete list of 11 factors based on 45 items accounted for 33%, 29%, and 32% of the outcome R^2, respectively, with each of the individual factors varying in their strength of association from 5% to 21% (Multimedia Appendix 4).

In a third EFA, an even briefer subset of items (35 items) was selected based on the item's ability to account for a minimum of 10% of any of the 3 symptom outcomes. This resulted in 5 factors and 27 items with a factor loading above 0.5 for any 1 factor. No item-factor cross-loadings were identified. A test of association between the resulting factor scores and each of the PHQ-9, GAD-7, and SWLS outcomes revealed a list of 5 factors based on 27 items, which accounted for 38%, 31%, and 33% of the outcomes R^2, respectively (Multimedia Appendix 4). The formation of factors largely followed the a priori clusters, and the resulting factors varied 3-fold in their strength of association with each of the outcomes (9%-28%).

A fourth EFA, which included items, identified a stringent R^2>15% criterion. This resulted in a brief set of items (8 items) found to represent a single composite factor combining key items from several domains. Although this result is interesting, the PHQ-9, GAD-7, and SWLS variance explained by a single factor (Multimedia Appendix 4) was suboptimal (R^2=36.6%, R^2=26.7%, and R^2=32.9%, respectively), and the representation of multiple domains, including activity planning, self-representation, and values through a single nonspecific factor, would inevitably limit the ability to assess and interpret the type of behavior.

Together, these EFAs illustrate that the 96-item list could be reduced to a more parsimonious and selective item list, with an optimal 5-factor, 27-item solution that retained >90% of the strength of association with each of the symptom outcomes, with only a quarter of the items (27/96, 28%). Each factor solution was further assessed for item reliability, mean item intercorrelation within each factor, and the ability of the identified factors to replicate each item (commonalities), as collated in Multimedia Appendix 5.

Generalizability of Factor Solutions Across Different Subgroups

A series of CFAs were conducted to examine the measurement invariance of the 5-factor solution as a brief but sensitive measure of psychological well-being across 5 categories, namely age groups, gender, depression and anxiety symptoms, education, and employment. The results from the CFAs are collated in Multimedia Appendix 6. The identification of TYDQ differences in the intercept and residual scores indicated that gender, age, and baseline symptoms were associated with different frequencies of weekly activity, although the TYDQ scores formed the same latent patterns across these groups. To check for the potential of alternate models, we reran the EFA within each of the age, sex, and PHQ-9 subgroups. The results identified the 5-factor solution reliably replicated as the most prominent latent solution when no parameter restraints were made or when the rotation methodology was changed (Multimedia Appendix 7). Together, the results from these CFAs illustrate that the latent 5-factor structure identified was statistically reliable, generalized across subgroups of clinical interest, and replicated across differing aspects of the statistical methodology (rotation and parameter constraint).

Discussion

Study 1 involved a comprehensive series of methodological and statistical steps designed to create a large list of actions associated with psychological health, test their strength of association with defined mental health outcomes, and develop a parsimonious list of items. To begin with, 96 items previously associated with psychological health were identified and administered to a large and diverse adult sample, along with measures of depression, anxiety, and satisfaction with life.

Several factor solutions were explored, and through a systematic process of dimension reduction, EFA solutions based on 56 items, 45 items, 27 items, and 8 items were examined. The 27-item solution, which comprised 5 factors, was found to be optimal, as it represented 38%, 31%, and 33% of the variance in scores of measures of depression, anxiety, and satisfaction with life, respectively. A series of CFAs indicated that the factor structure was robust across different subgroups, with these 5 factors interpreted as (1) realistic thinking, (2) meaningful activities, (3) goals and plans, (4) healthy habits, and (5) social connections.

Surprisingly, some items and factors that were expected to show strong relationships with psychological health, such as actions associated with kindness, gratitude, and spirituality, were not represented in the 5-factor solution. In contrast, items associated with skills commonly taught in psychological treatments, such as challenging unhelpful thoughts and engaging in pleasant activities, were very strongly associated with psychological health. Importantly, the results of study 1 not only identified items and factors that are strongly associated with psychological health but also found that the frequency of performing those actions was associated with improved well-being.

In summary, the results of study 1 provided preliminary psychometric evidence for the TYDQ as a measure of modifiable actions associated with psychological health. The reliability of the strong relationship between the 5-factor 27-item list found in study 1 needs to be replicated in further studies, which is one of the aims of study 2. Given the surprising finding that some items and factors were not strongly associated with psychological health, such items should also be retested.

Study 2

Overview

The objective of study 2 was to explore the reliability of the results of study 1 by replicating the design used in study 1 using a briefer version of the TYDQ. The aims were to (1) generate a shortened list of items based on the results of study 1 and explore the relationship between the weekly frequency of performing these items and outcomes of psychological health; (2) explore the underlying factor structures and rank the association between the items, factors, and psychological health; (3) explore the importance of different actions across different subgroups; and (4) develop a parsimonious list of items.

It should be noted that study 2 was conducted when major cities across Australia were locked down in response to the COVID-19 pandemic, which, on the one hand, represents a significantly changed context to that of study 1 but, on the other hand, provides an opportunity to test the items and the questionnaire in a psychologically challenging context.

Methods

Ethics Approval

The design of study 2 replicated that used in study 1. Ethics approval for data collection was obtained from the Macquarie University Human Research Ethics Committee (MQ HREC: 5201700988), and informed consent was obtained from all participants.

Questionnaire

Study 2 evaluated a version of the TYDQ comprising 59 of the 96 original items (Multimedia Appendix 8), with each item using the same 5-point rating scale used in study 1. This brief list was constructed in several steps. First, items that accounted for at least 5% of the PHQ-9, GAD-7, and SWLS were identified, resulting in an initial list of 67 items. Second, this list was edited to remove items with similar wording or duplicated items, resulting in the removal of 23 items. Third, 15 additional items from the original 96 item list were added to test the association between items, such as kindness to others and gratitude, which were considered clinically important and expected in study 1 to be associated with the outcomes but were not. We reviewed each of these steps.

Participants and Procedure

The procedure used in study 1 was repeated in study 2. As shown in Table 1, a total of 3756 participants consented to participate in the study and started filling the questionnaires, which were available between June and August 2020. These dates coincided with social and travel restrictions across Australia owing to the impact of the COVID-19 pandemic.

Other Measures

The outcome measures used in study 1 were also used in study 2.

Statistical Analyses

The three analytical steps in study 1 were repeated: (1) item-level analyses, (2) factor formation and analyses, and (3) analyses that explored the generalizability of factor solutions across different subgroups. Additional analyses were conducted to compare study 1 and study 2 samples on TYDQ scores, the strength of association between the items and symptom outcomes, and factors. For example, the comparison of TYDQ scores from the 2 studies was operationalized by treating the studies as a group variable and regressing this variable on the TYDQ item scores. Similarly, the association between TYDQ items and symptom outcomes was tested by regressing the TYDQ item score, group, and item-by-group interaction on outcome measures.

Results

Item-Level Analysis

A total of 3756 people consented to participate in study 2, and 3160 (84.13%) completed all the questionnaires. Participant characteristics and comparison with the study 1 sample are presented in Multimedia Appendix 8. These results indicate significant differences between the samples across most demographic variables, symptom severity, satisfaction with life, use of medications for anxiety or depression, and mental health service use.

The TYDQ score for each item, the association of each TYDQ item with the 3 symptom outcomes, and testing of differences between the 2 study samples are included in Multimedia Appendix 8. The results indicated a significant increase in the average score of most TYDQ items from study 1 to study 2. The results also indicate that the strength of association between most TYDQ items and each of the 3 outcomes was greater in study 2 than in study 1. However, the relative strength of association across and between the items and the outcomes was similar in magnitude across the samples, as seen in the heat map.

Factor Formation, Item Reduction, and Analysis

In the second step, the 59 items were subjected to a dimension-reduction analysis. The complete list of 59 items identified a 10-factor solution when no parameter constraints were imposed on the data, with 9 of these factors overlapping with the study 1 EFA solution (Multimedia Appendix 9). Using the inclusion criterion of R^2>10%, a 26-item solution was identified using the data from study 2, replicating the pattern of the 5 factors identified in study 1. Using the inclusion criteria of R^2>10% and reanalyzing the results of study 1 resulted in an even more parsimonious 21-item solution (Multimedia Appendix 9) that reliably retained the factor structure. The 5 factors and the 21-item solution identified in the reanalysis of study 1 results (Multimedia Appendix 9; inclusion criteria of R^2>10%) were

then compared with the same 21 items from the study 2 results, and associations between the 5 factors and the 3 outcome variables were compared across samples (Multimedia Appendix 10). This demonstrated close similarities in strength, directionality, significance, and optimal frequency cutoffs between the 2 samples.

A set of 3 EFAs were then conducted for item subsets with no selection criteria, inclusion criteria of R^2>10%, and inclusion criteria of R^2>15%. The 5-factor 21-item solution associated with the R^2>10% criteria in study 2 accounted for the highest amount of PHQ-9, GAD-7, and SWLS outcome variance (R^2=46.8%, R^2=38.2%, and R^2=38.1%, respectively). Consistent with this, a reanalysis of the data from study 1 using the same R^2>10% criteria demonstrated that this 5-factor 21-item solution accounted for the highest amount of PHQ-9, GAD-7, and SWLS outcome variance (R^2=37.1%, R^2=30.0%, and R^2=30.5%, respectively).

Generalizability of Factor Solutions Across Different Subgroups

CFAs seeking to verify a 5-factor 21-item solution across the study 1 and study 2 samples resulted in metric invariance (similarities in factors and item loading) but not scalar or strict invariance (Multimedia Appendix 11). This result is consistent with the interpretation that participants from the 2 studies differed in their means but not item or factor importance.

In brief, the overall results of study 2 replicated the results and 5-factor model obtained in study 1, albeit with a 21- rather than a 27-item solution. These results confirm the strength of the relationship between key items in the TYDQ and psychological health outcomes and the underlying factor structure.

Discussion

Study 2 examined the performance of a shortened version of the TYDQ. The survey was completed by 3160 people, and although the sample was significantly different from that of study 1, the overall patterns of results obtained in study 1 were replicated. The 5-factor structure observed in study 1 was supported and accounted for 46.8%, 38.2%, and 38.1% of the variance in the symptoms of depression, anxiety, and satisfaction with life, respectively. These solutions were achieved using a parsimonious 21-item list. Importantly, and again consistent with the results of study 1, the observations that some items and factors expected to show strong relationships with psychological health, such as showing kindness to others and expressing gratitude, were not found to be associated with well-being. The implications and limitations of these results are discussed in the subsequent section.

General Discussion

This paper describes 2 studies reporting on the initial development and evaluation of the TYDQ, which aims to measure modifiable actions that are strongly associated with psychological health. The detailed and systematic approach followed accepted scale development methods [32] and attempted to address the limitations of previous work by including a broad range of modifiable actions, exploring their frequency over a 1-week timeframe, identifying the actions most strongly related to the target outcomes, and testing the generalizability of results across age, gender, and baseline symptom severity.

Principal Findings

Study 1 examined the performance of a 96-item version of TYDQ. A 27-item 5-factor solution achieved an optimal balance between accounting for sufficient variance in outcome measures of depression, anxiety, and satisfaction with life and parsimony. These five factors were interpreted as actions concerned with (1) realistic thinking, (2) meaningful activities, (3) goals and plans, (4) healthy habits, and (5) social connections. There was a strong relationship between the frequency of performing the actions associated with these factors each week and improved psychological health, with results indicating that they should be performed at least half a week or more to optimize their psychological health. Importantly, the factor structure was consistent across demographic variables, including gender and age groups.

Overall, these findings were replicated in study 2, which evaluated a shortened 59-item version by using the same systematic analytic strategy. Study 2 was conducted during a period when participants were mostly in lockdown owing to the impact of the COVID-19 pandemic in Australia. However, the strength of the relationship between the frequency of target actions and outcomes increased relative to study 1. The same 5-factor solution found in study 1 was also found in study 2, although slight differences in the item loadings were observed across the 2 samples. An item-reduction exercise revealed that a 5-factor 21-item version solution achieved an optimal balance between item parsimony and accounting for sufficient variance across the 3 outcome measures. These results appear to confirm the robustness of the 5 factors but also suggest that the 21-item list (Multimedia Appendix 9) should not be considered final, and additional survey design efforts could further improve the brevity and validity of the items identified. In any case, the 96-item and 59-item lists are provided in Multimedia Appendix 11.

Comparison With Prior Work

The overall results are consistent with previous research regarding key items associated with psychological health, including cognitions and actions such as, *I did something enjoyable, I kept a realistic perspective on things, I had something to look forward to, and I treated myself with respect*. These 5 factors confirm the importance of many of the skills taught in psychological interventions, including cognitive therapy [10], behavior therapy [12,49], and others. Various combinations of these skills and actions have been recognized as contributing to psychological health [50-52] and are frequently included in public mental health campaigns. Thus, the contribution of our preliminary study is not to identify the actions and factors associated with psychological health. Instead, these studies extend the literature by comparing the relative benefits of different types of actions on psychological health and the frequency with which these factors are associated with psychological health. The findings indicate that performing

these actions for at least half the days of the week is important for psychological health across age, gender and demographic groups, as well as in different severity levels of mood symptoms. This proposed frequency needs to be carefully evaluated using a longitudinal study; however, if supported, it would indicate a key target for psychological interventions as well as for public health interventions designed to promote psychological health and prevent common mood disorders in the community.

An additional and important finding was that some actions were less strongly associated with psychological health than expected. For example, acts of kindness, practicing gratitude, and spirituality are widely considered important for psychological health; however, in this study, these actions were not found to be as important as those relevant to the 5 factors. This lack of findings is broadly consistent with the results of empirical studies [53-59].

For example, in a systematic review and meta-analysis of 27 experimental studies examining the relationships between acts of kindness and different measures of subjective well-being, including measures of psychological health, Curry et al [53] found the overall effect of kindness to be small to medium (δ=0.28); however, they noted that most of the reviewed studies were underpowered to detect small differences. In a series of meta-analyses examining the findings of 38 studies, Dickens [54] showed that when compared with waitlist, inactive, or no-treatment control groups, gratitude interventions were associated with decreased depressive symptoms and improved positive affect, well-being, happiness, and life satisfaction, albeit with small effect sizes (Cohen *d*=0.13−0.30). In a more recent review of the effect of gratitude traits and interventions, Jans-Beken et al [60] concluded that gratitude interventions were not found to reliably improve symptoms, such as depression and anxiety, but they were associated with improved happiness and life satisfaction. In a meta-analysis of 48 longitudinal studies, Garssen et al [56] reported an overall positive effect of religion or spirituality on mental health; however, similar to studies examining the relationships between mental health and acts of kindness and practicing gratitude, the effects were small, with a random weighted average effect size of *r*=0.08.

Given the large number and broad cross-section of people completing the 2 surveys, these findings appear to be robust. However, given the preliminary nature of the 2 studies reported here, we propose that the findings should not be interpreted as such actions are not helpful; rather, it is likely that different actions are important at different stages of psychological health and for particular individuals and age groups. For example, it is commonly recognized in clinical practice that people who are severely depressed are more likely to benefit from actions involving increasing reinforcement and pleasure than from practicing gratitude, but once their mood has lifted, actions such as practicing gratitude and showing kindness to others may help maintain psychological health. In addition, it should be noted that the results obtained here are relevant to the outcomes of interest, namely, symptoms of depression, anxiety, and satisfaction with life, and that choosing other outcome measures may result in other factors and items ascending in importance. This raises the possibility of future research exploring the relationship between everyday actions and a much broader range of mental health outcomes, including the general p factor, which extends beyond internalizing disorders to include externalizing and thought or psychotic disorders [61].

Limitations

The preliminary and cross-sectional nature of the 2 studies reported here are associated with several limitations. First, despite the large sample sizes and consistent results across subsamples, the work should be considered preliminary and tested in other samples, including people with different cultural and demographic backgrounds. In addition, the items and factors identified here need to be tested using longitudinal or interventional research designs to gauge their sensitivity to change and reliability over time. The robustness of the 5 factors should also be tested across differing statistical methods for identifying dimensionality, for example, using different statistical criteria for selecting items or using latent profiling.

We also acknowledge that the initial selection of items used to develop the TYDQ, despite the breadth of our inquiries, involved some subjective choices. We have shared lists to assist independent replication and further development (Multimedia Appendix 12). Related to these limitations, we also acknowledge that the terms we generated to describe the factors were based on our clinical and psychometric decisions, and that other research groups may have generated other labels for each of the factors. We also acknowledge that although most of the actions were relatively simple to understand by participants, some actions were more difficult for participants to measure; for example, the item *I treated myself with respect*. This raises important questions about how people interpreted some of the items, which may in turn affect both the reliability and validity of the results. This is further complicated by the observation that a person's interpretation of the questions may change over time and with their experience of the action. Finally, we acknowledge that the focus of this study is primarily on the relationship between actions and symptoms and may not be generalizable to other concepts of psychological well-being, happiness, or thriving [62].

Strengths

The strengths of this study include the development of a questionnaire containing items that were not obviously associated with any specific psychological model or approach, evaluated through a systematic and detailed analytic procedure that included multiple tests of generalizability by using 2 large sample sizes. This procedure contributes to the literature by comparing the relative strengths of several groups of modifiable actions and the minimum frequency required for such actions to affect psychological health. An additional strength is that the results were mostly replicated although the second sample was obtained during a challenging period for the community characterized by social restrictions related to the impacts of the COVID-19 pandemic.

Conclusions

In conclusion, these studies provide preliminary evidence for actions and factors strongly associated with psychological health and the frequency with which they should be performed. Future

studies are planned to further explore the patterns of change in brief versions of the TYDQ, including samples of people seeking treatment for mental health conditions. We hope that these efforts will assist in the development of new psychological interventions and provide an evidence base for public mental health campaigns designed to promote good mental health and prevent the emergence of common mental disorders.

Acknowledgments

The studies reported in this paper were partly funded by donations from the Mental Wheels Foundation, which had no role in study design or analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Items, primary and secondary cluster, item wording, and weekly score dispersion by frequency.
[DOCX File , 37 KB - formative_v6i7e38837_app1.docx]

Multimedia Appendix 2
A heat map of the strength of association of Things You Do Questionnaire item weekly frequency with Patient Health Questionnaire 9-item, Generalized Anxiety Disorder 7-item, and Satisfaction With Life Scale outcomes, and the behavioral frequency thresholds associated with optimal mental health gains.
[DOCX File , 111 KB - formative_v6i7e38837_app2.docx]

Multimedia Appendix 3
Exploratory factor analysis solution and assigned factor labels for items grouped with no item selection criteria (96 items), and item grouping under R2>5%, 10%, and 15% outcomes association criteria.
[DOCX File , 46 KB - formative_v6i7e38837_app3.docx]

Multimedia Appendix 4
A heat map of the identified factor association with Patient Health Questionnaire 9-item, Generalized Anxiety Disorder 7-item, and Satisfaction With Life Scale outcomes, and the behavioral frequency thresholds associated with optimal mental health gains.
[DOCX File , 73 KB - formative_v6i7e38837_app4.docx]

Multimedia Appendix 5
Exploratory factor analysis diagnostic statistics: item commonalities, factor eigenvalues, factor Cronbach α, and item mean intercorrelations.
[DOCX File , 70 KB - formative_v6i7e38837_app5.docx]

Multimedia Appendix 6
Invariance test solutions examining dimensionality in study 1 sample.
[DOCX File , 56 KB - formative_v6i7e38837_app6.docx]

Multimedia Appendix 7
Exploratory factor analysis parameter solution replications across key sample dimensions and methods of rotation methods.
[DOCX File , 62 KB - formative_v6i7e38837_app7.docx]

Multimedia Appendix 8
Means, study 1 and study 2 sample group differences, and association of Things You Do Questionnaire items with each of the Patient Health Questionnaire 9-item, Generalized Anxiety Disorder 7-item, and Satisfaction With Life Scale outcomes.
[DOCX File , 47 KB - formative_v6i7e38837_app8.docx]

Multimedia Appendix 9
Exploratory factor analysis solution and assigned factor labels for items grouped with no item selection criteria (96 items), and under R2>10% outcomes association criteria.
[DOCX File , 40 KB - formative_v6i7e38837_app9.docx]

Multimedia Appendix 10
A heat map of the identified factor association with Patient Health Questionnaire 9-item, Generalized Anxiety Disorder 7-item, and Satisfaction With Life Scale outcomes across study 1 and study 2 samples.
[DOCX File , 70 KB - formative_v6i7e38837_app10.docx]

Multimedia Appendix 11
Invariance test solutions examining dimensionality in study 2 sample.
[DOCX File , 55 KB - formative_v6i7e38837_app11.docx]

Multimedia Appendix 12
Things You Do Questionnaire item lists.
[DOCX File , 41 KB - formative_v6i7e38837_app12.docx]

References

1. Ciobanu LG, Ferrari AJ, Erskine HE, Santomauro DF, Charlson FJ, Leung J, et al. The prevalence and burden of mental and substance use disorders in Australia: findings from the Global Burden of Disease Study 2015. Aust N Z J Psychiatry 2018 May 11;52(5):483-490. [doi: 10.1177/0004867417751641] [Medline: 29325437]
2. GBD 2016 DiseaseInjury IncidencePrevalence 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]
3. Cuijpers P, Smit F. Subthreshold depression as a risk indicator for major depressive disorder: a systematic review of prospective studies. Acta Psychiatr Scand 2004 May;109(5):325-331. [doi: 10.1111/j.1600-0447.2004.00301.x] [Medline: 15049768]
4. Cuijpers P, Schuurmans J. Self-help interventions for anxiety disorders: an overview. Curr Psychiatry Rep 2007 Aug 11;9(4):284-290. [doi: 10.1007/s11920-007-0034-6] [Medline: 17880859]
5. Diagnostic And Statistical Manual Of Mental Disorders, Fifth Edition. Virginia, United States: American Psychiatric Association; 2013.
6. International statistical classification of diseases and related health problems. In: Encyclopedia of Clinical Neuropsychology. New York, United States: Springer; 2011.
7. Holmes EA, Ghaderi A, Harmer CJ, Ramchandani PG, Cuijpers P, Morrison AP, et al. The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. Lancet Psychiatry 2018 Mar;5(3):237-286. [doi: 10.1016/s2215-0366(17)30513-8]
8. Kazdin AE. Understanding how and why psychotherapy leads to change. Psychother Res 2009 Jul 22;19(4-5):418-428. [doi: 10.1080/10503300802448899] [Medline: 19034715]
9. Southward MW, Sauer-Zavala S. Experimental manipulations to test theory-driven mechanisms of cognitive behavior therapy. Front Psychiatry 2020 Dec 17;11:603009 [FREE Full text] [doi: 10.3389/fpsyt.2020.603009] [Medline: 33391056]
10. Beck AT, Dozois DJ. Cognitive therapy: current status and future directions. Annu Rev Med 2011 Feb 18;62(1):397-409. [doi: 10.1146/annurev-med-052209-100032] [Medline: 20690827]
11. Dimidjian S, Barrera M, Martell C, Muñoz RF, Lewinsohn PM. The origins and current status of behavioral activation treatments for depression. Annu Rev Clin Psychol 2011 Apr 27;7(1):1-38. [doi: 10.1146/annurev-clinpsy-032210-104535] [Medline: 21275642]
12. Spiegler M. Contemporary Behavior Therapy. Boston, Massachusetts, United States: Cengage Learning; 2015.
13. Gu J, Strauss C, Bond R, Cavanagh K. How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. Clin Psychol Rev 2015 Apr;37:1-12. [doi: 10.1016/j.cpr.2015.01.006] [Medline: 25689576]
14. Lipsitz JD, Markowitz JC. Mechanisms of change in interpersonal therapy (IPT). Clin Psychol Rev 2013 Dec;33(8):1134-1147 [FREE Full text] [doi: 10.1016/j.cpr.2013.09.002] [Medline: 24100081]
15. Hundt NE, Mignogna J, Underhill C, Cully JA. The relationship between use of CBT skills and depression treatment outcome: a theoretical and methodological review of the literature. Behav Ther 2013 Mar;44(1):12-26. [doi: 10.1016/j.beth.2012.10.001] [Medline: 23312423]
16. Jarrett RB, Vittengl JR, Clark LA, Thase ME. Skills of Cognitive Therapy (SoCT): a new measure of patients' comprehension and use. Psychol Assess 2011 Sep;23(3):578-586 [FREE Full text] [doi: 10.1037/a0022485] [Medline: 21319902]
17. Jacob KL, Christopher MS, Neuhaus EC. Development and validation of the cognitive-behavioral therapy skills questionnaire. Behav Modif 2011 Nov 05;35(6):595-618. [doi: 10.1177/0145445511419254] [Medline: 21893554]
18. Manos RC, Kanter JW, Luo W. The behavioral activation for depression scale-short form: development and validation. Behav Ther 2011 Dec;42(4):726-739. [doi: 10.1016/j.beth.2011.04.004] [Medline: 22036000]

19. Strunk DR, Hollars SN, Adler AD, Goldstein LA, Braun JD. Assessing patients' cognitive therapy skills: initial evaluation of the competencies of cognitive therapy scale. Cognit Ther Res 2014 Oct;38(5):559-569. [doi: 10.1007/s10608-014-9617-9] [Medline: 25408560]
20. Terides MD, Dear BF, Karin E, Jones MP, Gandy M, Fogliati VJ, et al. The frequency of actions and thoughts scale: development and psychometric validation of a measure of adaptive behaviours and cognitions. Cogn Behav Ther 2016 Apr 29;45(3):196-216. [doi: 10.1080/16506073.2016.1149876] [Medline: 26926484]
21. Terides M, Dear B, Fogliati V, Gandy M, Karin E, Jones MP, et al. Increased skills usage statistically mediates symptom reduction in self-guided internet-delivered cognitive-behavioural therapy for depression and anxiety: a randomised controlled trial. Cogn Behav Ther 2018 Jan;47(1):43-61. [doi: 10.1080/16506073.2017.1347195] [Medline: 28724338]
22. Jorm AF, Griffiths KM, Christensen H, Parslow RA, Rogers B. Actions taken to cope with depression at different levels of severity: a community survey. Psychol Med 2004 Feb 28;34(2):293-299. [doi: 10.1017/s003329170300895x] [Medline: 14982135]
23. Morgan AJ, Chittleborough P, Jorm AF. Self-help strategies for sub-threshold anxiety: a Delphi consensus study to find messages suitable for population-wide promotion. J Affect Disord 2016 Dec;206:68-76 [FREE Full text] [doi: 10.1016/j.jad.2016.07.024] [Medline: 27466744]
24. Pilkington K, Wieland LS. Self-care for anxiety and depression: a comparison of evidence from Cochrane reviews and practice to inform decision-making and priority-setting. BMC Complement Med Ther 2020 Aug 10;20(1):247 [FREE Full text] [doi: 10.1186/s12906-020-03038-8] [Medline: 32778171]
25. Shepardson RL, Tapio J, Funderburk JS. Self-management strategies for stress and anxiety used by nontreatment seeking veteran primary care patients. Military Med 2017 Jul;182(7):e1747-e1754. [doi: 10.7205/milmed-d-16-00378]
26. van Grieken RA, van Tricht MJ, Koeter MW, van den Brink W, Schene AH. Correction: the use and helpfulness of self-management strategies for depression: the experiences of patients. PLoS One 2018 Dec 13;13(12):e0209109 [FREE Full text] [doi: 10.1371/journal.pone.0209109] [Medline: 30543714]
27. Morgan AJ, Jorm AF, Mackinnon AJ. Self-help for depression via e-mail: a randomised controlled trial of effects on depression and self-help behaviour. PLoS One 2013 Jun 21;8(6):e66537 [FREE Full text] [doi: 10.1371/journal.pone.0066537] [Medline: 23805231]
28. Frisch M. QOLI Quality of Life Inventory: Manual and Treatment Guide. Minneapolis, MN: Pearson; 1994.
29. THE WHOQOL GROUP. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group. Psychol Med 1998 May 01;28(3):551-558. [doi: 10.1017/s0033291798006667] [Medline: 9626712]
30. Lindner P, Frykheden O, Forsström D, Andersson E, Ljótsson B, Hedman E, et al. The Brunnsviken Brief Quality of Life Scale (BBQ): development and psychometric evaluation. Cogn Behav Ther 2016 Apr 17;45(3):182-195 [FREE Full text] [doi: 10.1080/16506073.2016.1143526] [Medline: 26886248]
31. Mental health: strengthening our response. World Health Organisation. 2018. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response [accessed 2022-05-04]
32. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health 2018 Jun 11;6:149 [FREE Full text] [doi: 10.3389/fpubh.2018.00149] [Medline: 29942800]
33. 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]
34. Kroenke K, Spitzer RL, Williams JB, Löwe B. The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry 2010 Jul;32(4):345-359. [doi: 10.1016/j.genhosppsych.2010.03.006] [Medline: 20633738]
35. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. CMAJ 2012 Feb 21;184(3):E191-E196 [FREE Full text] [doi: 10.1503/cmaj.110829] [Medline: 22184363]
36. Titov N, Andersson G. Using brief measures to identify depression and other mental disorders: a challenge for research and clinical practice. Internet Interv 2022 Apr;28:100450 [FREE Full text] [doi: 10.1016/j.invent.2021.100450] [Medline: 35646605]
37. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Med Care 2008 Mar;46(3):266-274. [doi: 10.1097/MLR.0b013e318160d093] [Medline: 18388841]
38. 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]
39. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. J Pers Assess 1985 Feb 10;49(1):71-75. [doi: 10.1207/s15327752jpa4901_13] [Medline: 16367493]
40. Pavot W, Diener E, Colvin CR, Sandvik E. Further validation of the satisfaction with life scale: evidence for the cross-method convergence of well-being measures. J Pers Assess 1991 Aug;57(1):149-161. [doi: 10.1207/s15327752jpa5701_17] [Medline: 1920028]

41. Pavot W, Diener E. The Satisfaction With Life Scale and the emerging construct of life satisfaction. J Positive Psychol 2008 Apr;3(2):137-152. [doi: 10.1080/17439760701756946]
42. Haarman BC, Riemersma-Van der Lek RF, Nolen WA, Mendes R, Drexhage HA, Burger H. Feature-expression heat maps--a new visual method to explore complex associations between two variable sets. J Biomed Inform 2015 Feb;53:156-161 [FREE Full text] [doi: 10.1016/j.jbi.2014.10.003] [Medline: 25445923]
43. Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. Dev Rev 2016 Sep;41:71-90 [FREE Full text] [doi: 10.1016/j.dr.2016.06.004] [Medline: 27942093]
44. The R project for statistical computing. R Foundation. URL: https://www.r-project.org [accessed 2022-02-01]
45. Rosseel Y. lavaan: an R package for structural equation modeling. J Stat Softw 2012;48(2):36. [doi: 10.18637/jss.v048.i02]
46. Howard MC. A review of exploratory factor analysis decisions and overview of current practices: what we are doing and how can we improve? Int J Human Comput Interact 2015 Sep 14;32(1):51-62. [doi: 10.1080/10447318.2015.1087664]
47. Tabachnick B, Fidell L, Ullman J. Using Multivariate Statistics, 5th edition. New York: Pearson Education; 2007.
48. Best practices in exploratory factor analysis. In: Best Practices in Quantitative Methods. Thousand Oaks, California: SAGE Publications; 2008.
49. Lewinsohn P. A behavioral approach to depression. In: The Psychology of Depression: Contemporary Theory and Research. Hoboken, New Jersey, United States: Wiley; 1974.
50. Keyes CL. The mental health continuum: from languishing to flourishing in life. J Health Social Behav 2002 Jun;43(2):207. [doi: 10.2307/3090197]
51. Ryff CD. Happiness is everything, or is it? Explorations on the meaning of psychological well-being. J Personality Social Psychol 1989;57(6):1069-1081. [doi: 10.1037/0022-3514.57.6.1069]
52. Seligman M. PERMA and the building blocks of well-being. J Positive Psychol 2018 Feb 16;13(4):333-335. [doi: 10.1080/17439760.2018.1437466]
53. Curry O, Rowland L, Van LC, Zlotowitz S, McAlaney J, Whitehouse H. Happy to help? A systematic review and meta-analysis of the effects of performing acts of kindness on the well-being of the actor. J Exp Soc Psychol 2018;76:9. [doi: 10.31219/osf.io/ytj5s]
54. Dickens LR. Using gratitude to promote positive change: a series of meta-analyses investigating the effectiveness of gratitude interventions. Basic Applied Social Psychol 2017 May 30;39(4):193-208. [doi: 10.1080/01973533.2017.1323638]
55. Emmons RA, McCullough ME. Counting blessings versus burdens: an experimental investigation of gratitude and subjective well-being in daily life. J Personality Social Psychol 2003;84(2):377-389. [doi: 10.1037/0022-3514.84.2.377]
56. Garssen B, Visser A, Pool G. Does spirituality or religion positively affect mental health? Meta-analysis of longitudinal studies. Int J Psychol Religion 2020 Feb 27;31(1):4-20. [doi: 10.1080/10508619.2020.1729570]
57. Jans-Beken L, Jacobs N, Janssens M, Peeters S, Reijnders J, Lechner L, et al. Gratitude and health: an updated review. J Positive Psychol 2019 Aug 06;15(6):743-782. [doi: 10.1080/17439760.2019.1651888]
58. Peteet J. Spirituality and mental health: implications for ethics, medicine, and public health. Ethics Med Public Health 2019 Apr;9:75-79. [doi: 10.1016/j.jemep.2019.05.002]
59. Vitorino LM, Lucchetti G, Leão FC, Vallada H, Peres MFP. The association between spirituality and religiousness and mental health. Sci Rep 2018 Nov 22;8(1):17233 [FREE Full text] [doi: 10.1038/s41598-018-35380-w] [Medline: 30467362]
60. Jans-Beken L, Jacobs N, Janssens M, Peeters S, Reijnders J, Lechner L, et al. Gratitude and health: An updated review. The Journal of Positive Psychology 2019 Aug 06;15(6):743-782. [doi: 10.1080/17439760.2019.1651888]
61. Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S, et al. The p factor: one general psychopathology factor in the structure of psychiatric disorders? Clin Psychol Sci 2014 Mar 14;2(2):119-137 [FREE Full text] [doi: 10.1177/2167702613497473] [Medline: 25360393]
62. van Agteren J, Iasiello M, Lo L, Bartholomaeus J, Kopsaftis Z, Carey M, et al. A systematic review and meta-analysis of psychological interventions to improve mental wellbeing. Nat Hum Behav 2021 May 19;5(5):631-652. [doi: 10.1038/s41562-021-01093-w] [Medline: 33875837]

Abbreviations

CFA: confirmatory factor analysis
EFA: exploratory factor analysis
GAD-7: Generalized Anxiety Disorder 7-item
PHQ-9: Patient Health Questionnaire 9-item
SWLS: Satisfaction With Life Scale
TYDQ: Things You Do Questionnaire

Edited by A Mavragani; submitted 18.04.22; peer-reviewed by G Andersson, V Kaldo; comments to author 17.05.22; revised version received 31.05.22; accepted 01.06.22; published 05.07.22.

Please cite as:
Titov N, Dear BF, Bisby MA, Nielssen O, Staples LG, Kayrouz R, Cross S, Karin E
Measures of Daily Activities Associated With Mental Health (Things You Do Questionnaire): Development of a Preliminary Psychometric Study and Replication Study
JMIR Form Res 2022;6(7):e38837
URL: https://formative.jmir.org/2022/7/e38837
doi:10.2196/38837
PMID:35788101

Original Paper

Standardizing Primary Health Care Referral Data Sets in Nigeria: Practitioners' Survey, Form Reviews, and Profiling of Fast Healthcare Interoperability Resources (FHIR)

Emeka Chukwu[1,2], BEng, MSc; Lalit Garg[1], BSc, PhD; Nkiruka Obande-Ogbuinya[3], BSc, MSc, PhD; Vijay Kumar Chattu[4,5,6], MD, MPH, MPhil

[1]Department of Computer Information Systems, Faculty of Information Communication Technology, University of Malta, Msida, Malta
[2]Digital Health Interoperability Network, Abuja, Nigeria
[3]Department of Science Education, Alex-Ekwueme Federal University Ndufu-Alike, Ikwo, Ebonyi, Nigeria
[4]Department of Occupational Science & Occupational Therapy, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[5]Center for Transdisciplinary Research, Saveetha Institute of Medical and Technological Sciences, Saveetha University, Chennai, India
[6]Department of Community Medicine, Faculty of Medicine, Datta Meghe Institute of Medical Sciences, Wardha, India

Corresponding Author:
Emeka Chukwu, BEng, MSc
Department of Computer Information Systems
Faculty of Information Communication Technology
University of Malta
PG room A24, Level 0
Msida, MSD 2080
Malta
Phone: 356 99330888
Email: nnaemeka_ec@hotmail.com

Abstract

Background: Referral linkages are crucial for efficient functioning of primary health care (PHC) systems. Fast Healthcare Interoperability Resource (FHIR) is an open global standard that facilitates structuring of health information for coordinated exchange among stakeholders.

Objective: The objective of this study is to design FHIR profiles and present methodology and the profiled FHIR resource for Maternal and Child Health referral use cases in Ebonyi state, Nigeria—a typical low- and middle-income country (LMIC) setting.

Methods: Practicing doctors, midwives, and nurses were purposefully sampled and surveyed. Different referral forms were reviewed. The union of data sets from surveys and forms was aggregated and mapped to base patient FHIR resource elements, and extensions were created for data sets not in the core FHIR specification. This study also introduced FHIR and its relation to the World Health Organization's (WHO's) International Classification of Diseases.

Results: We found many different data elements from the referral forms and survey responses even in urban settings. The resulting FHIR standard profile is published on GitHub for adaptation or adoption as necessary to aid alignment with WHO recommendations. Understanding data sets used in health care and clinical practice for information sharing is crucial in properly standardizing information sharing, particularly during the management of COVID-19 and other infectious diseases. Development organizations and governments can use this methodology and profile to fast-track FHIR standards adoption for paper and electronic information sharing at PHC systems in LMICs.

Conclusions: We presented our methodology for profiling the referral resource crucial for the standardized exchange of new and expectant moms' information. Using data from frontline providers and mapping to the FHIR profile helped contextualize the standardized profile.

(JMIR Form Res 2022;6(7):e28510) doi:10.2196/28510

KEYWORDS

FHIR; COVID-19; digital health; eHealth; mHealth; BlockMom; Nigeria; primary health care; health information; health information exchange; interoperability

Introduction

Background

Health care is a complex sector that involves medical professionals, allied health workers, the information and communication technology (ICT) workforce, and various other stakeholders. The World Health Organization (WHO) highlights the importance of 6 building blocks of any health system: service delivery, health workforce, health information systems, access to essential medicines, financing, and leadership or governance [1]. Therefore, the health information systems block is very critical and plays an important role in data capture, processing, and usage. Substantial investments have been made in the development and strengthening of routine health information systems (RHIS) in many low- and middle-income countries (LMICs) over the past 2 decades [2,3]. Although early RHISs were produced using paper-based health facility reports, many LMICs have implemented newer web-based systems over the past decade [4,5]. Given that substantial investments have been made in strengthening RHISs in LMICs in recent years, researchers have a growing demand for more real-time data [6]. Besides, data for policy and operational decision-making in LMICs, including Nigeria, have been largely limited to report generation, process monitoring, and early surveillance responses. Reliable, quality, timely, and transparent health service data are essential for an efficient health system [7]. Globally, health care interoperability has been identified as vital to seamless care coordination among the different stakeholders.

Global Health Care Interoperability

According to the World Health Organization (WHO) Europe's 2016 e-Health in practice report, Estonia is the first country to implement electronic health records (EHRs) [8]. The famous X-Road facilitates Estonia's exchange network, an interoperability layer launched in 2001, with several different services added over the years. Estonia achieved success with over 99% of electronic medical subscriptions in 2018. Estonia's X-Road interoperability layer connects over 2700 services across 700 institutions and enterprises across several sectors, including health care. The United Kingdom's national service specification was based on the Health Level 7 (HL7) version 3 standard and is now transitioning to the HL7 Fast Healthcare Interoperability Resource (FHIR). However, local implementation was left to providers to determine, most of whom already run different versions of HL7 version 2 [9]. This National Health Service project started with 2 main use cases: the Summary Care Records and the Detailed Care Records [9]. Canada launched a national Infoway project to standardize and foster collaboration among pan-Canadian health care solutions [10]. After leveraging the CEN TC251 standards for referrals, discharge letters, laboratory, prescriptions, reimbursements, radiology requests, and reports, a national program was deemed successful in Denmark. These use cases were pilot-tested via 15 independently managed projects [10]. In 2008, a report highlighted that in the United States, "only 15 to 20 percent of medical doctors have access to computerized patient records and only a small fraction of the billions of medical transactions happen electronically" [10]. Such a low usage led to the creation of Health Information Technology for Economic and Clinical Health, which was launched in 2009 to incentivize digitization, and things have since changed.

Interoperability in LMICs

Some LMIC health systems services are still paper-dependent for recording and transmitting health information. Paper records are limited because only one person can access them at a time. Systematic digitization of health systems has driven the development and implementation of national digital health strategies in Nigeria and other LMICs [11,12]. From our literature search [13] and to our knowledge, LMICs still struggle with patient-level interoperability project implementations, which has limited recording successes. Nigeria is a typical LMIC because it has one of the highest global burdens of maternal mortality [14]. Furthermore, there are more primary health care (PHC) facilities (approximately 10 times) than hospitals in Nigeria; hence, here we focus on PHCs. PHCs have the highest potential for impact in the Nigerian health system because most health services are delivered at the PHC level. In a typical PHC network, the possible use cases for health information interchange may include the following:

- Interdepartmental care communication
- Inter-PHC or PHC to secondary hospital referral
- Reporting of decentralized laboratory results
- Triangulation of immunization and surveillance information
- Payment settlement
- Diagnostic information exchange

Nigeria used the DHIS2 for routine reporting of the delivery of health information system services. Routine health information systems (RHISs) continue to collect data on a wide range of diseases and conditions [6]. These RHIS data are analyzed to assess community-level initiatives such as policies to boost community engagement and strengthen referrals from traditional birth attendants to increase demand for maternal and child care [15-17]. The COVID-19 pandemic has further exposed the weakness in health systems worldwide and the value of linkages.

Health Care Interoperability Standards

The international organizations for certifying and ratifying widely used digital health standards are the ISO/TC (International Organization for Standards' Health Informatics Technical Committee) 215 and CEN/TC (European Committee for Standards' Health ICT Technical committee) 251. For instance, ISO 21090:2011 is a ratified HL7 version 3 data type for information interchange. Similarly, ISO 13606-1:2019 is a ratified description of archetype reference models. HL7 is a leading health care standard development organization that has facilitated many standards, including the HL7 version 2 messaging standard, HL7 version 3 Clinical Document Architecture document exchange standard, and the HL7 FHIR. FHIR was popularized because it supports REpresentational State Transfer (REST)–based web-based (real-time) transactions and its extension for services. FHIR is now emerging as the de facto global standard for health care data interchange. The FHIR community includes Microsoft, Google, Apple, and many electronic medical record and EHR vendors [18-20]. In addition, the WHO has recently published a digital adaptation kit to support countries deploying standards for antenatal care [21].

Terminologies

In addition to data interchange standards, terminology categorization helps guarantee consistent and uniform understanding (and meaning) of terms in health care systems (within and across geographies). The leading terminologies for disease, procedure, and other concept classification are Systematized Nomenclature of Medicine–Clinical Terms (SNOMED-CT) and the International Classification of Diseases (ICD). Other technology providers are Logical Observation Identifiers Names and Codes (LOINC) for laboratory result reporting and Digital Imaging and Communication in Medicine (DICOM) for imaging data reporting. This study used ICD because it uses a free license compared to the better developed SNOMED-CT for disease classification. Code systems such as the WHO ICD-10 use statistical classification of medical concepts and entities into coded groups, assigning identifiers [22]. Codes allow for the unique identification of these concepts in an information processing system. These codes classify diseases, procedures, billing, history or symptoms, and case summaries (jurisdictional and international aggregate reporting). Simultaneously, service providers, including clinicians, use clinical terms in information processing tools.

Study Objectives

The project's main objective is to use a referral use case to profile, validate, and present data elements relevant to exchanging health information at the PHC level of care. Profiling is the strategy for defining FHIR models by domesticating the international core standard through specific use cases by structured authoring and publishing. Global best practices facilitate digital health information exchange for better care by using standardized data. Digital tools can only communicate using data in certain formats (eg, XML or JSON), organized in an agreed structure [23].

Methods

Overview

We reviewed paper referral forms and surveyed frontline health workers, drawing inspiration from similar work conducted by Odisho et al [24]. We checked how consistent the referral data sets were. Aggregated referral data sets were then mapped to and FHIR extensions profiled. We also modeled data types and cardinalities, including references to other profiles, resources, and terminology binding to ICD-10.

Stakeholder Interviews and Data Set Identification

We established the research focus by addressing the data flow in the maternal and child health information flow value chain in Ebonyi State, Nigeria. Nigeria has between 28,000 and 36,000 health facilities overall. Ebonyi state is one of the 36 subregional governments in Nigeria with 171 “functional” PHC centers and 13 general hospitals [25]. Although from the National Health workforce Registry, there are up to 830 health facilities in the state [26]. Based on our use case, a strategic point of data exchange among multiple PHC centers or PHC centers and hospitals is the referral chain for pregnant women. We used the purposeful snowball sampling technique to identify health care providers in Ebonyi State and share the survey questionnaire.

We sent out questionnaires and a request for a copy of “referral forms” was used for 24 health workers in their respective health facilities in Ebonyi State. Between June 10 and 17, 2019, all 24 health workers completed and returned the questionnaires, and only 3 provided referral paper forms. Respondents were a mix of medical doctors, midwives, and nurses, as shown in Table 1.

We acknowledge the possibility of selection bias, and, for instance, these providers were mostly from health facilities in Abakaliki, the state capital and the main metropolitan city. We consider this bias insignificant as we measured consistency or variation in referral data sets among providers, which was significant. Each provider was from a different health facility (except the tertiary hospital doctors).

The structured questionnaire used asked the following questions:

1. What information is shared when referring-out a pregnant woman?
2. What information is expected when referring-in a pregnant woman?
3. What forms are used, and what are the contents of these forms?
4. What information is the client or caregiver expected to know or have at the recipient end?

Table 1. Distribution of respondents and their roles.

Workstation	Roles, n		
	Nurse	Midwife	Doctor
Primary health care clinics	3	3	4
Secondary health care (general hospitals)	4	2	3
Tertiary health care (teaching hospital)	0	1	3
State Ministry of Health	0	1	0

Profiling, Validation, and Publishing

We started by creating a default patient profile with no extension by using the Forge tool and uploading it on the simplifier.net web interface under the BlockMom project for validation [27]. This first step was to confirm that the example of the base patient resource instance is FHIR-conformant. From the stakeholder interviews, we aggregated information data sets. We then mapped them to the standard patient resource to create a referral resource with extensions that capture all the identified data points. We further created the bare XML schema for easy file-based resource instance validation. The codes in XML and

JSON formats are freely available on the GitHub directory [28]. Our steps and tools used are shown in Figure 1.

We modeled the FHIR referral use case profile of information flow regarding pregnant women from one PHC center—for example, PHC center 1 to PHC center 2—or general hospital. Afterward, these resource mapping outputs were then synthesized into JSON and XML machine-readable data formats on the basis of FHIR resources for antenatal referral. We have further indicated the resources category affected by our referral bundle in green in Figure 2.

Figure 1. Steps to profiling and publishing the Fast Healthcare Interoperability Resource.

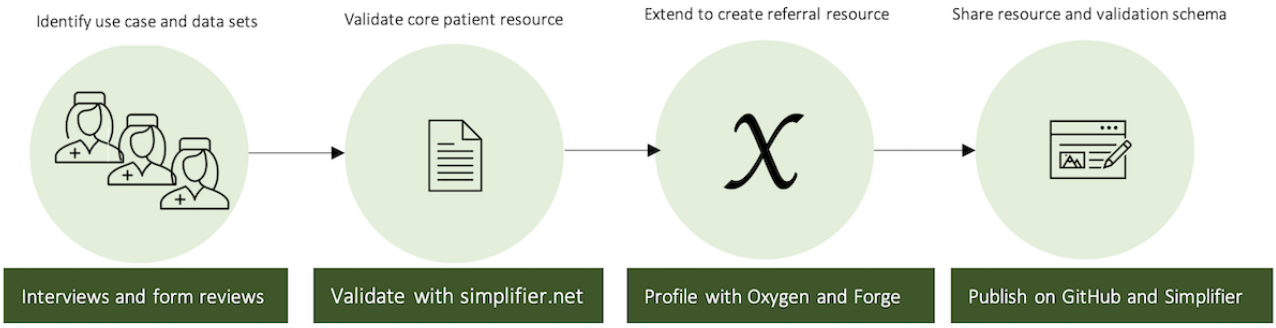
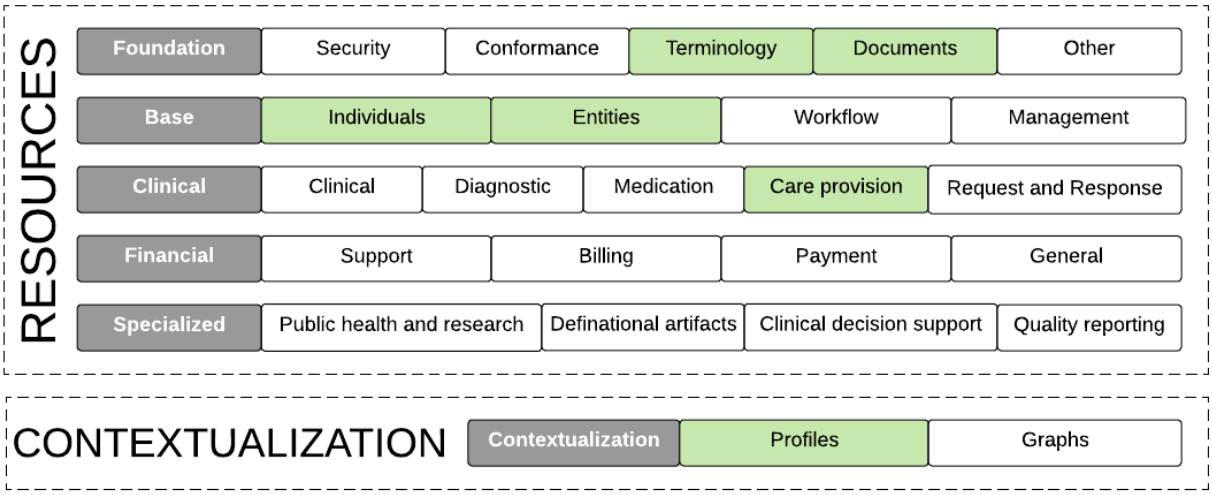


Figure 2. Resources considered for the Referral bundle.



Binding With ICD-10 Terminology

An example referral use case is described in Sierra Leone's digital health strategy 2018-2023, "Table 2- Scenario: The vision in Practice" [12]. Asuma, a pregnant woman described in the use case, was referred to the clinic from the community by a roaming community health extension worker. See page 25 of Sierra Leone's national digital health strategy for more information on this use case. In ICD-10, the "Personal History of malaria" code is Z8613. The code allows for unique identification in any information system using the same coding system, thus distinguishing this from, say, B500, which represents "Plasmodium *falciparum* malaria with cerebral complication," which is a case of complicated malaria with intermittent coma. We analyzed the 21 chapters of ICD-10. Table 2 highlights those that are most relevant for in-depth study when designing a Maternal and Child Health (MNCH) information management system [10].

Code O98.6 represents "Protozoal disease complicating pregnancy, childbirth, and the puerperium." This is synonymous with "Malaria in pregnancy" or "Maternal malaria during pregnancy," both not explicitly coded in ICD-10 [29]. "Malaria in pregnancy" is the scenario described in the Sierra Leone MNCH digital health use case from the preceding paragraph.

Table 2. Key Maternal and Child Health chapters of the International Classification of Diseases, Tenth Revision coding system.

Chapter	Description	Code range
XV	Pregnancy, childbirth, and the puerperium	O00xx to O99xx
XVI	Certain conditions originating in the perinatal period	P00xx to P96xx
XVIII	Symptoms, signs, and abnormal clinical and laboratory findings not elsewhere classified	R00xx to R99xx
XX	External causes of morbidity and mortality	V01xx to Y98xx
XXI	Factors influencing health status and contact with health services	Z00xx to Z99xx

Ethics Consideration

This study was exempted from ethics approval by the University of Malta ethics review board.

Results

Survey Outputs

In addition to responses from these surveys, 3 different referral forms for referral tracking among pregnant women used in the state were made available by respondents. The forms were then mapped to survey questions to generate a unified form with a union of content from the 3 forms in Table 3.

Survey responses from care providers varied widely and included extraneous information than in the referral forms. Based on the 24 health care providers' responses for the first 2 questions—"What Information is shared when referring a pregnant woman?" and "what information is expected when receiving a pregnant woman?"—none of the responses matched for all respondents. In response to question 3—"what forms are used?"—3 respondents said "referral letter" and 7 said "referral form." Three respondents noted that referral forms varied by health institutions, while one indicated that they do not use any forms for referral. Other forms listed by respondents are the consent form, investigation form, chemistry form, hematology form, results form, ultrasound form, laboratory form, radiology form, and virology form. In response to the question, "What information is the client, or their caregiver expected to have or know?" At the same time, 2 respondents said "nill," the rest of them listed information that completely varied. When we mapped the aggregated responses with the form details from Table 3, the list of data sets will be similar to that shown in Textbox 1.

Table 3. Form contents and their mapping.

Ebonyi State Ministry of Health referral slip	National Health Insurance Scheme (NHIS) referral form	Women, infants, and children referral form for pregnant women
Facility name	• Facility name • NHIS code	N/A[a]
N/A	• Date	• Measurement date
Patient number and social insurance number	• Health Management Organization (HMO) • NHIS ID number • HMO code	—
Name of patient	• Name	• Patient's name (last, first)
Age	• Date of birth	• Date of birth (MM/DD/YY)
Sex	• Sex	—
Address	N/A	• Address (state, city, zip code)
N/A	N/A	• Telephone number
Complaints	• Presenting complaint	—
Findings on examination	• Examination findings	• Height, weight • Hemoglobin (g/dL), hematocrit (%), and blood test date
Investigations performed, if any	• Investigation results	—
Provisional diagnosis	• Provisional diagnosis	—
N/A	• Reason for referral post medical history taking • Clinical warnings (allergies)	• Estimated date of confinement • Date when last pregnancy ended • Gravida Para • Pregravida weight (lbs) • Indicate any of the following medical conditions (diabetes, multiple pregnancies, hypertension, tuberculosis, previous poor pregnancy outcome, and history (specify) • If other, current history of the condition (specify)
Current and recent medication	N/A	• Current medication and supplements prescribed
N/A	• Other relevant information	• Impressions and • comments
Name of officer	• Referring doctor • Medical and Dental Council of Nigeria number • Receiving doctor’s Medical and Dental Council of Nigeria number • Date	• Name of the physician care provider group and clinic
Designation	N/A	—
Signature	• Signature and stamp	• Health care provider • Signature Date
To	• Health facility • NHIS code	—

[a]N/A: not applicable.

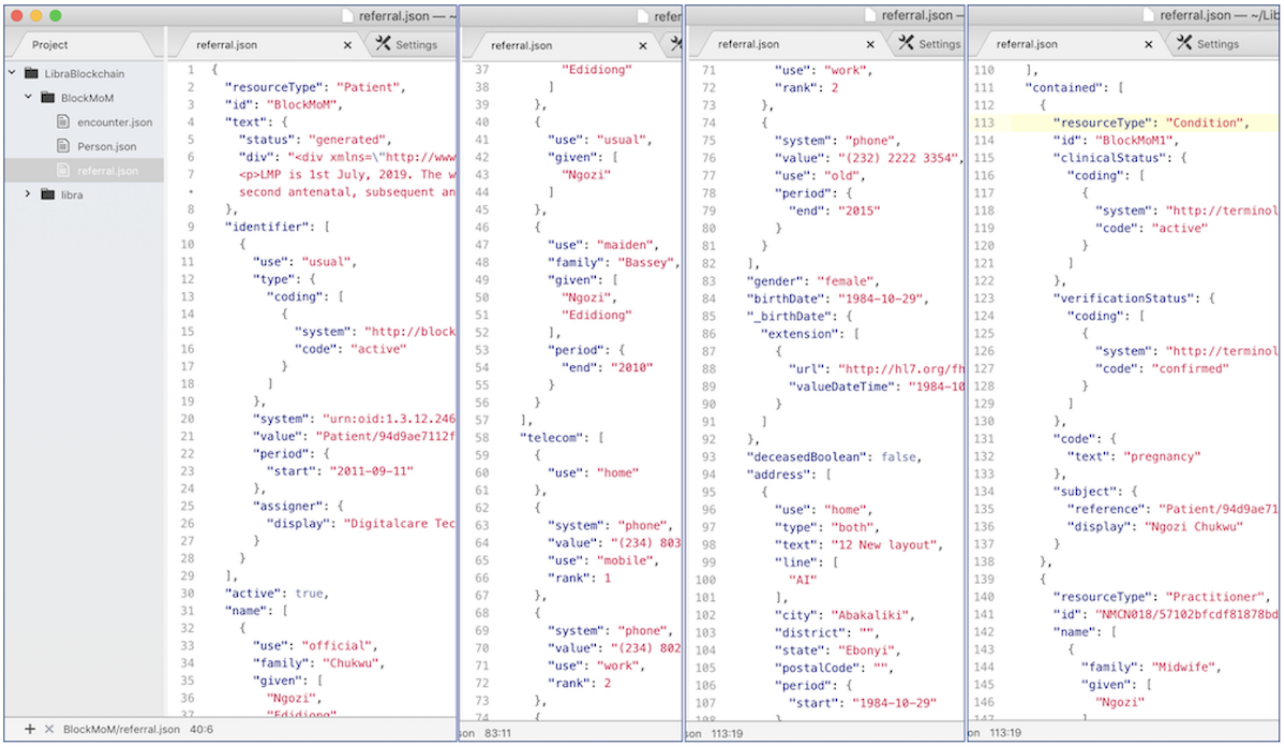
Textbox 1. List of data sets.

- Source health facility name
- Source health facility ID
- Destination health facility name
- Destination health facility ID
- Date of referral
- Patient name
- Patient number
- Patient no type (Health Management Organization, National Health Insurance Scheme, softwareVendor)
- Patient age
- Sex
- Address
- Complaint
- Presenting complaint
- Investigation done
- Findings on examination
- Provisional diagnosis
- Reason for referral
- Current medication
- Recent medication
- Name of referring officer
- ID of referring officer (Medical and Dental Council of Nigeria)
- Designation of the referring officer
- Other relevant information
- Referrals direction (in or out)
- Referrals by disease
- Malaria case referred for adverse drug reaction (Health Management Information System)
- Referral disease (tracked by age and case)

Referral FHIR Resource (Known as Referral Letter or Discharge Letter)

The profile developed in this paper is only considered a provisional national profile suggestion for consideration and should not be relied on for clinical decisions. The resource mapping's final output is in XML and JSON formats and is freely available on GitHub [28]. The resource file is shown in Figure 3.

Figure 3. The profiled referral resource in picture.



REST and FHIR

Our artifact assumes the use of the REST paradigm for information exchange. In this section, we explain the technicalities of REST. REST is the foundation for the scale of the internet as we know it today [30]. While FHIR supports many different communication paradigms, REST is responsible for its popularity. The REST paradigm leverages the HTTP protocol with the simple client-server architecture with variants of catch-less/catching, stateless/stateful, or n-tired architectures and hierarchies. For further technical details of these ICT concepts, please refer to Thomas Fielding's thesis introducing REST in 2000 [30].

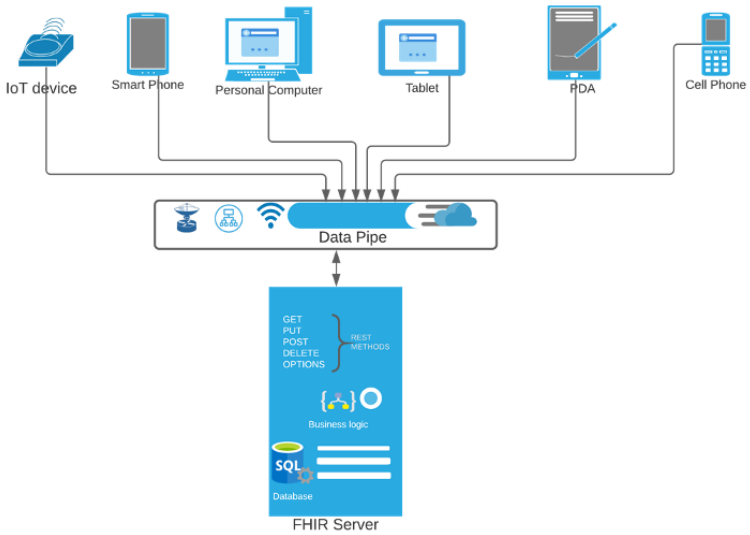
Similar to traditional REST, FHIR's REST paradigm considers entities and concepts as resources. Each of the resource instances is unique and is represented using a uniform resource identifier (URI). The URI may also be used for locating the resource if it points to the location on a given server (in which case, it can also be referred to as URL). There are a finite number of ways a client can manipulate entities and concepts (resources) located in an FHIR server using REST requests of get, put, post, delete, options, head, trace, and connect.

The server responds to the client's HTTP requests after performing internal business logics unique to each server implementation. The clients will receive the same response for similar HTTP requests irrespective of whether they are for a mobile app, web browser, computer application, or embedded device, as shown in Figure 4.

Both the client and server use the header and body components of the HTTP request or response for their information exchange, depending on the HTTP method used. The server always returns a status code indicating success or failure or a variant of either with further detail for each request. FHIR servers can use the OperationOutcome resource to provide structured details of request failure to the clients in the event of failure. There are over 100 different FHIR resources [31]. When the request succeeds, the client is sent the resource by the server.

Methods may or may not be allowed (or even implemented) by the server for a particular resource and may be specified by the client's server response. The header has many attributes that can be set, for instance, to indicate the data type it accepts, authorization credentials, connection, content encoding, caching, and more. The content-type attribute can indicate the resource as either XML or JSON format—both native to FHIR. FHIR, similar to REST, is an open standard and thus aligns with key principles of digital development [32].

Figure 4. A typical client-server REpresentational State Transfer interface. FHIR: Fast Healthcare Interoperability Resource; IoT: Internet of Things; PDA: personal digital assistant.



Discussion

Principal Findings

We could not use the core FHIR resource because it did not contain all the data sets as aggregated in Figure 4. To add and extend the outstanding data sets, FHIR designers provide for and allow extendibility using profile extensions. FHIR contributors and balloting-process use the 80-20 rule to determine what makes it into the core FHIR resource [33]. It is understandable if there are no contributions from Nigeria or many other LMICs because they are often not represented at HL7 FHIR balloting. Our referral form reviews and data set mapping also lay credence to our hypothesis that the core patient resource needs an extension for our use case.

Our study shows that traditional paper referral forms currently in use vary widely in Nigeria's PHC system as illustrated in the word art in Multimedia Appendix 1. The implication is that interinstitution care coordination will remain suboptimal as much of the essential information will be missing. This work will help policy makers and PHC centers in Nigeria understand the need to standardize or enforce agreed referral standards. In addition, the steps we have outlined in this work will help guide institutions as they standardize or adopt FHIR.

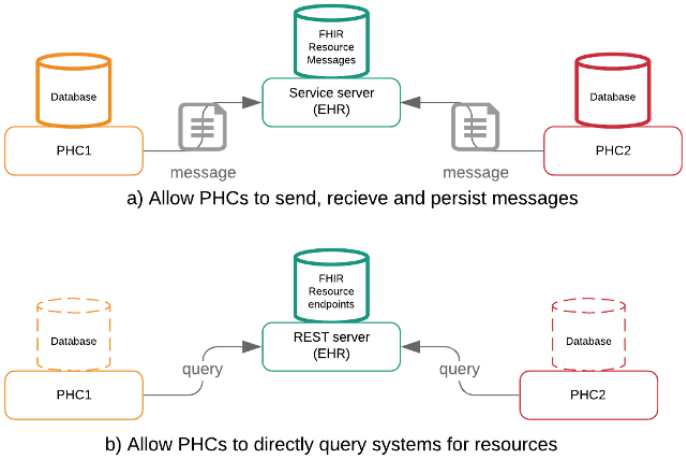
While the surveys were conducted before the COVID-19 pandemic, their findings are relevant for continued functioning of the PHC centers even amid and after the COVID-19 pandemic. The pandemic has exposed the weakness of health systems and shown the importance of interconnected and interoperable health systems. Emerging technologies are being proposed in response to the pandemic [34] and new models are emerging for health information interoperability [35] in LMICs. Even when PHC centers are digitized, referrals among health facilities in many LMICs with different software vendors do not happen seamlessly. Our research shows that referral practices for pregnant women varied significantly even in urban settings.

The key output of this study is the FHIR referral resource artifact, which will help vendors design consistent referral data sets and ensure out-of-the-box interoperability. This resource remains broad and from the core patient resource. FHIR allows for organizational or national extensions and adaptations [33]. Health authorities in many LMIC countries will benefit from standardizing and exposing its required referral data model for Women and Child Health, which encompasses MNCH [24]. Publishing a public FHIR specification that can be leveraged by MNCH solution implementers will help simplify interface implementations.

We here illustrate that our approach differs from traditional health information exchange approaches. Figure 5A shows the traditional document- and message-based information (which is still supported by FHIR) where both databases are required to retain the messages being exchanged. Figure 5B shows that the end point query approach using REST method calls is used to access or share FHIR resources being exchanged.

Women continue to die owing to preventable causes at the point of giving birth. Many of these deaths happen before, during, and after delivery. In addition, maternal health has been highlighted in some LMIC national digital health strategies (ie, Nigeria and Sierra Leone) as a priority health area [11]. Furthermore, our BlockMom model used the ICD terminology over the SNOMED-CT model owing to its favorable pricing license. For instance, SNOMED licenses are based on the number of health facilities using the terminology service, though there are requirements to apply for a waiver for certain implementations in LMICs. Moreover, the model's deliberate focus on FHIR over other HL7 or ISO standards was because it is free and open for adaptation, adoption, and testing in LMICs.

Figure 5. Fast Healthcare Interoperability Resource (FHIR) paradigms in a 2–public health care (PHC) system referral exchange scenario. EHR: electronic health record. REST: REpresentational State Transfer.



Limitations

We used on-file schema and the ICD-10 version 2021 text file to validate profiled resources. Even though the bare schema was provided in the GitHub directory, standard practice would be to set up an FHIR server for this purpose. This is the focus of our future study. Although many mobile-based solutions are available in PHC centers in LMICs, these profiles can only be used with a mobile-based solution that uses patient-specific information rather than aggregated information, as is the case with many community health worker information systems. Furthermore, ICD-11 2022 release has just been launched by the WHO, which emphasizes following the profiling process and not necessarily the output [36].

We also noted that unique identification metrics and characteristics, as explained by McFarlane et al [37] and Chukwu [38], were not part of the data sets proposed and made available during referral. This aspect was profiled as part of the FHIR resource. Besides, this may be the case as the patient-required abridged historical information is assumed to be comprehensive enough. Nevertheless, in a digital platform where a unique identifier is important, this assumption will not hold true. We assumed cryptographically generated unique identification mechanisms in this prototype.

A key limitation of our survey approach was that we did not prevalidate the questionnaire before use. In addition, our sampling methodology used a snowball strategy that targeted health workers from the most urban part of Ebonyi State, Nigeria. We are aware and acknowledge that this may seem inherently biased. However, our aim for the survey was to determine consistency or otherwise of the referral data sets, which we determine to vary widely across all respondents. In addition, since Abakaliki is the state capital and the main metropolitan city in the state, it is expected to have a standardized referral form; however, it does not.

Conclusions

Questionnaire responses were collected from health care providers, and referral forms from health institutions in Nigeria were reviewed. Survey responses and fields of referral forms show variability in referral data sets across respondents and forms. Here we have made a case for FHIR, an emerging health care data interchange standard, and have profiled a referral resource for PHC information exchange targeted at LMICs. This paper describes the profiling steps, including key questionnaire responses and mapping of referral forms. We have proposed the use of ICD-10 terminology and used file-based schema validation. The methodology and artifacts will be invaluable for the research and implementation community targeting LMICs. Our future work will set up the server and configure the appropriate binding for this and other resources.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey respondents' referral keywords.
[PNG File , 2318 KB - formative_v6i7e28510_app1.png]

References

1. World Health Organization. Monitoring the Building Blocks of Health Systems: A Handbook of Indicators and Their Measurement Strategies. Geneva: World Health Organization; 2011.

2. Mutale W, Chintu N, Amoroso C, Awoonor-Williams K, Phillips J, Baynes C, et al. Improving health information systems for decision making across five sub-Saharan African countries: Implementation strategies from the African Health Initiative. BMC Health Serv Res 2013 May 31;13(S2). [doi: 10.1186/1472-6963-13-s2-s9]
3. Warren AE, Wyss K, Shakarishvili G, Atun R, de Savigny D. Global health initiative investments and health systems strengthening: a content analysis of global fund investments. Global Health 2013 Jul 26;9(1):30 [FREE Full text] [doi: 10.1186/1744-8603-9-30] [Medline: 23889824]
4. Kiberu VM, Matovu JK, Makumbi F, Kyozira C, Mukooyo E, Wanyenze RK. Strengthening district-based health reporting through the district health management information software system: the Ugandan experience. BMC Med Inform Decis Mak 2014 May 13;14(1):40 [FREE Full text] [doi: 10.1186/1472-6947-14-40] [Medline: 24886567]
5. Hazel E, Wilson E, Anifalaje A, Sawadogo-Lewis T, Heidkamp R. Building integrated data systems for health and nutrition program evaluations: lessons learned from a multi-country implementation of a DHIS 2-based system. J Glob Health 2018 Dec;8(2):020307 [FREE Full text] [doi: 10.7189/jogh.08.020307] [Medline: 30356499]
6. Hung YW, Hoxha K, Irwin BR, Law MR, Grépin KA. Using routine health information data for research in low- and middle-income countries: a systematic review. BMC Health Serv Res 2020 Aug 25;20(1):790 [FREE Full text] [doi: 10.1186/s12913-020-05660-1] [Medline: 32843033]
7. Health Metrics Network, World Health Organization. Framework and standards for country health information systems (2nd edition). Geneva: World Health Organization; 2008.
8. European Programme of Work: United Action for Better Health. World Health Organization. URL: https://www.euro.who.int/en/countries/estonia/news/news/016/03/e-health-in-practice [accessed 2020-08-30]
9. Greenhalgh T, Stramer K, Bratan T, Byrne E, Russell J, Potts HWW. Adoption and non-adoption of a shared electronic summary record in England: a mixed-method case study. BMJ 2010 Jun 16;340(jun16 4):c3111-c3111. [doi: 10.1136/bmj.c3111] [Medline: 20554687]
10. Benson T, Grieve G. Coding and Classification Schemes. In: Principles of Health Interoperability. Cham: Springer; 2016:135-154.
11. National Health ICT Strategic Framework 2015-2020. Federal Ministry of Health. 2016. URL: http://www.health.gov.ng/doc/HealthICTStrategicFramework.pdf [accessed 2022-06-13]
12. National Digital Health Strategy 2018-2023. Ministry of Health and Sanitation. 2018. URL: https://mohs2017.files.wordpress.com/2019/02/sl-national-digital-health-strategy-nov-2018.pdf [accessed 2021-01-16]
13. Chukwu E, Garg L. A Systematic Review of Blockchain in Healthcare: Frameworks, Prototypes, and Implementations. IEEE Access 2020;8:21196-21214. [doi: 10.1109/access.2020.2969881]
14. World Health Organization. World Health Statistics 2018: Monitoring Health for the SDGs, Sustainable Development Goals. Geneva: World Health Organization; 2018.
15. Mochache V, Irungu E, El-Busaidy H, Temmerman M, Gichangi P. "Our voices matter": a before-after assessment of the effect of a community-participatory intervention to promote uptake of maternal and child health services in Kwale, Kenya. BMC Health Serv Res 2018 Dec 04;18(1):938 [FREE Full text] [doi: 10.1186/s12913-018-3739-9] [Medline: 30514292]
16. Kitui JE, Dutton V, Bester D, Ndirangu R, Wangai S, Ngugi S. Traditional Birth Attendant reorientation and Motherpacks incentive's effect on health facility delivery uptake in Narok County, Kenya: An impact analysis. BMC Pregnancy Childbirth 2017 Apr 21;17(1):125 [FREE Full text] [doi: 10.1186/s12884-017-1307-7] [Medline: 28431565]
17. Paudel D, Shrestha IB, Siebeck M, Rehfuess E. Impact of the community-based newborn care package in Nepal: a quasi-experimental evaluation. BMJ Open 2017 Oct 05;7(10):e015285 [FREE Full text] [doi: 10.1136/bmjopen-2016-015285] [Medline: 28982810]
18. A Free and Open Source Global Good: Powering Interoperability Around the World for 18 Years. HAPI FHIR. URL: https://hapifhir.io [accessed 2021-01-16]
19. HL7 Launches Argonaut Project to Advance FHIR. Health IT Answers. 2014. URL: https://www.healthitanswers.net/hl7-launches-argonaut-project-to-advance-fhir/ [accessed 2021-01-16]
20. Braunstein ML. FHIR. In: Health Informatics on FHIR: How HL7's New API is Transforming Healthcare. Cham: Springer; 2018:179-203.
21. Mehl G, Tunçalp Ö, Ratanaprayul N, Tamrat T, Barreix M, Lowrance D, et al. WHO SMART guidelines: optimising country-level use of guideline recommendations in the digital age. Lancet Digit Health 2021 Apr;3(4):e213-e216 [FREE Full text] [doi: 10.1016/S2589-7500(21)00038-8] [Medline: 33610488]
22. International Statistical Classification of Diseases and Related Health Problems 10th Revision. ICD-10 Version:2019. URL: https://icd.who.int/browse10/2019/en [accessed 2021-01-10]
23. Benson T. Prevention of errors and user alienation in healthcare IT integration programmes. Inform Prim Care 2007 Jan 01;15(1):1-7 [FREE Full text] [doi: 10.14236/jhi.v15i1.639] [Medline: 17612470]
24. Odisho AY, Lui H, Yerramsetty R, Bautista F, Gleason N, Martin E, et al. Design and development of referrals automation, a SMART on FHIR solution to improve patient access to specialty care. JAMIA Open 2020 Oct;3(3):405-412. [doi: 10.1093/jamiaopen/ooaa036] [Medline: 33215075]
25. Overview of the Ministry of Health, Abakaliki. URL: http://www.ebonyistate.gov.ng/Ministry/Health/resources/achievement.pdf [accessed 2019-06-30]

26. NIGERIA Health Facility Registry (HFR). URL: https://hfr.health.gov.ng/statistics/tables [accessed 2021-03-18]
27. MyPatient. Simplifier.net. URL: https://simplifier.net/BlockMom/MyPatient/~xml [accessed 2020-01-14]
28. BlockMom. GitHub. URL: https://github.com/EmekaC/BlockMom [accessed 2021-01-14]
29. Protozoal diseases complicating pregnancy, unspecified trimester. ICD10data.com. URL: https://www.icd10data.com/ICD10CM/Codes/O00-O9A/O94-O9A/O98-/O98.619 [accessed 2021-01-10]
30. REST: Architectural Styles and the Design of Network-based Software Architectures. URL: http://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm [accessed 2022-06-13]
31. HL7 FHIR. URL: http://hl7.org/fhir [accessed 2021-01-10]
32. Principles for Digital Development. URL: http://digitalprinciples.org [accessed 2019-05-31]
33. FHIR and Architectural Principles. HL7 FHIR. URL: https://build.fhir.org/overview-arch.html#:~:text=In%20addition%2C%20FHIR%20aligns%20to,80%25%20of%20the%20interoperability%20needs [accessed 2022-06-13]
34. Ye J. The Role of Health Technology and Informatics in a Global Public Health Emergency: Practices and Implications From the COVID-19 Pandemic. JMIR Med Inform 2020 Jul 14;8(7):e19866 [FREE Full text] [doi: 10.2196/19866] [Medline: 32568725]
35. El Jabari C, Macedo M, Al-jabari MO. Towards a New Paradigm of Federated Electronic Health Records in Palestine. Informatics 2020 Oct 05;7(4):41. [doi: 10.3390/informatics7040041]
36. ICD-11 2022 release. World Health Organization. URL: https://www.who.int/news/item/11-02-2022-icd-11-2022-release [accessed 2022-02-12]
37. McFarlane TD, Dixon BE, Grannis SJ. Client Registries: Identifying and Linking Patients. In: Health Information Exchange. Cambridge, MA: Academic Press; 2016:163-182.
38. Chukwu E. The Case for a Unique Digital Patient ID Scheme in Nigeria. J Health Med Inform 2017;08(03) [FREE Full text] [doi: 10.4172/2157-7420.1000267]

Abbreviations

CEN: European Committee for Standards' Health ICT
DICOM: Digital Imaging and Communication in Medicine
EHR: electronic health record
FHIR: Fast Healthcare Interoperability Resource
ICD: International Classification of Diseases
ICT: Information and Communication Technologies
ISO: International Organization for Standards
LMIC: low- and middle-income country
LOINC: Logical Observation Identifiers Names and Codes
MNCH: Maternal and Child Health
PHC: primary health care
REST: REpresentational State Transfer
RHIS: routine health information systems
SNOMED-CT: Systematized Nomenclature of Medicine–Clinical Terms
URI: uniform resource identifier
WHO: World Health Organization

Edited by A Mavragani; submitted 05.03.21; peer-reviewed by J Ye, B Nievas Soriano; comments to author 14.05.21; revised version received 23.05.21; accepted 14.03.22; published 07.07.22.

Please cite as:
Chukwu E, Garg L, Obande-Ogbuinya N, Chattu VK
Standardizing Primary Health Care Referral Data Sets in Nigeria: Practitioners' Survey, Form Reviews, and Profiling of Fast Healthcare Interoperability Resources (FHIR)
JMIR Form Res 2022;6(7):e28510
URL: https://formative.jmir.org/2022/7/e28510
doi:10.2196/28510
PMID:35797096

Original Paper

Understanding the Security and Privacy Concerns About the Use of Identifiable Health Data in the Context of the COVID-19 Pandemic: Survey Study of Public Attitudes Toward COVID-19 and Data-Sharing

Charlotte Summers[1], BSc; Frances Griffiths[2], PhD; Jonathan Cave[2,3,4], PhD; Arjun Panesar[1], MEng

[1]DDM Health, Coventry, United Kingdom

[2]Warwick Medical School, University of Warwick, Coventry, United Kingdom

[3]Department of Economics, University of Warwick, Coventry, United Kingdom

[4]Data Ethics Group, The Alan Turing Institute, London, United Kingdom

Corresponding Author:
Charlotte Summers, BSc
DDM Health
Technology House, Science Park
University of Warwick
Coventry, CV4 7EZ
United Kingdom
Phone: 44 7969091134
Email: charlotte@ddm.health

Abstract

Background: The COVID-19 pandemic increased the availability and use of population and individual health data to optimize tracking and analysis of the spread of the virus. Many health care services have had to rapidly digitalize in order to maintain the continuity of care provision. Data collection and dissemination have provided critical support for defending against the spread of the virus since the beginning of the pandemic; however, little is known about public perceptions of and attitudes toward the use, privacy, and security of data.

Objective: The goal of this study is to better understand people's willingness to share data in the context of the COVID-19 pandemic.

Methods: A web-based survey was conducted on individuals' use of and attitudes toward health data for individuals aged 18 years and older, and in particular, with a reported diagnosis of a chronic health condition placing them at the highest risk of severe COVID-19.

Results: In total, 4764 individuals responded to this web-based survey, of whom 4674 (98.1%) reported a medical diagnosis of at least 1 health condition (3 per person on average), with type 2 diabetes (n=2974, 62.7%), hypertension (n=2147, 45.2%), and type 1 diabetes (n=1299, 27.4%) being most prominent in our sample. In general, more people are comfortable with sharing anonymized data than personally identifiable data. People reported feeling comfortable sharing data that were able to benefit others; 66% (3121 respondents) would share personal identifiable data if its primary purpose was deemed beneficial for the health of others. Almost two-thirds (n=3026; 63.9%) would consent to sharing personal, sensitive health data with government or health authority organizations. Conversely, over a quarter of respondents (n=1297, 27.8%) stated that they did not trust any organization to protect their data, and 54% (n=2528) of them reported concerns about the implications of sharing personal information. Almost two-thirds (n=3054, 65%) of respondents were concerned about the provisions of appropriate legislation that seeks to prevent data misuse and hold organizations accountable in the case of data misuse.

Conclusions: Although our survey focused mainly on the views of those living with chronic health conditions, the results indicate that data sensitivity is highly contextual. More people are more comfortable with sharing anonymized data rather than personally identifiable data. Willingness to share data also depended on the receiving body, highlighting trust as a key theme, in particular who may have access to shared personal health data and how they may be used in the future. The nascency of legal guidance in this area suggests a need for humanitarian guidelines for data responsibility during disaster relief operations such as pandemics and for involving the public in their development.

(JMIR Form Res 2022;6(7):e29337) doi:10.2196/29337

KEYWORDS

COVID-19; data; data ethics; privacy; sharing; ethics; attitude; perception; data sharing; survey; understanding; security; health data; willingness

Introduction

The World Health Organization declared the COVID-19 outbreak a public health emergency on January 30, 2020; after 6 weeks, it was categorized as a pandemic [1]. Certain groups of people are particularly likely to have serious or severe symptoms of COVID-19 [2]. Preliminary data suggest that people with obesity are at an increased risk of severe COVID-19 [3]. Type 2 diabetes mellitus and hypertension are the most common comorbidities in patients with COVID-19 [4]. According to several reports, including those from the Centers for Disease Control and Prevention, patients with type 2 diabetes are at a greater risk of death than those without type 2 diabetes [5].

Digital health technologies are being used in the fight against COVID-19 [6]. Global health care systems have seen an influx in the incidence of the same novel condition, and the contagious nature of the condition has driven the shift to remote medicine. Many health care pathways have been rapidly digitalized with face-to-face services seeing a drop in usage [7]. This has increased the collection, sharing, and use of data in digital form. Technology is used for remote monitoring, general practitioner consultations, providing structured education, and tracking the spread of disease. As well as the technologies themselves, the data they generate are also useful [8].

Timely, secure, and reliable data access and sharing are critical to understanding COVID-19, controlling its spread, improving the effectiveness and acceptance of government policies, and fostering global cooperation in the race to develop and distribute effective therapies and vaccines. During the COVID-19 pandemic, data are being rapidly shared to understand the location of infections, confirmed cases, recoveries, and deaths. The main data points of interest for this are geolocation and biometric data, both of which are available from users' mobile devices. However, there are serious concerns regarding the objectivity and accuracy of these data, and their utility has been compromised by inconsistent collection and definitions. This, in turn, feeds back into individuals' trust in the collecting organizations and in the extent to which their shared data will actually be used to help others, and this needs to be matched by the trustworthiness of those organizations.

During an unprecedented time, some digital responses to the crisis have precipitated novel data governance and privacy challenges [9]. Governments are taking extraordinary measures to track, trace, and contain the spread of COVID-19 by transitioning to digital technologies and advanced analytics to collect, process, and share data for effective frontline responses. Government-mandated apps are bringing the fight against COVID-19 onto users' devices and have generally adopted pragmatic and contextualized approaches, but they have prompted concerns about security and privacy and the control and use of data beyond the pandemic [10]. There is a trade-off between effectiveness and privacy, centralized and decentralized implementations, and the links to trace and isolate policies.

While the exceptional measures implemented in some countries may prove effective in limiting the spread of the virus, some have provoked controversy in terms of privacy and other fundamental rights, particularly when they lack transparency and public consultation [11]. In South Korea, the specificity of publicly available anonymized data raised privacy concerns when some researchers found that data trails were so detailed that individuals could be identified [12]. In Italy, the Department of Prevention released specific guidelines on the application of the European Union's and national data privacy rules in the context of the COVID-19 pandemic [13]. Similarly, the United Kingdom's Information Commissioner's Office, an independent authority set up to uphold information rights in public interest, confirmed that there would be no regulatory action taken against organizations that fail to meet the data protection standards if noncompliance results from the COVID-19 pandemic [14,15]. In China, new arrivals to the country are tested for COVID-19, instructed to download a government-mandated app, and wear a wristband that is linked to the app to monitor movement with a technology similar to that used in Singapore [16,17]. The United Kingdom's Track and Trace app was the center of a debate on centralization of data [18]. On May 5, 2020, the Government revealed its first attempt at a contact-tracing app, but 6 weeks later admitted that the app was flawed and it would switch to a more privacy-preserving model devised by Apple and Google [19,20]. Transparency is a key theme. One of the most common misconceptions about the United Kingdom's Track and Trace app was that it could allow users to specifically identify and map COVID-19 cases among their contacts and in their vicinity [21].

This study seeks to understand the opinions of British people with long-term health conditions on the themes of data privacy and security, data ethics, and data misuse and to assess the possible trade-offs in data utilization to manage a crisis such as the COVID-19 pandemic [22]. It is important to understand the concerns of people with long-term health conditions such as type 2 diabetes and hypertension as these conditions have been shown to be key risk factors in the progression and prognosis of COVID‐19 [23,24].

Methods

Study Design and Setting

A web-based survey study was conducted with a mixed methods design conforming to the checklist for reporting results of internet electronic surveys [25]. An email invitation to participate, which included a weblink to the survey, was sent to 11,213 people who had consented to be contacted for research opportunities.

Quantitative information (closed and multiple-choice questions) was collected on four topics: (1) demographic characteristics, (2) COVID-19 symptoms and clinical diagnoses, (3) sharing and privacy of pre– and post–COVID-19 health data, and (4) COVID-19 lockdown behaviors. Responses from the final topic are not included in this analysis.

The survey contained 31 questions: 26 closed questions, 1 open question, and 4 demographic questions. Questions on sharing and privacy of pre– and post–COVID-19 health data were answered on a 5-point Likert scale with responses ranging from *strongly disagree* to *strongly agree* or from *not concerned at all* to *very concerned*.

Participants

People aged ≥18 years who had joined the Diabetes.co.uk community were surveyed. The survey commenced with 1 screening question: "Do you consent to take part in the study?" Respondents who consented went on to complete the survey.

Procedure

Data collection occurred between July 6 and August 31, 2020. The survey was administered through the Jisc Online Surveys software and comprised closed, open, and multiple-choice questions. The survey was designed to elicit individual responses to questions about retrospective data use and privacy prior to the COVID-19 pandemic and prospective use during the COVID-19 pandemic.

It is intended to have multiple windows of data collection for several reasons: people's recollections of pre–COVID-19 attitudes may be unreliable, and changes in the course, apparent seriousness, and confidence in scientific understanding of the pandemic will have evolved.

The type and wording of each question was composed by the research team. The order of questions was not randomized. The survey followed a predetermined logic where contingent questions were included or automatically skipped on the basis of responses. Qualitative data were collected with 1 open question exploring what respondents would like to see happening: "What would you like to see happen to improve the COVID-19 situation?" (question 30).

Analysis

We exported all data from Jisc and conducted data analysis using SPSS (version 22; IBM Corp). We conducted descriptive data analyses of sample distributions and characteristics. Pearson *r* correlation coefficients were used to determine the relation between prior data-sharing behavior and attitudes toward data-sharing activity in the context of the COVID-19 pandemic. The data from the open question were read through and then categorized into themes.

Ethical Considerations

Ethics approval was obtained from the Human Research Ethics Committee of the University of Warwick (BSREC 144/19-20). Web-based informed consent was required before the survey could be accessed.

Results

Survey Respondents

Of 11,213 people emailed, 10,705 clicked through to the survey; in total, 4764 gave their consent and began the survey. As indicated in Table 1, all of them completed the survey and were included in the analysis. All respondents were located in the United Kingdom. In total, 2287 (48.0%) respondents were male and 3083 (64.8%) were aged between 55 and 74 years. A total of 115 (2.8%) respondents reported having been clinically diagnosed with COVID-19. The majority of patients (n=4674, 98.1%) reported a prior clinical diagnosis of at least one health condition (on average 3 per person). There was a high prevalence of individuals living with type 2 diabetes (n=2974, 62.7%), hypertension (n=2147, 45.2%), type 1 diabetes (n=1299, 27.4), obesity (n=892, 18.8%), and depression (n=871, 18.3%). Respondent demographics are shown in Table 1.

Table 1. Respondent demographics (N=4764).

Characteristics	Respondents, n (%)
Gender	
Male	2287 (48.0)
Female	2435 (51.1)
Prefer not to say	42 (0.9)
Age (years)	
18 to 24	23 (0.5)
25 to 34	104 (2.2)
35 to 44	298 (6.3)
45 to 54	839 (17.6)
55 to 64	1550 (32.6)
65 to 74	1533 (32.2)
75 or older	410 (8.6)
Prefer not to say	7 (0.1)
Health conditions[a]	
Type 2 diabetes	2974 (62.7)
Hypertension	2147 (45.2)
Type 1 diabetes	1299 (27.4)
Arthritis	1002 (21.1)
Obesity	892 (18.8)
Depression	871 (18.3)
Employment[b]	
Full-time employment	1205 (24.9)
Part-time employment	571 (11.8)
Retired	2298 (47.5)
Student	39 (0.8)
Unemployment	464 (9.6)
Furloughed	188 (3.9)
Volunteering in my community (National Health Service, key services)	69 (1.4)
Ethnicity	
Indian or Pakistani	69 (1.4)
Black, British African, or Caribbean	46 (1.0)
Middle Eastern	5 (0.1)
Mixed groups	27 (0.6)
White	4,434 (93.1)
Other	44 (0.9)
Chinese, Japanese, or East Asian	7 (0.1)
Prefer not to say	132 (2.8)

[a]Frequently occurring health conditions selected.

[b]Respondents selected multiple statuses; for example, full-time employed and furloughed.

COVID-19 Symptoms and Clinical Diagnosis

Of the 4764 respondents who completed the survey, 494 stated they had had symptoms of COVID-19. The most common symptoms were the following: a continuous cough (n=467, 94%), fever (n=325, 65.7%), difficulty breathing (n=384, 77.7%), and loss of taste (n=324, 65.5%). In total, 384 (77.7%) respondents reported another symptom, predominantly fatigue (16.4%).

Of those reporting symptoms, 111 (22.5%) reported a clinical diagnosis of COVID-19. Of these respondents, 73 (63.5%) reported that their symptoms were severe or very severe. In total, 26 (22.6%) respondents reported that their symptoms were not severe at all. All respondents who reported a clinical diagnosis of COVID-19 reported at least one symptom, including loss of smell or taste (63.1%), fever (62.2%), difficulty breathing (61.3%), or continuous cough (53.2%).

A total of 131 (2.8%) respondents reported that a household member had been tested and was clinically diagnosed with COVID-19.

Sharing and Privacy of Pre–COVID-19 Health Data

Prior to the COVID-19 pandemic, almost half of the respondents (n=2313, 49.2%) agreed or strongly agreed that they often consented to anonymized sharing of their private health data, while only 608 (13%) respondents often consented to sharing of private health data without anonymization. Two-thirds of respondents (n=3113, 66.7%) disagreed or strongly disagreed with sharing their private health data without anonymization. Similarly, 3121 (66.3%) respondents would share their data if it keeps other people healthy; 3026 (63.9%) respondents agreed or strongly agreed to sharing private health data with the government or health authority; 1911 (40.7%) respondents agreed or strongly agreed to share their private health data with services that provide health services to the National Health Service (NHS) such as the Low Carb Program and PushDoctor. Only 232 (5%) participants agreed or strongly agreed to share private health data with social media platforms. Over a quarter of respondents (n=1297, 27.8%) agreed or strongly agreed that they did not trust *any* organization to protect their private health data. Just under a quarter of respondents (n=1094, 23.5%) agreed or strongly agreed that they were not concerned by the implications of sharing private health data. General health data–sharing responses are shown in Table 2. Respondents who reported that they felt “neutral” in response to the statements were excluded.

Table 2. General health data sharing responses.

Question	Disagree or strongly disagree, n (%)	Agree or strongly agree, n (%)
I often consent to share my private health data with any organisation as long as it is anonymised	1273 (27.1)	2313 (49.2)
I often consent to share my private health data to any organisation without anonymisation	3113 (66.7)	599 (13)
I am not concerned about the implications of sharing my private health data	2528 (54.2)	1094 (23.5)
I don't trust any organisation to protect my private health data	1541 (33.8)	1297 (27.8)
I'm happy to share my private health data if it helps keep other people healthy	526 (11.2)	3121 (66.3)

Sharing and Privacy of Post–COVID-19 Health Data

Over half (n=3026, 63.9%) agreed or strongly agreed to share their private data with the government or health authority if asked; 1911 (40.7%) respondents would happily consent to share their private data with services that provide health services to the NHS such as the Low Carb Program and PushDoctor, if asked. Only 232 (5%) participants agreed or strongly agreed that they would consent to sharing private data with social media if asked.

Almost half of respondents (n=2228, 47.1%) were concerned or very concerned about who would have access to their personal health data in the context of the COVID-19 pandemic and 2310 (49.1%) respondents were concerned or very concerned about how their personal health data may be used in the future. Almost two-thirds of respondents (n=3054, 65%) were concerned or very concerned around the legislation of data misuse.

Just over a third of respondents (n=1563, 33.4%) would consent to share their private data with any organization if it was providing essential COVID-19 support services such as the supermarkets, pharmacies, and banks. Responses toward the use of post–COVID-19 patient data is shown in Table 3, along with the sentiment toward the use of patient data in the context of the COVID-19 pandemic, and Table 4 shows the sentiment toward future use or misuse of data collected and used under the provisions of the COVID-19 pandemic.

Table 3. Responses toward the use of post–COVID-19 patient data.

Question	Disagree or strongly disagree	Agree or strongly agree
I would happily consent to share my private health data with the government or health authority	728 (15.4)	3026 (63.9)
I would happily consent to share my private health data with social media e.g Twitter, Facebook, Google	4023 (85.9)	232 (5)
I would happily consent to share my private health data with services that provide health services to the NHS[a] such as Low Carb Program, PushDoctor, Babylon Health	1351 (18.8)	1911 (40.7)

[a]NHS: National Health Service.

Table 4. Sentiment toward future use or misuse of data collected and used under the provisions of the COVID-19 pandemic.

Question	Not concerned at all	Concerned or very concerned
In light of COVID-19, how concerned are you about who would have access to your personal health data?	1138 (14)	2228 (47.1)
How concerned are you about how your personal health data may be used in the future?	1162 (24.7)	2310 (49.1)
How concerned are you around the legislation of data misuse?	644 (13.6)	3054 (65)

Prior Willingness to Share Data

Correlations between retrospective data-sharing that happened in the context of generalized concerns and attitude changes associated with the course of the pandemic were determined. Changes were not linked to any specific studies, policies, or measures. There were strong correlations in the attitudes of people exhibiting high levels of concern about future uses of shared data and concerns about access (r_{4685}=0.816; P<.001). There was a strong correlation between people exhibiting attitudes of concern that firmer legislation for data misuse is needed and concerns about future repurposing and reuse of personal health data collected during the COVID-19 pandemic (r_{4663}=0.636; P<.001). The Pearson r correlation coefficient as a normalized measure of the strength of a possible linear correlation, lying between –1 and +1. The Pearson r correlation coefficient measures nonlinear correlations (eg, when extreme views are highly correlated but more moderate ones are more independent).

Respondents agreed to share their personal data with roughly the same parties prior to the COVID-19 pandemic and within the context of the COVID-19 pandemic; governments and health authorities (r_{4710}=0.762; P<.001), health service providers such as the Low Carb Program and PushDoctor (r_{4662}=0.783; P<.001), and social media platforms such as Twitter, Facebook, and Google (r_{4662}=0.736; P<.001).

COVID-19 News and Information

Of the 4764 respondents, 2666 (56.1%) were concerned that they may be receiving misinformation about COVID-19 from trusted sources, 1079 (22.7%) were not concerned (genuinely unconcerned and those who feel that they are in control of the consumption of news and information), and 1006 (21.2%) had never considered it.

In total, 4237 responded to the open-ended question of what they would like to see happening to improve the COVID-19 situation. The majority of respondents shared a single response: 1348 (31.8%) stated they would like to see a reliable vaccine and treatment, 884 (20.8%) stated they would like to see balanced information from the government, and 485 (11.4%) wanted to see stricter measures to prevent the transmission of COVID-19.

Discussion

Principal Findings

Our study provides insights into public perception and attitudes toward the use of identifiable health data in the context of the COVID-19 pandemic; in particular, the perspectives of those living with chronic, long-term health conditions, with an average of 4 health conditions reported per respondent.

Our study suggests that data sensitivity is highly contextual. A significant proportion of people felt that their own attitudes have shifted as a result of the COVID-19 pandemic. More people reported being comfortable with sharing private health data with any organization during rather than before the COVID-19 pandemic. In order, people appear to trust their data with the government, health organizations, and social media. There is significant distrust of private health data use by social media organizations (eg, Twitter, Facebook, and Google) even though social media is used as a channel for communication by people caught up in crises such as emergency relief operations after earthquakes, tsunamis, and typhoons; where it provides a trusted and highly salient source of information about what is happening and what to do [26,27]. This is surprising as although users worldwide report that privacy and use of personal data are important issues, most rarely make an effort actively to protect these data and often even give them away voluntarily on social media where even innocuous data can reveal sensitive health information when suitably processed [28,29]. People treat data revelation and sharing differently depending on the perceived sensitivity of the data, and the sensitivity attached to different types of data is neither stable nor uniform.

When examining the correlations between retrospective views of data-sharing behavior and comfort regarding data-sharing in

the context of the COVID-19 pandemic, individuals were comfortable (or not) in sharing personal data with the same organizations prior to and during the COVID-19 pandemic, suggesting that COVID-19 has not drastically shifted people's willingness to share or withhold their personal data. This may be because attitudes have shifted both retrospectively and prospectively and also depends on whether people accurately remember and report their past views and actions. One of the strongest correlations observed in the analysis was between high levels of concern about the requirement for stronger legislation protecting individuals from data misuse and future repurposing and reuse. This highlights the need for improved communication, transparency, and potentially stronger regulation on how such data may be repurposed in the future, who will be accountable for inappropriate use of data, and a commitment to cease or reverse exceptional uses of data when the crisis is over. Individuals' data rights are protected by law in regulation such as General Data Protection Regulation 2018 in Europe and Health Insurance Portability and Accountability Act in the United States, which make clear the scope, purpose, and time limitations of data usage [30,31]. Concerns may therefore reflect ignorance of existing rules, doubts over enforcement, or a belief that current legislation does not go far enough (for instance, in the requirement of erasure after 3 years rather than a shorter time duration).

A key theme emerging from the literature that was confirmed in this study is the importance of trust [32-34]. Over a quarter of respondents stated they did not trust any organization to protect their data, over half reported concern about the implications of sharing personal information, and almost two-thirds were concerned about data misuse regulation not being strict enough. When asked during the pandemic (the United Kingdom's first wave), almost half of respondents were concerned about who would have access to their personal health data and a similar number were concerned about how their personal health data might be used in the future. This is consistent with prior research suggesting that public involvement in data policy is crucial to bolstering trust and provides support for legislation that is more enforceable [35]. Attitudes may have been perturbed by news stories relating to cybersecurity and privacy and by policy announcements (eg, around Huawei, the Online Harms Bill, etc) [36,37].

Although there are no directly comparable studies, the results from this study complement prior research on public perceptions about COVID‐19 and data-sharing. Data privacy and protection are important concepts [38]. Data policy tends to address human concerns about privacy by making rules about data protection; however, this can lead to category errors since data protection can undermine privacy.

Willingness to share anonymized personal health information varies depending on the degree to which the receiving body is trusted and the uses to which the data will be put [39,40]. The more commercial the objectives of the receiving institution appear, the less respondents are willing to share their personal health information. This in turn suggests that anonymization's disadvantages (in terms of confirming data and correlating shared with other data) might be offset by better (wider, deeper, and more accurate) sampling leading to greater validity of results. Further evidence comes from the interaction (or correlation) between these attitudinal responses and other characteristics, meaning that nonanonymized collection might lead to biased results.

Virus tracking apps are used at scale by governments; however, concerns about transparency, privacy, and morality remain [41,42]. There has been substantial research into the challenges involved in the digital response to the COVID-19 pandemic and proposed methodologies for the ethical design and use of digital public health tools [43,44]. Clear and effective data ethics is both a moral and a practical obligation. The nascency of legal guidance in this area combining ethics, law, and humanitarian impulses suggests the requirement for humanitarian guidelines for data responsibility during global crises such as pandemics. Therefore, rather than recalibrating the expectations of people with regard to their own privacy, the requirements for the use of data should be broader and more comprehensive as ethically collected big data could prove to be extremely useful in the prediction, monitoring, and mitigation of pandemics such as COVID-19 [45].

Strengths and Limitations

Despite the importance of the findings reported here, it is important to note that this study had several limitations. Conducting this study via a web-based survey carries a risk of response bias, simply because the respondents are likely to be more technology-savvy than the general population. However, the population studied (those with chronic health conditions) is of interest as these participants have a degree of awareness and the ability to self-manage their condition that is not (yet) typical of the population at large, and this sheds light on how policies that raise awareness may lead to greater effectiveness in terms of uptake of technical solutions and effectiveness of public health advice and other policies.

Participants were asked to rate retrospectively their perceptions of data sharing prior to the COVID-19 pandemic; these ratings may be inaccurate owing to faulty memories and response bias. There is some ambiguity between what people (now) thought they would have done had they been asked and how they responded to actual requests for consent to data-sharing. In particular, one could disagree with a statement like "I often consent to share my private health data to any organisation without anonymisation" simply because one rarely recalls being asked to share (even without a principled objection to such sharing should the occasion arise). In addition, the phraseology of the questionnaire refers to data that have either been anonymized or are identifiable. This dichotomous representation leaves out pseudonymized data. The participants were not educated about the concepts of anonymization prior to answering the questions potentially allowing ambiguity of the terminology to cause a strong bias in the response behavior.

In addition, the sample is concentrated on people with diabetes and those with other diagnosed health conditions, rather than the general population. This is a strength as well as a weakness as it focuses on a population with particular circumstances and perspectives and one that may be more representative of a post–COVID-19 population that has been sensitized to a continuing health concern than the current population. This in

turn means that a comparison of these findings with a similar survey of the general population can shed light on the potential impact of awareness-raising policies.

Another strength of the study is the high number of respondents who completed the questionnaire. In total, 4764 people participated in the study. This provides a unique insight into the views of a population deemed as being at the highest risk of severe disease and mortality related to COVID-19 [46]. There was a skew in the representation of the demographic distribution of individuals in the nationwide population of people living with chronic health conditions, since White people were overrepresented in our sample (n=4434, 93.4%) but not overrepresented among those with diabetes more generally. While our survey focused on those with diabetes, the results provide novel insight into concepts crucial for societal trust in data use and sharing initiatives.

While the study design did not allow us to ascertain whether technology use itself was correlated with higher acceptance of data-sharing, such an analysis is possible and will be an important topic for future research.

The study's findings suggest potential targets for further study and possible considerations for policy makers. There are two main implications: storing and processing data in pseudonymized form and emphasizing the use of synthetic data (generated from models estimated from real data but not involving any actual or identifiable human beings).

Understanding attitudes toward data sensitivities and trust can contribute to developing policies, improving transparency, and increasing the trust, speed, focus, and effectiveness of epidemic responses. Future practice should emphasize transparent data-sharing and privacy initiatives, while research should evaluate whether this does indeed lead to greater levels of trust and engagement. Encouraging ethical and relevant data-sharing can provide significant epidemic intelligence and support public health emergency relief operations [47].

Conclusions

Data sensitivity is highly contextual. More people are comfortable with sharing anonymized data than personally identifiable data. Willingness to share data also varied depending on the receiving body, highlighting trust as a key theme, who may have access to shared personal health data and how it may be used in the future. The nascency of legal guidance in this area suggests the requirement for humanitarian guidelines for data responsibility during disaster relief operations such as pandemics, and the requirement to involve the public in their development.

Acknowledgments

The authors acknowledge and thank the Diabetes.co.uk Forum community, which supported this study.

Conflicts of Interest

CS and AP are founding employees of DDM, who operate the Diabetes.co.uk Forum community.

References

1. Chan M, Yeo SE, Chong Y, Lee Y. Stepping Forward: Urologists' Efforts During the COVID-19 Outbreak in Singapore. Eur Urol 2020 Jul;78(1):e38-e39 [FREE Full text] [doi: 10.1016/j.eururo.2020.03.004] [Medline: 32192816]
2. Bornstein SR, Dalan R, Hopkins D, Mingrone G, Boehm BO. Endocrine and metabolic link to coronavirus infection. Nat Rev Endocrinol 2020 Jun 02;16(6):297-298 [FREE Full text] [doi: 10.1038/s41574-020-0353-9] [Medline: 32242089]
3. Stefan N, Birkenfeld AL, Schulze MB, Ludwig DS. Obesity and impaired metabolic health in patients with COVID-19. Nat Rev Endocrinol 2020 Jul;16(7):341-342 [FREE Full text] [doi: 10.1038/s41574-020-0364-6] [Medline: 32327737]
4. Sun Y, Guan X, Jia L, Xing N, Cheng L, Liu B, et al. Independent and combined effects of hypertension and diabetes on clinical outcomes in patients with COVID-19: A retrospective cohort study of Huoshen Mountain Hospital and Guanggu Fangcang Shelter Hospital. J Clin Hypertens (Greenwich) 2021 Feb 25;23(2):218-231 [FREE Full text] [doi: 10.1111/jch.14146] [Medline: 33369066]
5. Cole SA, Laviada-Molina HA, Serres-Perales JM, Rodriguez-Ayala E, Bastarrachea RA. The COVID-19 Pandemic during the Time of the Diabetes Pandemic: Likely Fraternal Twins? Pathogens 2020 May 19;9(5):389 [FREE Full text] [doi: 10.3390/pathogens9050389] [Medline: 32438687]
6. Scott BK, Miller GT, Fonda SJ, Yeaw RE, Gaudaen JC, Pavliscsak HH, et al. Advanced Digital Health Technologies for COVID-19 and Future Emergencies. Telemed J E Health 2020 Oct 01;26(10):1226-1233. [doi: 10.1089/tmj.2020.0140] [Medline: 32456560]
7. Virtual Solutions for Managing Cancer Care In a Pandemic Era: Lessons from COVID-19. A Rapid Evidence Review. NHS Peninsula Cancer Alliance. 2020. URL: https://peninsulacanceralliance.nhs.uk/wp-content/uploads/2020/10/Virtual-solutions-for-managing-cancer-care_FINAL-01Oct2020.pdf [accessed 2022-06-01]
8. Panesar A, Panesar H. Artificial intelligence and machine learning in global healthcare. In: Handbook of Global Health. Berlin: Springer; 2020:1-39.
9. Ensuring data privacy as we battle COVID-19. OECD. 2020 Apr 14. URL: http://www.oecd.org/coronavirus/policy-responses/ensuring-data-privacy-as-we-battle-covid-19-36c2f31e/ [accessed 2022-06-01]

10. Hoffman AS, Jacobs B, van Gastel B, Schraffenberger H, Sharon T, Pas B. Towards a seamful ethics of Covid-19 contact tracing apps? Ethics Inf Technol 2020 Sep 28;23(S1):1-11 [FREE Full text] [doi: 10.1007/s10676-020-09559-7] [Medline: 33013191]
11. Henderson J. Patient privacy in the COVID-19 era: Data access, transparency, rights, regulation and the case for retaining the status quo. Health Inf Manag 2021;50(1-2):6-8. [doi: 10.1177/1833358320966689] [Medline: 33176491]
12. Ahn NY, Park JE, Lee DH, Hong PC. Balancing Personal Privacy and Public Safety During COVID-19: The Case of South Korea. IEEE Access 2020;8:171325-171333. [doi: 10.1109/access.2020.3025971]
13. Torri E, Sbrogiò LG, Rosa ED, Cinquetti S, Francia F, Ferro A. Italian Public Health Response to the COVID-19 Pandemic: Case Report from the Field, Insights and Challenges for the Department of Prevention. Int J Environ Res Public Health 2020 May 22;17(10):3666 [FREE Full text] [doi: 10.3390/ijerph17103666] [Medline: 32456072]
14. Data protection and Coronavirus-19 – relaxation of government measures. Information Commissioner's Office. URL: https://ico.org.uk/global/data-protection-and-coronavirus-information-hub/data-protection-and-coronavirus/ [accessed 2021-03-27]
15. The ICO's regulatory approach. Information Commissioner's Office. URL: https://ico.org.uk/media/about-the-ico/policies-and-procedures/2617613/ico-regulatory-approach-during-coronavirus.pdf [accessed 2021-03-27]
16. Abuhammad S, Khabour OF, Alzoubi KH. COVID-19 Contact-Tracing Technology: Acceptability and Ethical Issues of Use. Patient Prefer Adherence 2020;14:1639-1647 [FREE Full text] [doi: 10.2147/PPA.S276183] [Medline: 32982188]
17. Simko L, Calo R, Roesner F, Kohno T. COVID-19 Contact Tracing and Privacy: Studying Opinion and Preferences. arXiv. Preprint posted online on May 12, 2020. [FREE Full text]
18. Sweeney Y. Tracking the debate on COVID-19 surveillance tools. Nat Mach Intell 2020 Jun 16;2(6):301-304. [doi: 10.1038/s42256-020-0194-1]
19. Alanzi T. A Review of Mobile Applications Available in the App and Google Play Stores Used During the COVID-19 Outbreak. JMDH 2021 Jan;14:45-57. [doi: 10.2147/jmdh.s285014]
20. Wise J. Covid-19: UK drops its own contact tracing app to switch to Apple and Google model. BMJ 2020 Jun 19;369:m2472. [doi: 10.1136/bmj.m2472] [Medline: 32561511]
21. Williams SN, Armitage CJ, Tampe T, Dienes K. Public attitudes towards COVID-19 contact tracing apps: A UK-based focus group study. Health Expect 2021 Apr;24(2):377-385 [FREE Full text] [doi: 10.1111/hex.13179] [Medline: 33434404]
22. Cairney P. The UK government's COVID-19 policy: assessing evidence-informed policy analysis in real time. Br Polit 2020 Nov 01;16(1):90-116. [doi: 10.1057/s41293-020-00150-8]
23. Guo W, Li M, Dong Y, Zhou H, Zhang Z, Tian C, et al. Diabetes is a risk factor for the progression and prognosis of COVID-19. Diabetes Metab Res Rev 2020 Mar 31:e3319 [FREE Full text] [doi: 10.1002/dmrr.3319] [Medline: 32233013]
24. Lippi G, Wong J, Henry BM. Hypertension in patients with coronavirus disease 2019 (COVID-19): a pooled analysis. Pol Arch Intern Med 2020 Apr 30;130(4):304-309 [FREE Full text] [doi: 10.20452/pamw.15272] [Medline: 32231171]
25. 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]
26. Palen L, Hughes AL. Social Media in Disaster Communication. In: Handbook of Disaster Research. Cham: Springer; 2017:497-518.
27. Muniz-Rodriguez K, Ofori SK, Bayliss LC, Schwind JS, Diallo K, Liu M, et al. Social Media Use in Emergency Response to Natural Disasters: A Systematic Review With a Public Health Perspective. Disaster Med Public Health Prep 2020 Feb 09;14(1):139-149. [doi: 10.1017/dmp.2020.3] [Medline: 32148219]
28. Gerber N, Gerber P, Volkamer M. Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. Comput Secur 2018 Aug;77:226-261. [doi: 10.1016/j.cose.2018.04.002]
29. Raij A, Ghosh A, Kumar S, Srivastava M. Privacy risks emerging from the adoption of innocuous wearable sensors in the mobile environment. 2011 Presented at: CHI '11: CHI Conference on Human Factors in Computing Systems; May 7-12, 2011; Vancouver, BC. [doi: 10.1145/1978942.1978945]
30. Goddard M. The EU General Data Protection Regulation (GDPR): European Regulation that has a Global Impact. Int J Mark Res 2017 Nov 01;59(6):703-705. [doi: 10.2501/ijmr-2017-050]
31. Cohen IG, Mello MM. HIPAA and Protecting Health Information in the 21st Century. JAMA 2018 Jul 17;320(3):231-232. [doi: 10.1001/jama.2018.5630] [Medline: 29800120]
32. Bargain O, Aminjonov U. Trust and compliance to public health policies in times of COVID-19. J Public Econ 2020 Dec;192:104316 [FREE Full text] [doi: 10.1016/j.jpubeco.2020.104316] [Medline: 33162621]
33. Ienca M, Vayena E. On the responsible use of digital data to tackle the COVID-19 pandemic. Nat Med 2020 Apr 27;26(4):463-464 [FREE Full text] [doi: 10.1038/s41591-020-0832-5] [Medline: 32284619]
34. Bunker D. Who do you trust? The digital destruction of shared situational awareness and the COVID-19 infodemic. Int J Inf Manage 2020 Dec;55:102201 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102201] [Medline: 32836649]
35. Liabo K, Boddy K, Bortoli S, Irvine J, Boult H, Fredlund M, et al. Public involvement in health research: what does 'good' look like in practice? Res Involv Engagem 2020 Mar 31;6(1):11 [FREE Full text] [doi: 10.1186/s40900-020-0183-x] [Medline: 32266085]
36. Lysne O. Containment of Untrusted Modules. In: The Huawei and Snowden Questions. Cham: Springer; 2018:99-107.

37. Consultation outcome: Online Harms White Paper. Department for Digital, Culture, Media & Sport. Government of the United Kingdom. 2020 Dec 15. URL: https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper [accessed 2022-06-01]
38. Zwitter A, Gstrein OJ. Big data, privacy and COVID-19 – learning from humanitarian expertise in data protection. Int J Humanitarian Action 2020 May 18;5(1). [doi: 10.1186/s41018-020-00072-6]
39. Ghafur S, Van Dael J, Leis M, Darzi A, Sheikh A. Public perceptions on data sharing: key insights from the UK and the USA. Lancet Digit Health 2020 Sep;2(9):e444-e446. [doi: 10.1016/s2589-7500(20)30161-8]
40. Panesar A. Machine Learning and AI Ethics. In: Machine Learning and AI for Healthcare: Big Data for Improved Health Outcomes. Berkeley, CA: Apress; Dec 16, 2020:207-247.
41. Williams SN, Armitage CJ, Tampe T, Dienes K. Public attitudes towards COVID-19 contact tracing apps: A UK-based focus group study. Health Expect 2021 Apr;24(2):377-385 [FREE Full text] [doi: 10.1111/hex.13179] [Medline: 33434404]
42. Ahmed N, Michelin RA, Xue W, Ruj S, Malaney R, Kanhere SS, et al. A Survey of COVID-19 Contact Tracing Apps. IEEE Access 2020;8:134577-134601. [doi: 10.1109/access.2020.3010226]
43. Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020 Aug 07;26(8):1183-1192. [doi: 10.1038/s41591-020-1011-4] [Medline: 32770165]
44. Gasser U, Ienca M, Scheibner J, Sleigh J, Vayena E. Digital tools against COVID-19: taxonomy, ethical challenges, and navigation aid. Lancet Digit Health 2020 Aug;2(8):e425-e434. [doi: 10.1016/s2589-7500(20)30137-0]
45. Santosh KC. COVID-19 Prediction Models and Unexploited Data. J Med Syst 2020 Aug 13;44(9):170 [FREE Full text] [doi: 10.1007/s10916-020-01645-z] [Medline: 32794042]
46. Rajpal A, Rahimi L, Ismail-Beigi F. Factors leading to high morbidity and mortality of COVID-19 in patients with type 2 diabetes. J Diabetes 2020 Dec 02;12(12):895-908 [FREE Full text] [doi: 10.1111/1753-0407.13085] [Medline: 32671936]
47. Kostkova P. Disease surveillance data sharing for public health: the next ethical frontiers. Life Sci Soc Policy 2018 Jul 04;14(1):16 [FREE Full text] [doi: 10.1186/s40504-018-0078-x] [Medline: 29971516]

Abbreviations

NHS: National Health Service

Edited by A Mavragani; submitted 02.04.21; peer-reviewed by M Abdelhamid, M Jungkunz; comments to author 21.04.21; revised version received 10.05.22; accepted 22.05.22; published 07.07.22.

Please cite as:
Summers C, Griffiths F, Cave J, Panesar A
Understanding the Security and Privacy Concerns About the Use of Identifiable Health Data in the Context of the COVID-19 Pandemic: Survey Study of Public Attitudes Toward COVID-19 and Data-Sharing
JMIR Form Res 2022;6(7):e29337
URL: https://formative.jmir.org/2022/7/e29337
doi:10.2196/29337
PMID:35609306

Original Paper

Development of a Quality Management Model and Self-assessment Questionnaire for Hybrid Health Care: Concept Mapping Study

Rosian Tossaint-Schoenmakers[1,2,3], MSc; Marise J Kasteleyn[2,3], PhD; Anneloek Rauwerdink[4], MSc, MD; Niels Chavannes[2,3], MD, PhD; Sofie Willems[2,3], MD, PhD; Esther P W A Talboom-Kamp[2,3,5], MBA, MD, PhD

[1]Saltro Diagnostic Centre, Unilabs Netherlands, Utrecht, Netherlands

[2]National eHealth Living Lab, Leiden University Medical Centre, Leiden, Netherlands

[3]Department of Public Health and Primary Care, Leiden University Medical Centre, Leiden, Netherlands

[4]Department of Surgery, Gastroenterology and Metabolism, Amsterdam University Medical Centre, Amsterdam, Netherlands

[5]Unilabs Group, Geneve, Switzerland

Corresponding Author:
Rosian Tossaint-Schoenmakers, MSc
Saltro Diagnostic Centre, Unilabs Netherlands
Mississippidreef 83
Utrecht, 3565 CE
Netherlands
Phone: 31 655895361
Email: rtossaint@saltro.nl

Abstract

Background: Working with eHealth requires health care organizations to make structural changes in the way they work. Organizational structure and process must be adjusted to provide high-quality care. This study is a follow-up study of a systematic literature review on optimally organizing hybrid health care (eHealth and face to face) using the Donabedian Structure-Process-Outcome (SPO) framework to translate the findings into a modus operandi for health care organizations.

Objective: This study aimed to develop an SPO-based quality assessment model for organizing hybrid health care using an accompanying self-assessment questionnaire. Health care organizations can use this model and a questionnaire to manage and improve their hybrid health care.

Methods: Concept mapping was used to enrich and validate evidence-based knowledge from a literature review using practice-based knowledge from experts. First, brainstorming was conducted. The participants listed all the factors that contributed to the effective organization of hybrid health care and the associated outcomes. Data from the brainstorming phase were combined with data from the literature study, and duplicates were removed. Next, the participants rated the factors on importance and measurability and grouped them into clusters. Finally, using multivariate statistical analysis (multidimensional scaling and hierarchical cluster analysis) and group interpretation, an SPO-based quality management model and an accompanying questionnaire were constructed.

Results: All participants (n=39) were familiar with eHealth and were health care professionals, managers, researchers, patients, or eHealth suppliers. The brainstorming and literature review resulted in a list of 314 factors. After removing the duplicates, 78 factors remained. Using multivariate statistical analyses and group interpretations, a quality management model and questionnaire incorporating 8 clusters and 33 factors were developed. The 8 clusters included the following: Vision, strategy, and organization; Quality information technology infrastructure and systems; Quality eHealth application; Providing support to health care professionals; Skills, knowledge, and attitude of health care professionals; Attentiveness to the patient; Patient outcomes; and Learning system. The SPO categories were positioned as overarching themes to emphasize the interrelations between the clusters. Finally, a proposal was made to use the self-assessment questionnaire in practice, allowing measurement of the quality of each factor.

Conclusions: The quality of hybrid care is determined by organizational, technological, process, and personal factors. The 33 most important factors were clustered in a quality management model and self-assessment questionnaire called the Hybrid Health Care Quality Assessment. The model visualizes the interrelations between the factors. Using a questionnaire, each factor can be

assessed to determine how effectively it is organized and developed over time. Health care organizations can use the Hybrid Health Care Quality Assessment to identify improvement opportunities for solid and sustainable hybrid health care.

(JMIR Form Res 2022;6(7):e38683) doi:10.2196/38683

KEYWORDS

quality assessment; hybrid health care; blended health care; eHealth; digital health; structure; process; outcome; concept mapping

Introduction

Background

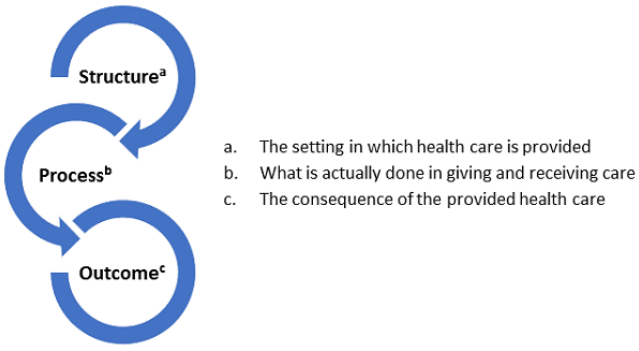
In recent years, the use of eHealth has expanded, encouraged by the increasing pressure on health care [1,2] and growing interest in patient empowerment [3,4]. On the one hand, an aging population and an increase in chronic diseases are causing a higher and more complex demand for health care. In addition, the COVID-19 pandemic has accelerated pressure on health care [5-8]. Therefore, innovations such as eHealth are required to maintain accessibility and high quality of health care [9-12]. On the other hand, digital health technologies have significantly accelerated patients' involvement [13-16]. In line with these developments, health care organizations have intensively integrated eHealth into traditional face-to-face consultations [17]. The combination of eHealth and face-to-face consultations can be defined as hybrid health care [18,19]. A few examples of hybrid health care are telemonitoring systems for patients with chronic diseases [20,21], web-based video coaching [22,23], and direct web-based access to medical records of patients [24,25], all of which are integrated into traditional health care.

Although health care organizations are increasingly providing hybrid health care, integrating eHealth into the daily care process is challenging. Working with hybrid health care requires organizations to change the way they work. The roles of health care providers and patients are changing, and the available resources are used differently [4,22,26,27]. Organizational structure and work processes must be adapted to ensure high-quality hybrid care [28-31]. Several studies have examined ways to promote eHealth adoption, such as increasing the adaptability of the technology or stakeholders' value [32,33]. However, it remains challenging to organize hybrid health care effectively and sustainably [17]. There is a need for further research on how hybrid health care can be improved to add value to patients and health care providers when they work with eHealth. Therefore, we recently performed a systematic literature review to optimally organize hybrid health care [17].

In the systematic literature review, the Donabedian Structure-Process-Outcome (SPO) framework was used to identify indicators related to the integration of eHealth into health care organizations [17,34-36] (Figure 1). According to Donabedian, health care quality is based on the aspects of these 3 categories and their relationships. The SPO framework and its categories are described in detail in a literature review [17].

In the literature review, we identified 111 potential indicators under the SPO categories that impact eHealth integration. The study demonstrated that 3 principles are important for successful integration. First, the patient's role must be centrally placed in the organization of hybrid care. Second, technology must be well attuned to the organizational structure and daily care process. Third, the deployment of human resources must be aligned with desired results [17].

Figure 1. Donabedian Structure-Process-Outcome framework.



Objectives

To translate the findings from the literature study into a modus operandi for health care organizations, we aimed to develop a model that can help health care organizations organize hybrid health care and identify improvement opportunities for a solid and sustainable integration of eHealth. To achieve this aim, the objectives of the concept mapping study included the following: (1) enrich and validate evidence-based knowledge from the literature review with practice-based knowledge from experts and (2) develop an SPO-based model for organizing hybrid health care with an accompanying self-assessment questionnaire.

Methods

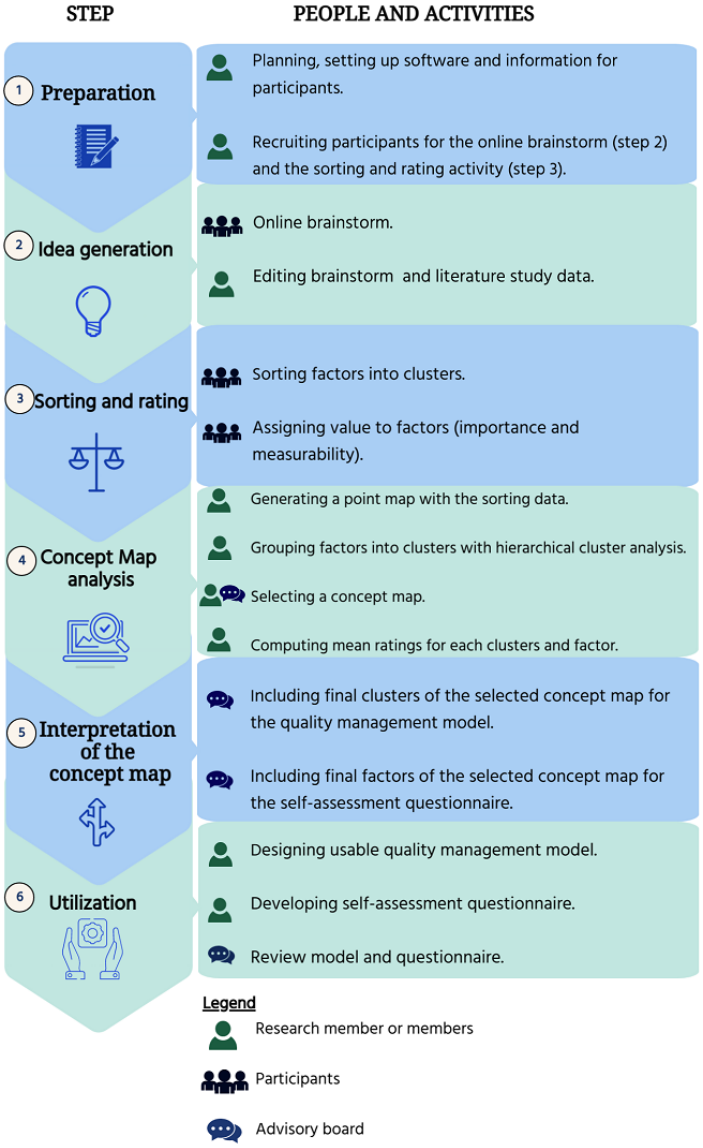
Concept Mapping

Concept mapping is a highly structured methodology for organizing ideas from different stakeholders and other data sources to produce a common framework for complex topics that can be used for evaluation or planning [37-40]. The method

integrates qualitative data collection with quantitative analysis to construct an interpretable pictorial view of different ideas and concepts and how these are interrelated [41,42]. Concept mapping has been used worldwide, for a diverse range of health care projects and studies to develop conceptual frameworks, as well as health and eHealth evaluations [43-49].

In this study, the 6-step concept mapping approach of Trochim and McLinden [42] was followed [49] to develop a usable, tailored, SPO-based quality management model for hybrid health care and an accompanying questionnaire. The six steps of concept mapping are as follows: (1) preparation, (2) idea generation, (3) sorting and rating, (4) concept mapping analysis, (5) map interpretation, and (6) utilization. Each step involves different activities leading to an output, which serves as an input for the next step. The steps and activities are explained in Figure 2 and in the paragraphs below. All the steps were supported by the GroupWisdom webtool [41,42].

Figure 2. Concept mapping steps and study activities.



Step 1: Preparation

Concept mapping is most effective when multiple stakeholders participate in all the steps of the concept mapping process [50]. There is no strict limitation to the number of participants, ranging from small groups of 8 to 15 people to groups of hundreds of participants [50]. For this study, participants with eHealth experience, those employed by health care organizations, and patients with eHealth experience were recruited. The amount or kind of eHealth experience, health care setting, or disease was not relevant for inclusion. The goal was to create a diverse group in which different experiences, perceptions, and viewpoints complemented each other. We aimed to include a mix of health care professionals, patient experts (patients and caregivers), managers, directors, project leaders, researchers, and eHealth suppliers.

Potential participants were approached to attend both brainstorming in step 2 and sorting and rating in step 3. Participants were invited via the research team's network, social media, and snowballing. Before agreeing to participate, participants received an information letter about the concept mapping method, the study's purpose, and the SPO framework. None of the potential participants were familiar with our

previous literature study results. A selected group was asked to participate in step 4 (concept mapping), step 5 (interpretation), and step 6 (utilization), which will be explained in the subsequent sections.

Step 2: Idea Generation

Web-Based Brainstorming

In step 2, data from the participants were collected and combined with data from the literature study. Idea generation with participants was organized by brainstorming. Brainstorming is the most common method used in concept mapping, and can be either group brainstorming or individual brainstorming [42]. In this study, web-based brainstorming was conducted by the participants. Participants received a link via email with instructions, giving them access to the web-based brainstorm program of the GroupWisdom webtool. Before starting the brainstorming session, informed consent was provided, and participant characteristics (age, eHealth experience, professional background, and work setting) were collected to generate general background information about the participants. When the brainstorming started session, the following instruction was presented: "Name all factors, which you believe contribute to effective organization of patient care with eHealth, and what the outcomes of this care should be. Keep the 'Structure-Process-Outcome' framework in mind."

For 23 days, the participants could list as many factors they considered essential contributors to effective hybrid health care. Participants could see each other's inputs and save their brainstorming results in the meantime. They received reminders after 10 and 15 days.

Editing Brainstorming and Literature Study Data

After closing the web-based brainstorming session, the brainstorming and literature study data were combined for sorting and rating. A manageable amount of data for sorting and rating is ideally ≤100 to prevent redundancy and a loss of participants' motivation [51,52]. To generate a final set of up to 100 factors, duplicates and factors that did not match the brainstorming instructions were removed. For this purpose, each factor was assessed independently by the authors, RT-S and ET-K. The assessments were compared, and disagreements were resolved by discussion between RT-S and ET-K. Next, RT-S edited the remaining factors for grammar and spelling.

Authors, MK and AR reviewed the editing process to check whether they would conclude the same selection and wording and made recommendations where appropriate. Finally, the set was entered into the GroupWisdom webtool, serving as an input for the sorting and rating activities.

Step 3: Sorting and Rating

At the beginning of step 3, the participants received instructions for the sorting and rating tasks. For the sorting task, the participants were asked to cluster the factors into self-created clusters and assign names to the clusters. The participants were instructed to keep the Donabedian SPO categories in mind while sorting each factor into self-created clusters. For the rating task, each participant was asked to rate each factor by relevancy on a 5-point Likert scale, ranging from 1 (*not important at all* or *not feasible to measure*) to 5 (*very important* or *very feasible to measure*) by answering the questions, "How important is this factor for effective patient care with eHealth?" and "How feasible to measure is this factor?"

The participants had the opportunity to sort and rate over 3 weeks. They could save their activities and return later and received reminders after 10 and 15 days. The sorting data were approved for concept mapping analysis for participants who completed 75% of the sorting activity and created at least three clusters [41]. The rating data were included when the participant rated at least one factor.

Step 4: Concept Mapping Analysis

Concept mapping analysis consisted of four main activities: (1) generating a point map with the sorting data, (2) grouping factors into clusters using hierarchical cluster analysis, (3) selecting a concept map from the hierarchical cluster analysis, and (4) computing average ratings for each factor and cluster of the selected concept map [50]. All computations were based on the concept mapping approach of Kane et al [53,54] and conducted using the GroupWisdom webtool.

Generating a Point Map With the Sorting Data

Data from the rating step were analyzed to create a point map [45,53,55,56]. A point map is a 2-dimensional point map, in which each point represents a factor [53]. The point map visually displayed the locations of all factors. Factors closer to each other on the point map were sorted together more frequently by the participants, whereas more distant factors on the map were sorted together less frequently [42,50,53]. The point map was constructed using a similarity matrix and multidimensional scaling algorithm. First, the similarity matrix indicated the number of times various factors were grouped together. Next, a multidimensional scaling algorithm plotted factors as points on a point map [42,54,55]. Subsequently, a stress value (0-1) was calculated, indicating the degree to which the distances on the point map fit the original similarity matrix [38,54]. The better the fit, the lower is the stress value.

Grouping Factors Into Clusters With Hierarchical Cluster Analysis

The point map provided the input for the hierarchical cluster analysis. The hierarchical cluster analysis grouped factors into clusters [44] using Ward algorithm [57]. The algorithm proposed several concept map solutions, where 2 clusters were merged at each following the proposed solution.

Selecting a Concept Map

From the proposed concept map solutions, a concept map that made sense for conceptualization was selected. There is no single correct number of clusters or mathematical decision criterion for selecting a concept map solution [38,56]. This study selected the number of clusters for the concept map by determining the range of the highest and lowest number of clusters. The range was the average number of clusters made by the participant and its SD.

Subsequently, the cluster solutions in this range were reviewed to select the cluster level by following the cluster tree in the Methods section of the studies by Trochim [53] and Kane et al

[54]. Finally, in a meeting, 2 authors (RT-S and ET-K) and 2 participants reviewed the merging of clusters, beginning with the highest number of clusters and moving to the lowest. The 2 study participants were asked to join this meeting because of their extensive experience with eHealth, daily care processes, research, operational management, and concept mapping.

After establishing the number of clusters in the concept map, each factor was reviewed for compatibility with the cluster and to determine whether it was appropriate to move the factor to a different cluster. A cluster and its content were appropriate for inclusion when they were considered essential and usable for the quality management model [53].

In addition, each cluster received a name and description based on the cluster names that emerged from the sorting activity.

Computing Mean Ratings for Each Cluster and Factor of the Selected Concept Map

After the cluster map was selected, the relationships between ratings were computed using pattern-match and Go-zones [42].

Pattern-match and its Pearson product-moment (*r* value) were calculated to compare how the clusters of the selected concept map were rated on importance and measurability. The pattern-match visualized the mean ratings of each cluster in a ladder graph, connecting lines between the mean ratings on importance and measurable of each cluster [50,57]. The *r* value represented the correlation strength between the 2 mean ratings of all clusters [50,57].

Finally, multiple Go-zones were computed: a Go-zone of the total point map and Go-zones per cluster of the selected concept map. Go-zone is a 4-quadrant graph with an x-y graph [50], visualizing the mean ranking results of each factor on the questions “How important is this factor” and “How feasible to measure is this factor.” The minimum and maximum values for each axis were the minimum and maximum average Likert scores, respectively. The upper-right quadrant is called the G*o-zone* because it shows factors rated above the mean for both importance and measurability [42,58]. The pattern-match and Go-zone showed how important and measurable each cluster and its factors were rated for quality assessment by the individual participants during the step, sorting and rating.

The selected concept map, with its calculation of importance and measurability for each cluster and factor, formed the basis of interpretation in the next step [53].

Step 5: Interpretation of the Concept Map

The selected concept map, with its pattern-match and Go-zones, was discussed with an advisory board. On the basis of the pattern-match and Go-zones, the advisory board decided which clusters and factors should be included in the quality management model and the accompanying questionnaire. The advisory board consisted of 4 study participants from the brainstorming and sorting step, of whom, 2 also participated in step 4, concept mapping analysis. The advisors were chosen because they could be future model users. In addition, all had extensive experience with eHealth, health care business, and as health care professionals (general practitioners, nurses, anesthetists, and clinical psychologists) in different health care settings.

The advisors voted individually on which clusters and factors of the selected concept map should be included in the quality management model and questionnaire to ensure usability. Using a web-based survey, the following questions were asked: “Which cluster should be included in the quality management model based on the mean cluster rating scores of the pattern matches? Please, specify your choice.” and “On which factors should the questionnaire give focus? Guide your choice by the Go-zones of each cluster and the Go-zone of the total point map. Please specify your choice.” The advisors could not see each other’s votes. By 75% (3/4) agreement or more, the concerned clusters and factors were operationalized in the quality assessment model and questionnaire. Where there was less agreement, the advisors viewed all responses, including the comments, and were asked to vote again. This process was repeated until a 75% consensus was reached. The web-based survey results were used as inputs to develop the quality management model and its questionnaire.

Step 6: Utilization

Quality Management Model

The remaining clusters and their positions in the selected concept map provided the blueprint for the quality management model. First, the excluded clusters and factors were removed from the concept map. Second, the concept map with the remaining clusters was used to produce a logic model. A logic model is a framework that visualizes the interrelations between the clusters in graphic form and is therefore valuable for quality evaluation [59]. The SPO framework [34,35] was used to identify logical interrelationships between the clusters. Accordingly, noticeable SPO connections between the clusters were drawn on the map by RT-S. A simplified version of the logic model was designed for clarity and readability. Authors SW, ET-K, and RT-S discussed the design of the quality management model to ensure the usability and clarity of the model.

Self-assessment Questionnaire

The questionnaire was drafted by RT-S with the remaining factors, taking the advisors’ comments into account. The questionnaire should give care organizations insight into the quality of hybrid care and how quality develops over time. On the one hand, the questionnaire must be easy to use and uniformly independent of the type of health care organization, type of eHealth, and disease. On the other hand, the questionnaire results must provide specific guidance to improve the quality of specific clusters and factors.

The concept model and questionnaire were submitted to the advisors for peer review of usability and clarity. Their comments were processed by RT-S, resulting in an improved draft. Finally, ET-K and SW peer reviewed the last draft to ensure that the representatives’ comments were implemented entirely in the quality management model and the related questionnaire.

Ethics Approval

Approval by an ethics committee was not needed because no intervention or trial has occurred in the sense that the research

participants were subjected to actions or had modes of behavior imposed on them [60].

Results

Participant Characteristics (Step 1)

A total of 39 people participated in this study. The participants had a mean age of 45.2 (SD 11.1) years and were mainly working at the family medicine clinic (12/39, 31%) or hospital (10/39, 26%) within a management function (16/39, 41%) or as a health care professional (14/39, 36%). A total of 59% (23/39) of the participants estimated their eHealth experience to be extensive. The 3 most commonly used eHealth tools were apps (37/147, 25.2% participants), web portals (35/147, 23.8% participants), and video communication (34/147, 23.1% participants). An overview of the participants' characteristics is shown in Table 1.

Of the 39 participants, 38 (97%) completed the brainstorming sessions. In all, 18% (7/38) of the participants dropped out after the brainstorming session, and a new participant joined the sorting and rating phase. In total, 79% (31/39) of the participants completed the sorting and rating phase (Figure 3).

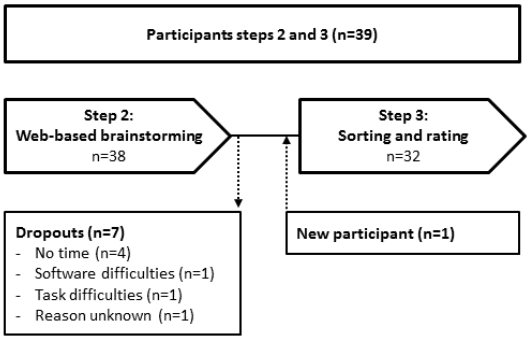
Table 1. Participant characteristics (N=39).

Variables	Values
Age (years), mean (SD)	45.2 (11.1)
Main work setting, n (%)	
Family medicine	12 (31)
Hospital	10 (26)
Mental health clinic	5 (13)
Nursing and residential care	5 (13)
eHealth supplier	4 (10)
Research institute	2 (5)
Patient experts (self-employed)	1 (3)
Main profession, n (%)[a]	
Manager, director, or project leader	16 (41)
Health care professional (eg, physician, nurse, therapist, or psychologist)	14 (36)
Patient expert (eg, patient or caregiver)	5 (13)
Researcher	3 (8)
Unknown	1 (3)
eHealth technology experience, n (%)[b]	
Apps	37 (25.2)
Web portals (eg, electronic health records or personal care records)	35 (23.8)
Video communication	34 (23.1)
Sensors and wearables	23 (15.6)
Artificial intelligence	13 (8.8)
Domotica and robotica	10 (6.8)
Estimated level of experience with eHealth, n (%)	
Extensive experience	23 (59)
Moderated experience	15 (38)
Limited experience	1 (3)

[a]Many participants had dual roles, from which they were asked to choose one role.

[b]Participants could select multiple answers.

Figure 3. Number of participants at steps 2 and 3.



Idea Generation (Step 2)

Brainstorming during idea generation resulted in a list of 203 factors. A total of 111 potential indicators were extracted from the literature study [17]. Both lists were aggregated, resulting in a list of 314 factors. Editing of the data led to a final list of 78 factors. These 78 factors served as inputs for the sorting and rating activity. The list of 78 factors is provided in Multimedia Appendix 1.

Sorting and Rating (Step 3)

The rating data of the 32 participants were included in this study. All factors received mean rating scores of >3.1, for both importance and measurability. The mean ratings on the questions, "How important is this factor for successful integration of eHealth?" and "How feasible to measure is this factor" are described in Multimedia Appendix 1.

The sorting data of 8 people were excluded, with the reason "less than 75% sorted" (n=4, 50%) or "sorted in two clusters" (n=4, 50%). The mean number of clusters of the approved data was 7 (SD 3.5) with a range of 3 to 15 clusters.

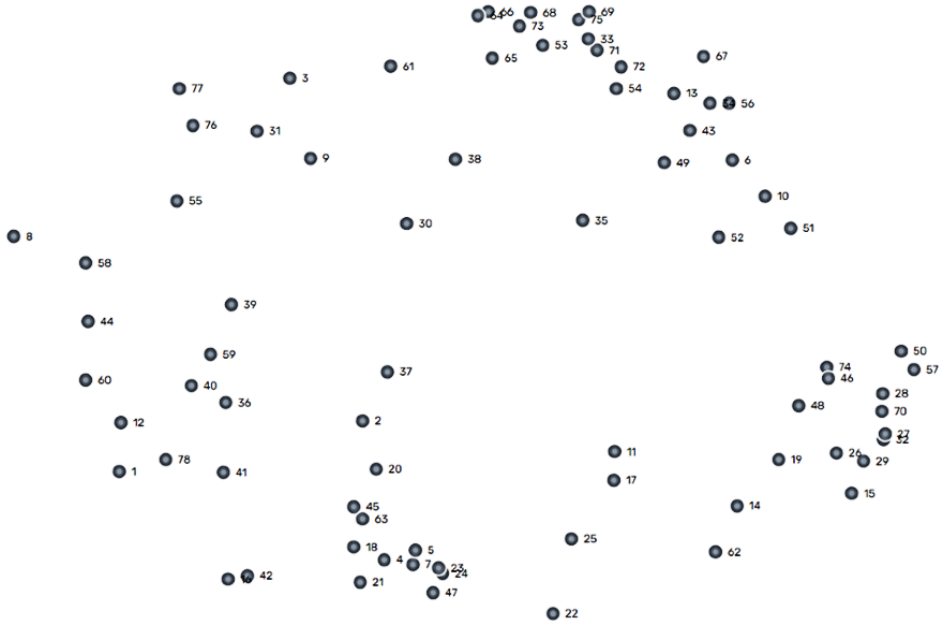
Concept Mapping Analysis (Step 4)

Visual Representation

The point map in Figure 4 shows how the 78 factors are related according to the sorting data. The point map had a stress value of 0.26, indicating that it had a good fit with the original similarity matrix [38,54].

The point map displays the locations of all factors that were frequently sorted closer together by the participants, whereas unrelated factors were plotted farther from each other. The number of points corresponds to the number of factors presented in Multimedia Appendix 1.

Figure 4. Point map.



Selecting the Concept Map

Concept map solutions ranging from 11-cluster to 3-cluster options were reviewed (mean 7, SD 3.5). The 9-cluster concept map was selected to make the most sense of conceptualization. A few factors (n=14) were unanimously replaced, leading to the concept map shown in Figure 5. Replaced factors and their reasons are presented in Multimedia Appendix 2. The 9 clusters were labeled and received a short description, as described in Table 2. The number of points corresponds to the number of factors presented in Multimedia Appendix 1. The clusters represent how the participants sorted the factors into self-created clusters using the proposed cluster labels.

Figure 5. Nine-cluster concept map. IT: information technology.

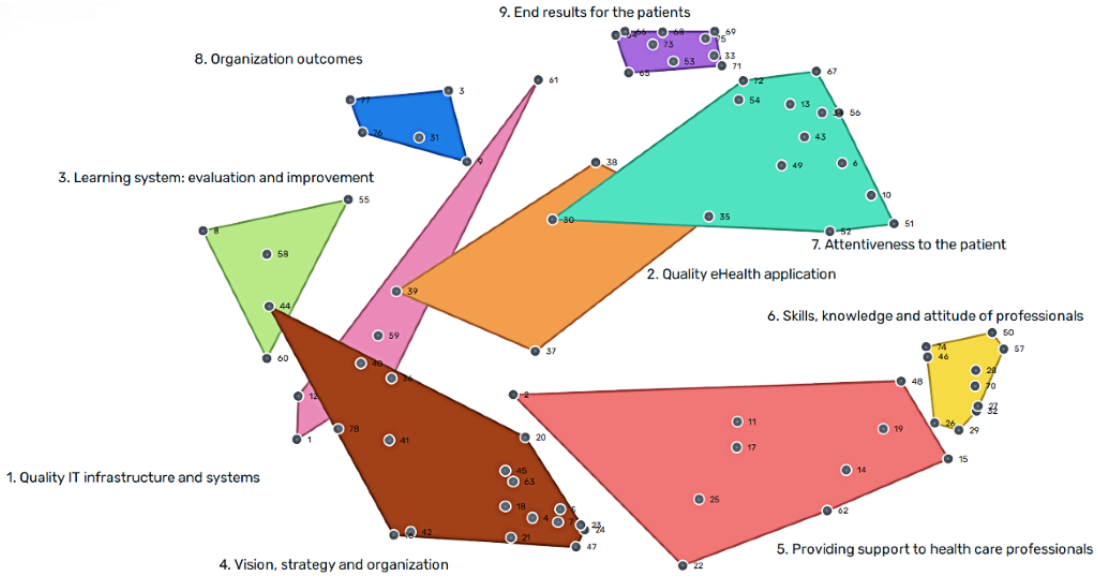


Table 2. Clusters labels and descriptions.

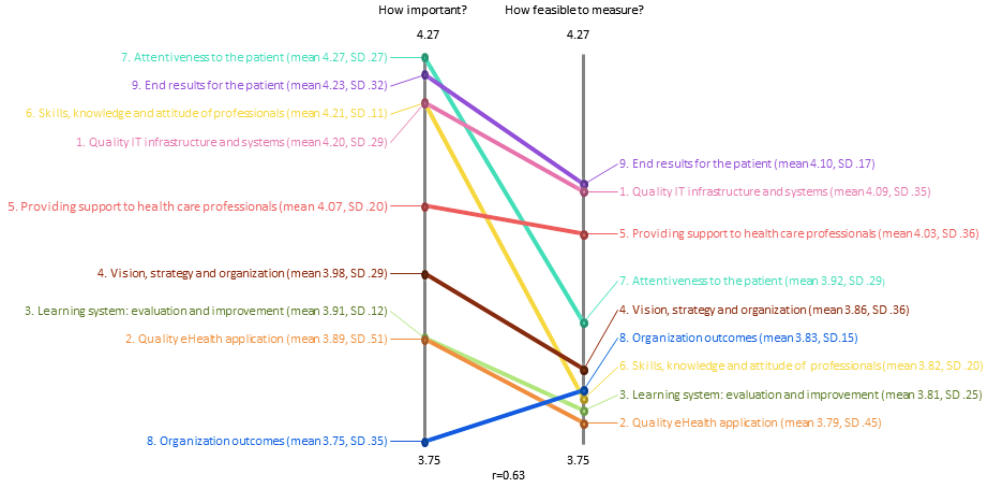
Cluster number[a]	Cluster label	Description	Included factors, n
1	Quality information technology infrastructure and systems	Conditions concerning technology, information technology systems, and data.	6
2	Quality eHealth application	Conditions concerning the eHealth application.	4
3	Learning system: evaluation and improvement	Evaluation and realignment with stakeholders and the patient care objectives for a continuous development.	4
4	Vision, strategy, and organization	Responsibilities of the health care organization concerning vision, strategy, policy, leadership, funding, and work process designs.	16
5	Providing support to health care professionals	Conditions arranged by the health care organization to encourage the use of eHealth among its health care professionals.	10
6	Skills, knowledge, and attitude of health care professionals	Health care professionals' ability to provide hybrid care.	10
7	Attentiveness to the patient	Organize the daily care process in line with the patient's needs, demand for care, and its capacity.	13
8	Organization outcomes	Outcomes for the health care organization; for example, quality health care provision and health care logistics.	5
9	End results for the patient	Outcomes for the patients; for example, health, added value, satisfaction, ownership, and convenience.	10

[a]Number corresponds with the number of the concerning cluster in Figure 5.

Mean Ratings for Each Cluster and Factor of the Selected Concept Map

The pattern-match showed that all clusters had a mean score between 3.75 and 4.27 on the importance and a mean score between 3.79 and 4.10 on measurability (Figure 6). The cluster with the highest mean score on importance was *Attentiveness to the patient* (mean 4.27, SD 0.27), and the cluster with the highest mean score on measurability was *End results for the patients* (mean 4.10, SD 0.17). On the contrary, the cluster with the lowest mean score on importance was *Organization outcomes* (mean 3.75, SD 0.36), whereas the cluster *Quality eHealth application* (mean 3.79, SD 0.45) had the lowest mean score on measurability. The *r* value was 0.63, indicating a predictable alignment between the rating of importance and the rating of measurability. The mean ratings of the factors and Go-zones per cluster are included in Multimedia Appendix 3.

Figure 6. Pattern-match between the cluster-mean scoring on importance and measurability, with Pearson product-moment. IT: information technology.



Interpretation of the Concept Map (Step 5)

The pattern-match and Go-zones were input to determine which clusters and factors of the selected concept map should be included in the quality management model and questionnaire. Decisions were made in 2 voting rounds. Of the 9 clusters, the cluster *Organization outcomes* was not included in the quality management model, based on the voting (3/4, 75%) of the advisors had doubts about including the cluster in the model) and after discussion with the research team. The factors included in the questionnaire concerned those placed in the Go-zone of the total point map or the Go-zone of the clusters. As a result, 8 clusters remained in the model and 33 factors in the questionnaire remained as a manageable utility for quality assessment (Textbox 1). Multimedia Appendix 3 presents the responses and comments of the advisory board during the voting rounds.

Textbox 1. The included clusters and factors.

Quality Information technology infrastructure and systems (1)

- Information technology architecture available within the health care organization (1).
- Back-up scenario during technical problems (12).

Quality eHealth application (2)

- The eHealth application is user-friendly (35).

Learning system: evaluation and improvement (3)

- Cocreation: eHealth is developed, implemented and redeveloped with different stakeholders (8).
- Monitoring and evaluation of service and treatment results (58).

Vision, strategy, and organization (4)

- Support the implementation and development of eHealth in the organization with good project management (4).
- Mobilizing funding for working with eHealth (16).
- Clear internal policies regarding the use of eHealth (18).
- Vision supported by the line, “Why are we doing this?” (21).
- Care delivery with eHealth complies with laws and regulations (41).
- Financial reimbursements for eHealth deployment (42).
- Redesign the current work process and review what contributes to the desired care outcomes (47).

Providing support toward health care professionals (5)

- Health care professionals have easy access to information technology resources; for example, device, internet, screen, or headset (2).
- Embedding eHealth in the daily practice of health care professionals (11).
- Training and supervision for health care professionals (15).
- Help desk for health care professionals (17).
- Information on the treatment with eHealth is clear and accessible to the health care professional (19).

Skills, knowledge, and attitude of health care professionals (6)

- Good balance between face to face and eHealth for the health care professional (46).
- The health care professional has confidence in the eHealth application (70).
- The health care professional is satisfied with working with eHealth (74).

Attentiveness to the patient (7)

- Clear communication to the patient about how care is offered (10).
- Personalized care, considering patient needs with regard to (deployment of) eHealth (13).
- The patient has easy access to the necessary information technology resources; for example, device, Internet, and so on (30).
- Patients receive practical support in using the eHealth application; for example, a help desk (49).
- The patient has confidence in the eHealth application (67).
- The patient has the flexibility to use eHealth wherever and whenever it is convenient (72).

End results for the patient (9)

- The patient can integrate the use of eHealth in their daily life (33).
- Treatment with eHealth has a positive influence on the patient’s health (64).
- Treatment with eHealth contributes to the patient’s self-reliance (65).
- The patient is satisfied (68).
- The patient has easy access to care (71).
- eHealth provides logistical convenience for the patient (73).

- eHealth has added value for the patient (75).

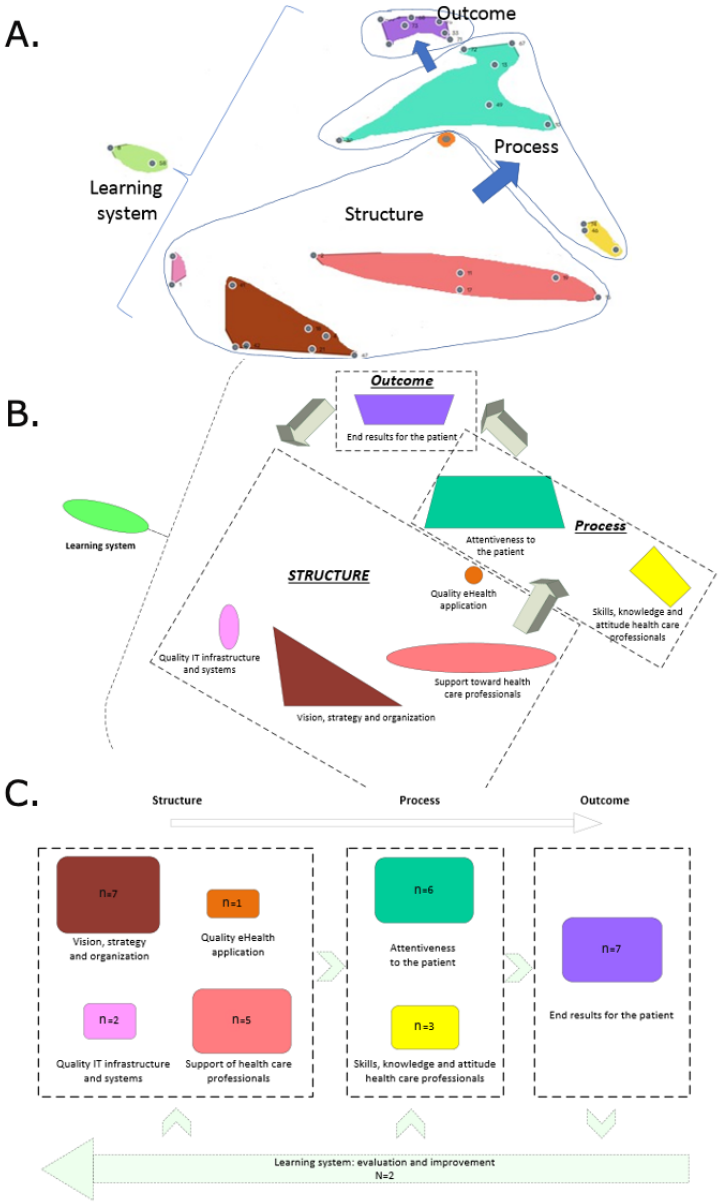
Utilization (Step 6)

Utilization Model

The clusters and factors excluded from the voting rounds were removed from the selected concept map. The remaining clusters (n=8) and their factors (n=33) led to nonoverlaying clusters on the concept map. Above the clusters, the SPO categories were positioned as overarching themes to emphasize the interrelations between the clusters. In addition, a complex cluster map can be simplified into a logic model. Figures 7A-C show the simplification of the model.

The overarching categories, *structure*, *process* and *outcomes* and the clusters' interconnections refer to the Donabedian SPO framework [34,35]. The cluster *Learning system* is visualized in the arrows with the dashed line. The numbers inside the clusters represent the number of factors included.

Figure 7. Simplification of the model. (A) Removing the excluded cluster and factors from the selected concept map and adding the overarching categories' structure, process, and outcome. (B) Drawing a logic interrelationship with structure, process, and outcome categories. (C) Simplification into a quality management model. IT: information technology.



Utilization Questionnaire

The remaining 33 factors were included in the questionnaire, where each factor can be measured on how effectively it is organized and developed over time. The advisory board noted that measuring the quality progress of hybrid health care is very important, in addition to learning and continuous improvement with stakeholders. Subsequently, the idea was to enrich the questionnaire with a quality progress tracker based on the plan-do-check-act (PDCA) cycles of Deming [61]. Incorporating the PDCA cycle makes it possible to assess the quality easily

and uniformly with tailored feedback for health care organizations. PDCA is a well-known cycle method for continuous improvement and quality measurement [61]. The PDCA cycles assess each factor's quality by measuring the extent to which *The objective is tangible?* (plan), *The plan is implemented?* (do), *To what extent is the plan realized?* (check), and *Providing feedback on the quality of the execution to make improvements* (act) [61]. Each factor can be monitored on the quality level of the PDCA cycles using a Likert score (0-10). A score of 0 means there is *no plan to improve the concerning factor*, and a score of 10 means *continue improvement with stakeholders*. The Likert scoring is based on the PDCA cycles and the 2 factors of the cluster Learning system, which include the following: (1) *Cocreation: eHealth is being developed and implemented with various stakeholders* and (2) *Monitoring and evaluation of service- and treatment outcomes*. Using the PDCA cycles in combination with a Likert score provides a health care organization insight into improvement possibilities for each factor or cluster.

Finally, the model and questionnaire obtained a more convenient workname Hybrid Health Care Quality Assessment (HHQA). The HHQA model and questionnaire with suggestions on how to use it are explained in Multimedia Appendix 4.

Discussion

Principal Findings

In this concept mapping study, we aimed to develop an SPO-based model and an accompanying self-assessment questionnaire for hybrid health care. By combining practice-based knowledge from eHealth users with an evidence-based literature review, we found that organizational, technological, and process and personal factors affect the quality of hybrid health care. Health care organizations must understand that these factors play a role in organizing hybrid health care and should be familiar with ways to improve them. The authors developed the HHQA, which can be used to systematically assess and improve the quality of hybrid health care.

The HHQA model includes 8 clusters. Cluster 1 (*Vision, strategy, and organization*) includes the responsibilities of the management to set the vision, strategy, policy, leadership, finance, and project management. Cluster 2 (*Quality information technology infrastructure and systems*) focuses on information technology infrastructure and back-up scenarios by information technology issues. Cluster 3 (*Quality eHealth application*) concerns the user-friendliness of the digital health application itself. Cluster 4 (*Providing support toward care professional*) and cluster 5 (*Skills, knowledge, and attitude of health care professionals*) include factors concerning health care providers. Cluster 4 focuses on factors that should be arranged for the individual health care professional by the care organization, and cluster 5 includes the responsibilities of the professional. The patient is central in cluster 6 (*Attentiveness to the patient*). This cluster contains the measurement of factors that allow patients to increase their self-management and consider the individual patient's needs. Patient centeredness is also reflected in cluster 7 (*Patient outcomes*), including factors such as patient's health outcomes, added value, satisfaction, ownership, and convenience. Finally, cluster 8 (*Learning system*), forms the relationship between the continued development of hybrid health care with stakeholders and health care provision objectives. The factors in cluster 8 provide insight into where alignment can be improved with other organizational criteria and actions, such as cost-benefit or capacity management.

The interdependencies of the clusters are logically expressed in the HHQA model because of the overarching categories of the Donabedian SPO framework. Moreover, according to eHealth users, clusters consist of the most important factors for the quality of hybrid health care. Using the questionnaire, each factor (33 in total) was measured to determine how effectively it was organized and developed over time. Subsequently, the main results of the questionnaire were shown at the cluster level. It was possible to zoom in on the relevant factors for each cluster.

Comparison With Literature

In our previous literature review [17], we concluded that the capabilities of patients, health care professionals, and technology play a crucial role in the quality of hybrid health care. We also concluded that offering hybrid health care requires adjusting the daily care process and appropriate process monitoring. The conclusions from the literature review are reflected in the HHQA clusters, namely, the patient's role is visible in the clusters *Attentiveness to the patient* and *Patient outcomes*; the health care professional's role is central in the clusters *Providing support toward health care professionals* and *Skills, knowledge, and attitude of professionals*; and technology is covered in the clusters *Quality information technology infrastructure and systems* and *Quality eHealth application*. The adjustment of the daily care processes is elaborated in the cluster *Vision, strategy, and organization.* Finally, monitoring is embedded in the cluster *Learning system* and the PDCA-progress tracker.

The 8 clusters of the HHQA model fit the 3 overarching categories of the Donabedian SPO framework. According to Donabedian [34], health care quality is based on aspects of these 3 categories and their relationships. The interaction between the categories can be bidirectional and is an "unbroken chain of antecedents, followed by intermediate ends, which are themselves the means to still further ends" [35]. Our research translated the complex interaction between the categories, structure, process, and outcome into user language.

The HHQA connects essential contributions to the quality of hybrid health care using a progress tracker. The relationship between quality contributors and continuous improvement also appears in the European Foundation for Quality Management Model (EFQM) [62,63]; nonadoption, abandonment, scale-up, spread, sustainability (NASSS) [32]; and the Consolidated Framework for Implementation Research (CFIR) [64,65]. All models approach the organizational structure, process, and outcomes with continuous improvement in a structured manner, but with different focus areas. For example, the EFQM is not specified for health care, in contrast to the NASSS and CFIR. The NASSS focuses on the adoption of technology and reduces implementation complexity, whereas the CFIR emphasizes on implementation in general. However, none of them have been

specified for quality assessment and improvement of hybrid health care.

Nevertheless, it is interesting to conduct a detailed examination of the assessment questionnaires of the EFQM and NASSS. The EFQM deployed the Results-Approach-Deployed-Assessment-Refinement (RADAR) method [66,67], a questionnaire to assess the quality improvement at each EFQM criteria, which incorporates the continued improvement circle. The assessment using the RADAR method is similar to the PDCA cycle in our questionnaire, as both monitor continuous quality improvement by completing the cycle plan-executing-monitoring and refining. However, the RADAR, similar to the EFQM model, is not specified for hybrid health care. In addition, the NASSS comes with a questionnaire to monitor the complexity of technology implementation in health care [68], but the focus is on project management instead of the hybrid health care process itself. Furthermore, there are other questionnaires measuring the quality of eHealth [69-72] or the quality of health care [73,74]. However, these questionnaires are concerned with the quality assessment of eHealth nationwide [68,70], the quality of a specific digital health application [70,72], or measuring the quality of a specific disease pathway [73,74]. To the best of our knowledge, HHQA is the first questionnaire measuring the quality of hybrid health care at an organizational level, taking the role of the patient, health care professionals, and technology into account, accompanied by an improvement progress tracker. Therefore, the authors recommend using the HHQA to measure and improve the quality of hybrid health care.

Strengths and Limitations

This study has several strengths. First, the HHQA was developed in cocreation with stakeholders who are direct users of eHealth. Therefore, the HHQA content was drawn from inside the health care system itself and not conceived or imposed outside the health care organizations. Second, stakeholders choose the included clusters and factors. The researcher only played a facilitating role. Consequently, the clusters and factors accurately reflect stakeholders' views and values, expressed in their own words and visual representations. Third, the stakeholder group was diverse and consisted of representatives of health care professionals, patients, managers, researchers, and eHealth designers. Nevertheless, the stress value of the point map shows that the stakeholders' outcomes are highly compatible. Therefore, the study results are likely to be generalizable to everyday practices. Fourth, the model and questionnaire were developed by combining scientific and practice-based knowledge. Together, these strengths result in important factors for effective hybrid health care covering different users' needs and organization requirements.

Our study had some limitations. First, the questionnaire had not yet been tested in health care organizations. This will be conducted in a follow-up study. Although eHealth users from different health care organizations have reviewed the model and questionnaire, the model and questionnaire may still be too abstract for daily practice, as is often the case in scientific research [75-77]. A follow-up study could provide concrete recommendations on how to use the HHQA. Second, it is conceivable that other factors and clusters could be included in other participants and health care environments. We attempted to overcome this problem by creating diverse groups of participants with different backgrounds, various eHealth experiences, and different kinds of health care settings. In addition, combining idea generation through brainstorming with results from a systematic literature review reduces the risk of bias. Third, based on the analysis of the concept mapping phase, 14 factors were moved to other clusters. However, some of these factors were moved far across the map, which was not entirely in line with the spirit of group concept mapping. Nevertheless, we deemed it necessary to move these factors for substantive reasons. Fourth, the advisory group consisted of 4 participants. We wanted to avoid overquestioning the participants and, therefore, deliberately selected a group of delegates who reflected on the diversity among the participants and who also had experience with quality management and concept mapping. Combined with in-depth preparation and discussion among the research groups, this appeared to be the most feasible solution.

Finally, it is worth pointing out that the HHQA gives a first general impression of improvement, as there is much to be gained in taking the role of the patient, health care professionals, and used technology into account [17]. Furthermore, the authors will continue with follow-up research and warm-heartedly welcome repetition of the study to improve the HHQA, taking into account the different users and health care environments.

Conclusions

This study developed a quality management model and an accompanying self-assessment questionnaire tailored for hybrid health care, the HHQA. A quality model for hybrid care is indispensable for effectively integrating eHealth into regular care and delivering high-quality health care. The HHQA covers all relevant aspects for the assessment and sustainable improvement of hybrid health care and the interrelations of eHealth with organizational, technical, and human factors. The next step is to validate and apply the HHQA model and questionnaire in practice.

Acknowledgments

The authors thank the experts who participated in the brainstorming and sorting phase, Arjen Huizinga (MiGuide), Bart van Pinxteren (eHuisarstenkompas; Huisartsen Oog in Al; Huisartsen Utrecht Stad; Nederlands Huisartsen Genootschap; Spindok); Bart Timmers (Groepspraktijk Huisartsen Bergh), Beverly Rose (Vereniging van Ervaringsdeskundigen), Corine van Barneveld (Saltro), Caroline Meijer (Leiden University Medical Centre), Erwin van Boxtel (Thebe), Folkert de Winter (Sint Antonius Ziekenhuis), Geert-Jan van Hal (self-employed), Hans in het Veen (Franciscus Gasthuis and Vlietland), Heleen Krabben (Medicinemen), Ineke Kamp, Irvin Talboom (Huisartsen Zorggroep Breda), Jan Frans Mutsaerts (Het Huisartsenteam; Familiedokters), Jeanette Ploeger (Minddistrict), Joris Arts (DiSofa; Geestelijke GezondheidsZorg Noord Holland Noord), Joyce

Bierbooms (Geestelijke GezondheidsZorg Eindhoven), Judie Knol (de Neckar), Kim Brons (Leiden University Medical Centre; Process in progress), Leonoor van Dam van Isselt (Leiden University Medical Centre, Pieter van Foreest; Vilente), Maarten Ellenbroek (Topaz), Marijke de Vries (Leiden University Medical Centre), Marjolein Eldenhorst (Leiden University Medical Centre), Marjon Peters (Insight4u), Maryse Spapens (Zorgkompaz), Melchior Nierman (Atal Medial), Mieke Klerkx (self-employed), Lya van der Veen, Pascale Schure (Huisartsen Linschoten; Unilabs Group), Paul Haarkamp (Carinova), Ryan Esser (Incluzio), Stephanie Wouthuis (Gezondheidscentrum De Boog; Huisartsenpraktijk Nijdam; Saltro), Stijn de Ruijter (doccs huisartsenpraktijk), Wim Green (Leiden University Medical Centre).

The authors extend special thanks to Cynthia Hallensleben (Gezondheidscentrum Spoorlaan; Leiden University Medical Centre), Fred de Boer (Leiden University Medical Centre), Nathan Bachran (Geestelijke GezondheidsZorg Oost Brabant), and Tobias Bonten (Huisartsenpraktijk L Broek; Leiden University Medical Centre) for their advice on the development of the model and questionnaire, to Marc Smelik (IE Business School) for reviewing the manuscript textually, and to Léon Tossaint (former chief executive officer of the European Foundation for Quality Management Model) for his explanation of the Results-Approach-Deployed-Assessment-Refinement method of the European Foundation for Quality Management Model.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mean (SD) rating scores of clusters and factors.
[DOCX File , 38 KB - formative_v6i7e38683_app1.docx]

Multimedia Appendix 2
Relocation factors and their reasons.
[DOCX File , 16 KB - formative_v6i7e38683_app2.docx]

Multimedia Appendix 3
Results voting “which clusters and factors to include” and given comments.
[DOCX File , 292 KB - formative_v6i7e38683_app3.docx]

Multimedia Appendix 4
Suggestion utilization Hybrid Health Care Quality Assessment questionnaire.
[XLSX File (Microsoft Excel File), 84 KB - formative_v6i7e38683_app4.xlsx]

References

1. van der Kleij RM, Kasteleyn MJ, Meijer E, Bonten TN, Houwink EJ, Teichert M, et al. SERIES: eHealth in primary care. Part 1: concepts, conditions and challenges. Eur J Gen Pract 2019 Oct;25(4):179-189. [doi: 10.1080/13814788.2019.1658190] [Medline: 31597502]
2. Nijland N. Grounding eHealth: towards a holistic framework for sustainable eHealth technologies. University of Twente. 2011 Jan 21. URL: https://research.utwente.nl/en/publications/grounding-ehealth-towards-a-holistic-framework-for-sustainable-eh [accessed 2020-07-30]
3. Hibbard JH, Mahoney ER, Stock R, Tusler M. Do increases in patient activation result in improved self-management behaviors? Health Serv Res 2007 Aug;42(4):1443-1463. [doi: 10.1111/j.1475-6773.2006.00669.x] [Medline: 17610432]
4. Boers SN, Jongsma KR, Lucivero F, Aardoom J, Büchner FL, de Vries M, et al. SERIES: eHealth in primary care. Part 2: exploring the ethical implications of its application in primary care practice. Eur J Gen Pract 2020 Dec;26(1):26-32. [doi: 10.1080/13814788.2019.1678958] [Medline: 31663394]
5. van Hattem NE, Silven AV, Bonten TN, Chavannes NH. COVID-19's impact on the future of digital health technology in primary care. Fam Pract 2021 Nov 24;38(6):845-847. [doi: 10.1093/fampra/cmab081] [Medline: 34268563]
6. Golinelli D, Boetto E, Carullo G, Nuzzolese AG, Landini MP, Fantini MP. Adoption of digital technologies in health care during the COVID-19 pandemic: systematic review of early scientific literature. J Med Internet Res 2020 Nov 06;22(11):e22280 [FREE Full text] [doi: 10.2196/22280] [Medline: 33079693]
7. Thulesius H. Increased importance of digital medicine and eHealth during the Covid-19 pandemic. Scand J Prim Health Care 2020 Jun;38(2):105-106. [doi: 10.1080/02813432.2020.1770466] [Medline: 32484725]
8. Health at a Glance 2021. Organisation for Economic Cooperation and Development. 2021. URL: https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2021_ae3016b9-en [accessed 2022-03-08]
9. Zorgkeuzes in Kaart 2020: Analyse van beleidsopties van politieke partijen voor de zorg. Centraal Planbureau. 2020 Jul 24. URL: https://www.cpb.nl/zorgkeuzes-in-kaart-2020 [accessed 2021-11-25]

10. Mann DM, Chokshi SK, Kushniruk A. Bridging the gap between academic research and pragmatic needs in usability: a hybrid approach to usability evaluation of health care information systems. JMIR Hum Factors 2018 Nov 28;5(4):e10721 [FREE Full text] [doi: 10.2196/10721] [Medline: 30487119]
11. Heckemann B, Wolf A, Ali L, Sonntag SM, Ekman I. Discovering untapped relationship potential with patients in telehealth: a qualitative interview study. BMJ Open 2016 Mar 02;6(3):e009750 [FREE Full text] [doi: 10.1136/bmjopen-2015-009750] [Medline: 26936904]
12. Aardoom JJ, van Deursen L, Rompelberg CJ, Standaar LM, Suijkerbuijk AW, van Tuyl LH, et al. Indicatoren E-healthmonitor 2021-2023 en doelstellingen voor e-health. Rijksinstituut voor Volksgezondheid en Milieu. 2021. URL: https://rivm.openrepository.com/handle/10029/624864 [accessed 2022-03-09]
13. Snyder H, Engström J. The antecedents, forms and consequences of patient involvement: a narrative review of the literature. Int J Nurs Stud 2016 Jan;53:351-378. [doi: 10.1016/j.ijnurstu.2015.09.008] [Medline: 26602069]
14. Bentvelsen RG, van der Vaart R, Veldkamp KE, Chavannes NH. Systematic development of an mHealth app to prevent healthcare-associated infections by involving patients: 'Participatient'. Clin eHealth 2021;4:37-44 [FREE Full text] [doi: 10.1016/j.ceh.2021.03.001]
15. Bruce CR, Harrison P, Nisar T, Giammattei C, Tan NM, Bliven C, et al. Assessing the impact of patient-facing mobile health technology on patient outcomes: retrospective observational cohort study. JMIR Mhealth Uhealth 2020 Jun 26;8(6):e19333 [FREE Full text] [doi: 10.2196/19333] [Medline: 32589161]
16. Matamala-Gomez M, Maisto M, Montana JI, Mavrodiev PA, Baglio F, Rossetto F, et al. The role of engagement in teleneurorehabilitation: a systematic review. Front Neurol 2020 May 6;11:354 [FREE Full text] [doi: 10.3389/fneur.2020.00354] [Medline: 32435227]
17. Tossaint-Schoenmakers R, Versluis A, Chavannes N, Talboom-Kamp E, Kasteleyn M. The challenge of integrating eHealth into health care: systematic literature review of the Donabedian model of structure, process, and outcome. J Med Internet Res 2021 May 10;23(5):e27180 [FREE Full text] [doi: 10.2196/27180] [Medline: 33970123]
18. Chan SR, Torous J, Hinton L, Yellowlees P. Mobile tele-mental health: increasing applications and a move to hybrid models of care. Healthcare (Basel) 2014 May 06;2(2):220-233 [FREE Full text] [doi: 10.3390/healthcare2020220] [Medline: 27429272]
19. Hughes MC, Gorman JM, Ren Y, Khalid S, Clayton C. Increasing access to rural mental health care using hybrid care that includes telepsychiatry. J Rural Ment Health 2019 Jan;43(1):30-37. [doi: 10.1037/rmh0000110]
20. van Buul AR, Derksen C, Hoedemaker O, van Dijk O, Chavannes NH, Kasteleyn MJ. eHealth program to reduce hospitalizations due to acute exacerbation of chronic obstructive pulmonary disease: retrospective study. JMIR Form Res 2021 Mar 18;5(3):e24726 [FREE Full text] [doi: 10.2196/24726] [Medline: 33734091]
21. Dijkstra A, Heida A, van Rheenen PF. Exploring the challenges of implementing a Web-based telemonitoring strategy for teenagers with inflammatory bowel disease: empirical case study. J Med Internet Res 2019 Mar 29;21(3):e11761 [FREE Full text] [doi: 10.2196/11761] [Medline: 30924785]
22. Hinman RS, Nelligan RK, Bennell KL, Delany C. "Sounds a bit crazy, but it was almost more personal:" a qualitative study of patient and clinician experiences of physical therapist-prescribed exercise for knee osteoarthritis via Skype. Arthritis Care Res (Hoboken) 2017 Dec;69(12):1834-1844. [doi: 10.1002/acr.23218] [Medline: 28217864]
23. Hadjistavropoulos HD, Nugent MM, Dirkse D, Pugh N. Implementation of Internet-delivered cognitive behavior therapy within community mental health clinics: a process evaluation using the consolidated framework for implementation research. BMC Psychiatry 2017 Sep 12;17(1):331 [FREE Full text] [doi: 10.1186/s12888-017-1496-7] [Medline: 28899365]
24. Talboom-Kamp E, Tossaint-Schoenmakers R, Goedhart A, Versluis A, Kasteleyn M. Patients' attitudes toward an online patient portal for communicating laboratory test results: real-world study using the eHealth impact questionnaire. JMIR Form Res 2020 Mar 04;4(3):e17060 [FREE Full text] [doi: 10.2196/17060] [Medline: 32024632]
25. Tossaint-Schoenmakers R, Kasteleyn M, Goedhart A, Versluis A, Talboom-Kamp E. The impact of patient characteristics on their attitudes toward an online patient portal for communicating laboratory test results: real-world study. JMIR Form Res 2021 Dec 17;5(12):e25498 [FREE Full text] [doi: 10.2196/25498] [Medline: 34927593]
26. Chavennes NH. eHealth in Disease Management: doel of tool? Leiden University. 2015. URL: https://scholarlypublications.universiteitleiden.nl/handle/1887/51560 [accessed 2021-04-29]
27. Mitchell M, Getchell M, Nkaka M, Msellemu D, Van Esch J, Hedt-Gauthier B. Perceived improvement in integrated management of childhood illness implementation through use of mobile technology: qualitative evidence from a pilot study in Tanzania. J Health Commun 2012;17 Suppl 1:118-127. [doi: 10.1080/10810730.2011.649105] [Medline: 22548605]
28. Budhwani S, Fujioka JK, Chu C, Baranek H, Pus L, Wasserman L, et al. Delivering mental health care virtually during the COVID-19 pandemic: qualitative evaluation of provider experiences in a scaled context. JMIR Form Res 2021 Sep 21;5(9):e30280 [FREE Full text] [doi: 10.2196/30280] [Medline: 34406967]
29. Swinkels IC, Huygens MW, Schoenmakers TM, Oude Nijeweme-D'Hollosy W, van Velsen L, Vermeulen J, et al. Lessons learned from a living lab on the broad adoption of eHealth in primary health care. J Med Internet Res 2018 Mar 29;20(3):e83 [FREE Full text] [doi: 10.2196/jmir.9110] [Medline: 29599108]

30. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
31. Granja C, Janssen W, Johansen MA. Factors determining the success and failure of eHealth interventions: systematic review of the literature. J Med Internet Res 2018 May 01;20(5):e10235 [FREE Full text] [doi: 10.2196/10235] [Medline: 29716883]
32. 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]
33. 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. [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
34. Donabedian A. The quality of care. How can it be assessed? JAMA 1988 Sep 23;260(12):1743-1748. [doi: 10.1001/jama.260.12.1743] [Medline: 3045356]
35. Donabedian A. Evaluating the quality of medical care. 1966. Milbank Q 2005;83(4):691-729. [doi: 10.1111/j.1468-0009.2005.00397.x] [Medline: 16279964]
36. Rademakers J, Delnoij D, de Boer D. Structure, process or outcome: which contributes most to patients' overall assessment of healthcare quality? BMJ Qual Saf 2011 Apr;20(4):326-331. [doi: 10.1136/bmjqs.2010.042358] [Medline: 21339310]
37. Kane M, Trochim WM. Using concept mapping in planning. In: Kane M, Trochim WM, editors. Concept Mapping for Planning and Evaluation. Thousand Oaks, CA, USA: Sage Publications; 2011:135-156.
38. Rosas SR, Kane M. Quality and rigor of the concept mapping methodology: a pooled study analysis. Eval Program Plann 2012 May;35(2):236-245. [doi: 10.1016/j.evalprogplan.2011.10.003] [Medline: 22221889]
39. Trochim WM. Concept mapping: soft science or hard art? Eval Program Plann 1989 Jan;12(1):87-110. [doi: 10.1016/0149-7189(89)90027-x]
40. van Bon-Martens MJ, Achterberg PW, van de Goor IA, van Oers HA. Towards quality criteria for regional public health reporting: concept mapping with Dutch experts. Eur J Public Health 2012 Jun;22(3):337-342. [doi: 10.1093/eurpub/ckr016] [Medline: 21398660]
41. Groupwisdom. URL: https://groupwisdom.com [accessed 2021-03-08]
42. Trochim WM, McLinden D. Introduction to a special issue on concept mapping. Eval Program Plann 2017 Feb;60:166-175. [doi: 10.1016/j.evalprogplan.2016.10.006] [Medline: 27780609]
43. Minkman M, Ahaus K, Fabbricotti I, Nabitz U, Huijsman R. A quality management model for integrated care: results of a Delphi and Concept Mapping study. Int J Qual Health Care 2009 Feb;21(1):66-75. [doi: 10.1093/intqhc/mzn048] [Medline: 18945745]
44. Bonten TN, Rauwerdink A, Wyatt JC, Kasteleyn MJ, Witkamp L, Riper H, EHealth Evaluation Research Group. Online guide for electronic health evaluation approaches: systematic scoping review and concept mapping study. J Med Internet Res 2020 Aug 12;22(8):e17774 [FREE Full text] [doi: 10.2196/17774] [Medline: 32784173]
45. Rauwerdink A, Kasteleyn MJ, Haafkens JA, Chavannes NH, Schijven MP, steering committee, of the Citrien fund program eHealth. A national eHealth vision developed by University Medical Centres: a concept mapping study. Int J Med Inform 2020 Jan;133:104032. [doi: 10.1016/j.ijmedinf.2019.104032] [Medline: 31778886]
46. van Engen-Verheul M, Peek N, Vromen T, Jaspers M, de Keizer N. How to use concept mapping to identify barriers and facilitators of an electronic quality improvement intervention. Stud Health Technol Inform 2015;210:110-114. [Medline: 25991112]
47. Svobodova I, Filakovska Bobakova D, Bosakova L, Dankulincova Veselska Z. How to improve access to health care for Roma living in social exclusion: a concept mapping study. Int J Equity Health 2021 Feb 12;20(1):61 [FREE Full text] [doi: 10.1186/s12939-021-01396-4] [Medline: 33579295]
48. Hargett CW, Doty JP, Hauck JN, Webb AM, Cook SH, Tsipis NE, et al. Developing a model for effective leadership in healthcare: a concept mapping approach. J Healthc Leadersh 2017 Aug 28;9:69-78 [FREE Full text] [doi: 10.2147/JHL.S141664] [Medline: 29355249]
49. van Bon-Martens MJ, van de Goor LA, Holsappel JC, Kuunders TJ, Jacobs-van der Bruggen MA, te Brake JH, et al. Concept mapping as a promising method to bring practice into science. Public Health 2014 Jun;128(6):504-514. [doi: 10.1016/j.puhe.2014.04.002] [Medline: 24923995]
50. Kane M, Trochim WM. An introduction to concept mapping. In: Kane M, Trochim WM, editors. Concept Mapping for Planning and Evaluation. Thousand Oaks, CA, USA: Sage Publications; 2011:1-26.
51. Kane M, Trochim WM. Preparing for concept mapping. In: Kane M, Trochim WM, editors. Concept Mapping for Planning and Evaluation. Thousand Oaks, CA, USA: Sage Publications; 2011:28-48.
52. Kane M, Trochim WM. Generating the ideas. In: Kane M, Trochim WM, editors. Concept Mapping for Planning and Evaluation. Thousand Oaks, CA, USA: Sage Publications; 2011:49-66.
53. Trochim WM. An introduction to concept mapping for planning and evaluation. Eval Program Plann 1989 Jan;12(1):1-16 [FREE Full text] [doi: 10.1016/0149-7189(89)90016-5]

54. Kane M, Trochim WM. Concept mapping analysis. In: Kane M, Trochim WM, editors. Concept Mapping for Planning and Evaluation. Thousand Oaks, CA, USA: Sage Publications; 2011:87-110.
55. Group Concept Mapping Steps. Groupwisdom. URL: https://groupwisdom.com/GCMRG#GCM [accessed 2021-03-08]
56. Kane M, Trochim WM. Concept Mapping for Planning and Evaluation. Thousand Oaks, CA, USA: Sage Publications; 2007.
57. Kane M, Trochim WM. Interpreting the maps. In: Kane M, Trochim WM, editors. Concept Mapping for Planning and Evaluation. Thousand Oaks, CA, USA: Sage Publications; 2011:112-134.
58. Trochim W, Kane M. Concept mapping: an introduction to structured conceptualization in health care. Int J Qual Health Care 2005 Jun;17(3):187-191. [doi: 10.1093/intqhc/mzi038] [Medline: 15872026]
59. Kane M, Trochim WM. Using concept mapping in evaluation. In: Kane M, Trochim WM, editors. Concept Mapping for Planning and Evaluation. Thousand Oaks, CA, USA: Sage Publications; 2011:157-174.
60. Medical Research Involving Human Subjects Act (WMO). Central Committee on Research Involving Human Subjects. URL: https://english.ccmo.nl/investigators/legal-framework-for-medical-scientific-research/laws/medical-research-involving-human-subjects-act-wmo [accessed 2021-11-04]
61. Harteloh PP, Casparie AF. Kwaliteit van zorg. Van een zorginhoudelijke benadering naar een bedrijfskundige aanpak. Utrecht, The Netherlands: Elsevier; Jul 1, 1994.
62. Fonseca L, Amaral A, Oliveira J. Quality 4.0: the EFQM 2020 model and industry 4.0 relationships and implications. Sustainability 2021 Mar 12;13(6):3107. [doi: 10.3390/su13063107]
63. European Foundation for Quality Management. URL: https://www.efqm.org/ [accessed 2022-03-04]
64. Sarkies M, Long JC, Pomare C, Wu W, Clay-Williams R, Nguyen HM, et al. Avoiding unnecessary hospitalisation for patients with chronic conditions: a systematic review of implementation determinants for hospital avoidance programmes. Implement Sci 2020 Oct 21;15(1):91 [FREE Full text] [doi: 10.1186/s13012-020-01049-0] [Medline: 33087147]
65. Versluis A, van Luenen S, Meijer E, Honkoop PJ, Pinnock H, Mohr DC, et al. SERIES: eHealth in primary care. Part 4: addressing the challenges of implementation. Eur J Gen Pract 2020 Dec;26(1):140-145. [doi: 10.1080/13814788.2020.1826431] [Medline: 33025820]
66. European Foundation for Quality Management. The EFQM Excellence Model: Large Company, Operational and Business Unit Version. Brussels, Belgium: European Foundation for Quality Management; 2003.
67. Sokovic M, Pavletic D, Pipan MK. Quality improvement methodologies – PDCA Cycle, RADAR Matrix, DMAIC and DFSS. J Achiev Mater Manuf Eng 2010;43(1):476-483 [FREE Full text]
68. Greenhalgh T, Maylor H, Shaw S, Wherton J, Papoutsi C, Betton V, et al. The NASSS-CAT tools for understanding, guiding, monitoring, and researching technology implementation projects in health and social care: protocol for an evaluation study in real-world settings. JMIR Res Protoc 2020 May 13;9(5):e16861 [FREE Full text] [doi: 10.2196/16861] [Medline: 32401224]
69. Mousavi SM, Takian A, Tara M. Design and validity of a questionnaire to assess national eHealth architecture (NEHA): a study protocol. BMJ Open 2018 Dec 22;8(12):e022885 [FREE Full text] [doi: 10.1136/bmjopen-2018-022885] [Medline: 30580265]
70. Vis C, Bührmann L, Riper H, Ossebaard HC. Health technology assessment frameworks for eHealth: a systematic review. Int J Technol Assess Health Care 2020 Jun;36(3):204-216. [doi: 10.1017/S026646232000015X] [Medline: 32297588]
71. Currie WL. TEMPEST: an integrative model for health technology assessment. Health Policy Technol 2012 Mar;1(1):35-49. [doi: 10.1016/j.hlpt.2012.01.004]
72. ISO/TS 82304-2:2021 – Health software — Part 2: Health and wellness apps — Quality and reliability. International Organization for Standardization. 2021. URL: https://www.iso.org/standard/78182.html [accessed 2022-03-17]
73. Campmans-Kuijpers MJ, Lemmens LC, Baan CA, Gorter KJ, Groothuis J, van Vuure KH, et al. Defining and improving quality management in Dutch diabetes care groups and outpatient clinics: design of the study. BMC Health Serv Res 2013 Apr 05;13:129 [FREE Full text] [doi: 10.1186/1472-6963-13-129] [Medline: 23561032]
74. Zonneveld N, Vat LE, Vlek H, Minkman MM. The development of integrated diabetes care in the Netherlands: a multiplayer self-assessment analysis. BMC Health Serv Res 2017 Mar 21;17(1):219 [FREE Full text] [doi: 10.1186/s12913-017-2167-6] [Medline: 28320415]
75. Banks GC, Barnes CM, Jiang K. Changing the conversation on the science–practice gap: an adherence-based approach. J Manag 2021 Feb 22;47(6):1347-1356. [doi: 10.1177/0149206321993546]
76. Wandersman A. Community science: bridging the gap between science and practice with community-centered models. Am J Commun Psychol 2003 Jun;31(3-4):227-242. [doi: 10.1023/a:1023954503247]
77. Disler RT, Gallagher RD, Davidson PM. Factors influencing self-management in chronic obstructive pulmonary disease: an integrative review. Int J Nurs Stud 2012 Feb;49(2):230-242. [doi: 10.1016/j.ijnurstu.2011.11.005] [Medline: 22154095]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
EFQM: European Foundation for Quality Management Model

HHQA: Hybrid Health Care Quality Assessment
NASSS: nonadoption, abandonment, scale-up, spread, sustainability
PDCA: plan-do-check-act
RADAR: Results-Approach-Deployed-Assessment-Refinement
SPO: Structure-Process-Outcome

Edited by A Mavragani; submitted 14.04.22; peer-reviewed by M Antoniou, G Deckard; comments to author 08.05.22; revised version received 01.06.22; accepted 07.06.22; published 07.07.22.

Please cite as:
Tossaint-Schoenmakers R, Kasteleyn MJ, Rauwerdink A, Chavannes N, Willems S, Talboom-Kamp EPWA
Development of a Quality Management Model and Self-assessment Questionnaire for Hybrid Health Care: Concept Mapping Study
JMIR Form Res 2022;6(7):e38683
URL: https://formative.jmir.org/2022/7/e38683
doi:10.2196/38683
PMID:35797097

Original Paper

Assessing the Views and Needs of People at High Risk of Gestational Diabetes Mellitus for the Development of Mobile Health Apps: Descriptive Qualitative Study

Beibei Duan[1*], BSc; Zhe Liu[1*], BSc; Weiwei Liu[1], MD; Baohua Gou[2], BSc

[1]School of Nursing, Capital Medical University, Beijing, China

[2]Beijing Youyi Hospital, Capital Medical University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Weiwei Liu, MD
School of Nursing
Capital Medical University
No. 10, Xi Tou tiao, You'an Men Wai, Feng tai District
Beijing, 100069
China
Phone: 86 13581694127
Email: lwwhlxy@ccmu.edu.cn

Abstract

Background: Early prevention of gestational diabetes mellitus (GDM) can reduce the incidence of not only GDM, but also adverse perinatal pregnancy outcomes. Moreover, it is of great significance to prevent or reduce the occurrence of type 2 diabetes. Mobile health (mHealth) apps can help pregnant women effectively prevent GDM by providing risk prediction, lifestyle support, peer support, professional support, and other functions. Before designing mHealth apps, developers must understand the views and needs of pregnant women, and closely combine users' needs to develop app functions, in order to better improve user experience and increase the usage rate of these apps in the future.

Objective: The objective of this study was to understand the views of the high-risk population of gestational diabetes mellitus on the development of mobile health apps and the demand for app functions, so as to provide a basis for the development of gestational diabetes mellitus prevention apps.

Methods: Fifteen pregnant women with at least one risk factor for gestational diabetes were recruited from July to September 2021, and were interviewed via a semistructured interview using the purpose sampling method. The transcribed data were analyzed by the traditional content analysis method, and themes were extracted.

Results: Respondents wanted to develop user-friendly and fully functional mobile apps for the prevention of gestational diabetes mellitus. Pregnant women's requirements for app function development include: personalized customization, accurate information support, interactive design, practical tool support, visual presentation, convenient professional support, peer support, reasonable reminder function, appropriate maternal and infant auxiliary function, and differentiated incentive function.These function settings can encourage pregnant women to improve or maintain healthy living habits during their use of the app

Conclusions: This study discusses the functional requirements of target users for gestational diabetes mellitus prevention apps, which can provide reference for the development of future applications.

(JMIR Form Res 2022;6(7):e36392) doi:10.2196/36392

KEYWORDS

gestational diabetes mellitus; high-risk groups; mobile health; mHealth; applications; user-centered design; qualitative research

Introduction

Gestational diabetes mellitus (GDM) refers to the first occurrence of abnormal glucose tolerance of varying degrees during pregnancy, which is one of the most common complications of pregnancy [1]. Due to differences in diagnostic criteria and race, the global incidence of GDM is between 1% and 25% [2]. A previous study has shown that the incidence of GDM in China is 14.8%, and with the adjustment of the fertility

policy and the increase in elderly people and obese pregnant women, the incidence of GDM shows a trend of increasing every year [3,4]. GDM brings serious threats and challenges to both maternal and infant health. GDM can increase the risks of pregnancy hypertension, abortion, polyhydramnios, premature delivery, dystocia, and cesarean section [4-6]. Women who have a history of GDM have a 50%-73% risk of recurrence of GDM when they get pregnant again and a 10-fold higher risk of developing type 2 diabetes in 5 to 10 years after delivery, and the risk of cardiovascular disease is also higher in these women than in normal women [7]. Compared with the offspring of healthy women, the offspring of GDM patients have higher risks of fetal malformation, macrosomia, large for gestational age, hyperinsulinemia, neonatal hypoglycemia, pathological jaundice, and respiratory distress syndrome [8,9]. In addition, these children have increased risks of obesity, abnormal glucose tolerance, and type 2 diabetes when they become adults, bringing heavy economic burden to countries around the world [10]. Therefore, early preventive measures for pregnant women at high risk of GDM are of great significance in reducing the incidence of GDM, reducing adverse perinatal pregnancy outcomes, and preventing or reducing the occurrence of type 2 diabetes.

Mobile health (mHealth) refers to medical treatment and health management through mobile devices, such as mobile phones, patient health data monitoring devices, palm computers, and other wireless devices [11]. Mobile information technology can bridge the communication between medical service providers and users, helping doctors understand the health status of users and providing clinical decision support, so as to achieve remote diagnosis and treatment. At the same time, users' needs for self-management support can be met to improve their compliance and self-management behaviors [12-14]. In addition, mobile information technology can greatly save users' time and transportation costs, relieve the burden of medical treatment faced by hospitals and community health service institutions, and make the limited medical service resources the most effective. Therefore, mHealth has significant potential in the medical and health fields [15,16].

As one of the main forms of mHealth management, mHealth apps refer to health service platforms that use smartphones, tablets, and other mobile devices as terminals and rely on mobile internet technology to provide services for patients and medical staff [17]. With the popularity of mobile devices, more than 500 million smartphone users worldwide use mHealth apps for health management [18]. Due to their unique biofeedback function, mHealth apps can carry out real-time health assessment and provide feedback on patients' health status, which can be considered beneficial for medical staff to implement personalized and precise health management for patients, and can have good application prospects [19]. At present, mHealth apps have been widely used in the prevention and health management of diabetes patients and have achieved satisfactory results [20]. However, currently, apps are mostly developed for commercial needs and purposes, the needs of patients are seldom taken into account during development, and few patients are invited to participate in the design process of apps [21-23]. Studies have shown that lack of interest in apps, fees, and fear of personal information disclosure are the main reasons that hinder users from downloading and sticking with apps [24]. Therefore, when developing mHealth apps, it is necessary to take users as the center, pay attention to users' preferences and needs for using apps, and design apps that can meet users' needs, so as to improve users' experiences, increase users' stickiness in using apps, and improve patients' compliance with self-management.

User-centered design (UCD), which is part of human-computer interaction, takes user needs into account at every stage of product development. It is an important step and a relatively mature design method for building application programs [25]. Its core is to understand users and build their psychological model into system functions, so as to provide customized high-quality nursing services for each user more effectively [26]. In the process of app development, UCD gives priority to the needs of users, which can improve the stickiness and autonomy of users, and bring positive emotional experience to users, so as to meet the needs, preferences, and goals of users, and improve the quality of apps. The World Health Organization recommends that it be integrated into the whole process of an mHealth intervention to ensure the effectiveness of the intervention [27-30]. In recent years, UCD technology has been applied in the development of mHealth apps and has achieved certain effects in the areas of lifestyle intervention for patients with chronic diseases [25]. Studies have shown that most health management apps are not designed with user-centered methods, leading to poor usability and less use of these apps [31].

At present, some apps for GDM health management have been developed by scholars, but few studies have applied mHealth apps to GDM prevention management [32]. As mentioned above, understanding the usage preferences of mHealth apps among people at risk of gestational diabetes is crucial for the development of GDM prevention apps. However, there is no evidence that people at risk of gestational diabetes have a preference for mHealth apps. Qualitative research methods can dig deep into the inner needs of pregnant women. Therefore, the purpose of this study was to use a user-centered qualitative research method to interview pregnant women at high risk of GDM in early pregnancy, in order to explore the needs and preferences of these pregnant women for the functions and design of mHealth apps. The results can provide a reference for the design and development of mHealth apps for the prevention of gestational diabetes, and can help to adjust intervention measures, improve the acceptance and effectiveness of app use by pregnant women, and improve user engagement.

Methods

Study Design

In this study, the descriptive qualitative research method was adopted and the semistructured in-depth interview method was used to collect data. According to the purpose of the study, the interview outline was designed based on previous experience and reference to relevant literature. Before the formal interview, 2 pregnant women were preinterviewed, and the interview outline was modified appropriately according to the results of

the preinterview analysis. The formal interview outline has been presented in Textbox 1.

Textbox 1. Semistructured interview guide used in this study.

Guide
1. Which mobile health management apps have you used before? What features do you find particularly useful in these apps? Or which features are not very useful and could be improved further?
2. If an app for gestational diabetes prevention is developed to assist your daily health management, what functions do you expect it to have?
3. What are your requirements for the interface design of an app?
4. What other requirements that you would like us to consider for developing an app for gestational diabetes prevention?

Participants and Recruitment

The maximum difference sampling strategy was adopted to select pregnant women with different age, occupation, parity, and gestational diabetes risk factors as far as possible for the purpose of sampling, so as to ensure diversity among the respondents. From July to September 2021, in the obstetrics clinic of a third-class hospital in Beijing, a researcher publicized the project to pregnant women, who made an appointment to establish routine health records, and invited pregnant women to share their experiences and opinions in depth in the form of face-to-face interviews. The inclusion criteria were as follows: (1) at least one risk factor for gestational diabetes (eg, advanced maternal age, overweight or obesity, family history of diabetes, history of GDM, history of macrosomia, and impaired fasting glucose); (2) gestational age <12 weeks; (3) experience of using sports apps, maternal and child health apps, or other health management apps; (4) good communication skills; and (5) informed consent and voluntary participation in this study. The exclusion criteria were as follows: (1) age <18 years and (2) presence of mental disorders. The interviews in this study were stopped when no new topics appeared, that is, the data were saturated. A total of 16 pregnant women were interviewed, and 1 of the women dropped out because she had something else to do halfway. Finally, 15 women participated in semistructured interviews (participant #1 to participant #15). The mean age of the respondents was 32 (SD 3.44) years, and the mean gestational age was 6.8 (SD 0.74) weeks. There were 12 first-time pregnancies and 3 second-time pregnancies. One of the pregnant women was a stay-at-home mother, and the other 14 were from different occupations. Six had 2 or more risk factors, and the remaining 9 had only 1 risk factor. Among the 15 pregnant women, 1 was of Hui nationality and the other 14 were of Han nationality. With regard to the education level, 3 had junior college or below education, 10 had undergraduate education, and 2 had graduate or above education. In terms of cost, 1 of the 15 pregnant women was self-paid, while the rest were covered by medical insurance.

Data Collection

Face-to-face interviews were conducted. Each interview was conducted in a quiet outpatient lounge with no third person to disturb. The interview duration was 20-40 minutes. Before the interview, interviewers introduced to interviewees the definition, harm, and intervention status of gestational diabetes. The purpose and significance of the interviews were also informed. Interviewees were told that their privacy would be protected by the researchers. The interview was recorded with the consent of the interviewees. The interview was conducted according to the interview outline. During the interview, the interviewers carefully listened to the statements; appropriately responded to them with questioning, repetition, clarification, response, and summary; encouraged participants to fully express their ideas; avoided inductive questioning; and timely recorded the key information of the interview. At the same time, they paid attention to observe and record the interviewees' nonverbal information, such as a pause, a smile, body language, and mood change. After each interview, a reflective diary was written to reflect on the problems in the interview and correct them in the next interview. After the interview, the interviewees were thanked for their participation and were informed of the possibility of contacting them again for further information.

Data Analysis

Data collection and analysis were conducted simultaneously. After each interview, interviewers listened to the original materials repeatedly. The materials were transcribed within 24 hours. Uncertainties were clarified in time in combination with on-site notes. The traditional content analysis method was used for data analysis. The final transcribed text was merged into a single text by topic. Interviewers read the text several times to get a sense of the whole text. Selection criteria were determined according to research objectives and research questions. Based on this standard, the text content was classified, and meaningful statements were extracted and coded. Interviewers read and analyzed the semantic units carefully, and distinguished and summarized the theme. Researchers also looked for the relationship between subjects, and formed a theme group. This cycle continued until saturation (no new themes or subthemes were present) [33].

Quality Control

Before the interviews, the researchers received training in qualitative research, read a large number of relevant literature and books, and learned the analytical methods of qualitative research. The interviewers were involved in the obstetrics clinic as student nurses. As a research tool, researchers always remained neutral. Interviewers truthfully recorded the information provided by interviewees and analyzed their body language and facial expressions. Two researchers with training independently analyzed and discussed the data until the coding information reached a consensus. In the process of data analysis, researchers paid attention to the use of suspension to avoid interference caused by researchers. This study has been reported according to the requirements of the Consolidated Criteria for

Reporting Qualitative Research (COREQ) checklist for qualitative studies [34].

Ethical Considerations

Before each interview, the interviewee was given an explained on the research objectives, methods, expected benefits, and potential risks. Interviewers informed interviewees that relevant information would be strictly confidential. Interviewees could choose to accept or refuse participation in the study, and they could withdraw from the study at any time during the interview. During the interview, interviewees were told that they could refuse to answer any questions that they did not want to answer. Interviewees voluntarily participated in the interviews. In addition, interviewees were told that the content would be used only for scientific research. The interview information was coded, and the researchers did not compromise the privacy of the interviewees.

Ethics Approval

This study was approved by the Ethics Committee of Capital Medical University (batch number: Z2019SY037) and the Ethics Committee of the hospital conducting the interviews (batch number: 2019-P2-204-02).

Results

Design

Theme 1: User-Friendly Interface Design

Pregnant women expected clear logic between modules on the app interface and human-computer interaction.

> *Apps should have logic. What are the second-level interface and third-level interface after opening app? This hierarchy and interface framework should be clear. Make sure the entry and return routes are clear, and then the modules are clear.* [Participant #1]

> *Categorize weight management, diet and exercise so that they can be easily seen on the home page. As clear as the app module of hospital appointment. Don't make me look for it, because it's too hard to look for it. A lot of apps these days are really annoying.* [Participant #4]

> *Hopefully the app doesn't lag, is smooth to use and doesn't have too many ads.* [Participant #6]

> *App can be divided into modules, like what's this, what's that. And when it updates content, don't always change the location.* [Participant #7]

> *The app guidelines should be clear. For example, interface modules can be divided into early pregnancy, middle pregnancy and delivery, which are mainly practical and simple.* [Participant #10]

Theme 2: Rich Functionality

Pregnant women hoped that the health management apps they use would cover the common functions of blood glucose management during pregnancy, so as to minimize the use of multiple apps.

> *Now everyone has several apps in their mobile phones. I hope that in the future there will be an app that contains all pregnancy programs, as convenient as possible.* [Participant #4]

> *If the software can provide a one-stop service and integrate these commonly used functions together, I will use only one app. I think it will be much better.* [Participant #11]

Functional Requirements

Theme 1: Personalized Customization

Pregnant women hoped to obtain personalized recipes customized by medical staff according to their own dietary preferences in mHealth apps, and hoped that these apps have a complete food bank and that they can search for the glycemic index of foods in the apps, so that they could flexibly carry out dietary replacement to reduce their dietary decision confusion.

> *I've always fantasized about an app that recommends foods I love. I don't have to think about matching my diet myself. It's all done for me, so I just follow the app. But only if I can choose what I like to eat, or if the app removes several ingredients I don't like to eat from the recipe. It would be nice if the app had a search function. For example, if I want to eat strawberries, I can see what the glycemic index of strawberries is and whether it's recommended for me to eat them.* [Participant #1]

> *I find it convenient to have recommended recipes in the app so I don't have to think about whether the food is edible or not. It's not realistic for me to follow the recipes exactly, but I'm free to mix and match, as long as the total calories are right, and I think it's better.* [Participant #4]

In addition, some pregnant women also expressed a need for exercise customization.

> *I hope the app can let me choose the exercise I want to do every day. For example, today I want to do yoga, and the app can calculate how long I need to do yoga according to my current weight and calorie intake to reach the goal of burning all calories.* [Participant #4]

Theme 2: Accurate Information Support

As the pace of work and life is accelerating, some pregnant women stated that popular science articles recommended by mHealth apps should have attractive titles, be relevant to them, and be more accurate.

> *At my age, when I have to juggle work and family, I really don't have much free time. For popular science articles, the content is already boring, and if the volume is any longer, users will have no interest or time to read it.* [Participant #1]

> *I hope that the app can set up a function of a prenatal assistant to tell me what I should do at different gestational weeks and whether I should have an empty*

stomach, etc. This is my favorite feature. [Participant #1]

The app should recommend relevant contents for pregnant women based on their individual weight and test values. I think this will be better, because I think people are extremely busy nowadays, especially when they have children. [Participant #4]

We hope that the app can push suitable exercise types for pregnant women according to their gestational weeks. [Participant #7]

If it's a video, it can be put in the appropriate module and I'll pay attention to it when I get to that stage of pregnancy. For example, if it's about breastfeeding, it'll be put in the third trimester and I'll click on it when I'm in labor. [Participant #10]

Only when the title of an article is clear and relevant to me will I read it. For example, if I'm eight weeks pregnant, what do I need to be aware of at this stage? I might click on it. But if it's about what to eat and control, I'm probably not going to read it. [Participant #11]

Theme 3: Interactive Design

Pregnant women hoped that mHealth apps have the function of setting daily calorie and exercise goals. After the pregnant women input their own diet, exercise, and blood sugar monitoring data, the system should provide immediate feedback and professional feedback according to their health management standards, so that the pregnant women can make self-adjustments to achieve their daily management goals.

I hope the app can record the number of steps on the day, and tell me whether the amount of exercise today meets the standard, so as to give me an evaluation. [Participant #1]

If I have a problem with my blood sugar, I may check my blood sugar every once in a while and record the result of each time on the app. Then the application platform will give feedback based on my data and tell me what I need to pay attention to in the next step. I think it would be nice to have that kind of feedback. [Participant #4]

The app can score me according to my exercise level every week, and then give me an adjustment plan, which forms a closed loop from input, scoring and feedback. [Participant #10]

Theme 4: Practical Tool Support

Pregnant women hoped that mHealth apps can be used as tools to assist them in self-management and that these apps can record their steps to help them know whether their exercise meets the standard, or can remind and motivate them to exercise.

Although I am in a first-tier city like Beijing, there are few pregnant women around me who do yoga or swim. Especially when you're expecting your first child, families are very cautious. The elders in the family are so ingrained in their beliefs that they may not be comfortable with their children doing yoga during pregnancy, and most would probably prefer to stick to walking during pregnancy. Therefore, it is sufficient for the app to record the number of steps of motion. [Participant #5]

For example, I can punch in the app to record my steps today. [Participant #7]

In addition, pregnant women hoped that mHealth apps have the function of weight recording to show the trend of their own weight change and that these apps can automatically compare the actual weight increase of pregnant women with the recommended weight increase range to determine whether the weight gain is reasonable, so as to help pregnant women manage their weight during pregnancy.

The app needs to have the function of weight graph, through which we can see the recent trend of pregnant women's weight. I remember when I was pregnant with my first child, there was an app that told me how much weight I needed to gain in my current gestational week. [Participant #4]

I am using an app for weight monitoring. I need to input my weight into the app every day to observe the change of my weight. [Participant #9]

In addition, some pregnant women hoped that mHealth apps can be used as practical tools to assist doctors in pregnancy management.

The app itself is also a tool to help doctors manage pregnancy health for pregnant women. Since everyone is different, doctors will definitely need to manage pregnancy according to their different physical conditions. It would be better if the app could be linked to hospital records and checklists. [Participant #6]

Theme 5: Visual Presentation

Some pregnant women stated that they do not have time to watch live broadcasts because of the fast pace of life. They hoped that the information on mHealth apps can be presented in the form of cartoon pictures, risk assessment forms, recorded videos, and short videos, so as to obtain relevant information more quickly.

The recommended article can be in the form of text with cartoon pictures, or it can be one of those risk self-assessment scales for pregnant women to rate their recent status. If the risk score is higher than a certain point, the woman has gestational diabetes. [Participant #10]

The live broadcast lasts a long time, and I may not be able to listen to it, because the pace of life is fast now, and life is so busy every day. Let alone live broadcast, I may skip to watch the recorded broadcast. [Participant #5]

I think recording is better than live broadcasting. Medical staff can put the video of the lecture in the corresponding module. If I had time, I would listen. [Participant #10]

> *I may not have time to watch the health education live broadcast, if there is a replay, I will watch it. It's better to put subtitles on the video because it's slow and I tend to skip to the subtitles or just read the document.* [Participant #6]
>
> *If experts' lectures are made into some short videos, I may watch them at any time.* [Participant #4]

Theme 6: Convenient Professional Support

Some pregnant women hoped that mHealth apps can provide the online consultation function of experts, so that it is more convenient to contact experts and solve some minor problems that are not too urgent, in order to avoid the cost of time and energy from the round trip to the hospital.

> *It's more difficult for pregnant women to get to the hospital. If I have any minor questions, I would like to consult one of the experts while they are giving an online lecture. I hope the app can provide free online consultation.* [Participant #1]
>
> *Pregnant women may have some simple problems, but it is not convenient to come to the hospital for registration. It would be better if the app could provide online consultations with doctors. Doctors can choose to reply to pregnant women's information when they have time, there is no rush to reply immediately.* [Participant #4]
>
> *It would be better if pregnant women had some problems that did not need to go to the hospital and could be answered online at a lower cost than going to the hospital.* [Participant #9]

Theme 7: Peer Support

Some pregnant women hoped to set up a function like WeChat Moments or a forum, in order to facilitate access to other pregnant women who have encountered problems and facilitate the exchange of experiences between pregnant women, so as to better deal with the problems of pregnancy.

> *When I was first pregnant, I went to the forum to see what other people were like when they were pregnant. Many people would share their early pregnancy experiences on the forum, such as problems with their checklists.* [Participant #1]
>
> *If I have some problems, I will look at other users' posts to see how they solve the problems. For example, I have a tummy ache at the early stage of pregnancy, so I will search for users who have the same experience on the Internet. If most users say this is normal, it means that tummy ache is ok.* [Participant #8]

Theme 8: Reasonable Reminder Function

Pregnant women hoped to set reminders according to their own needs. Excessive reminders can disturb pregnant women.

> *We may not always want to record our diet every day. I hope the app can remind me when I forget to fill in.* [Participant #7]
>
> *I hope the app has a reminder function, such as when I should do activities, when I should drink water, timely reminding me will be better.* [Participant #10]
>
> *I think it is still necessary to set reminders, because the pace of work and life is fast now, pregnant women cannot remember everything every day.* [Participant #14]

However, some pregnant women believed that an app's reminder function would cause some trouble, so they had little demand for the reminder function.

> *Because there are too many reminders in the app, I don't set any reminders in all my apps. I turn them off. If I want to keep track of something important, I'll set up a reminder myself.* [Participant #9]

Theme 9: Appropriate Maternal and Infant Auxiliary Function

Some pregnant women wanted to add mother-baby support tools to mHealth apps to make it through the pregnancy better.

> *I hope the app can tell me the growth of my baby, so that I can know the development of the baby. In addition, I hope the app can tell me what changes a pregnant woman will have during pregnancy and what changes are normal.* [Participant #1]
>
> *I think the app can add a function to record the gestational weeks.* [Participant #2]
>
> *I have heard that some apps can record contractions and fetal movements. Is it possible to add these functions?* [Participant #3]
>
> *I think there should be a keyword search function, I can query relevant articles, just like Baidu, I want to know can be found in it.* [Participant #9]

Theme 10: Differentiated Incentive Function

Different pregnant women had different views on the redemption function of points. Some thought that it would be better if the points can be exchanged for items they need, while some thought that it has no incentive effect on them.

> *It is possible to get points by checking in, posting, reading articles, ranking points, and exchanging points for small things, etc. Some people may prefer this function.* [Participant #5]
>
> *I think it would be nice if I could exchange my accumulated points for something to use during pregnancy or in the future.* [Participant #6]
>
> *I don't like those things very much. They are not very useful. I sometimes check to see how many points I have, but I don't think the points are of any use to me.* [Participant #7]

Discussion

Principal Findings

From the perspective of users, this study discussed the views and function requirements of pregnant women at high risk of GDM with regard to the development of mHealth apps for

preventing GDM, which can provide a reference for app developers to develop apps in line with users' usage habits in the future. In this study, pregnant women hoped that when using apps for management, they could receive personalized health management strategies according to their different stages of pregnancy and could receive immediate feedback on their daily management, helping them make adjustments. When pregnant women encountered difficulties in self-management, they hoped to obtain expert and peer resources in the apps to provide them with social support. In addition, pregnant women hoped that developers and professionals would consider the need for quick and accurate information acquisition when designing mHealth apps and would develop user-friendly apps. They stated that the content should be visual and attractive, and should provide both valuable and accurate information to meet the needs of information support.

Comparison With Prior Work

This study found that it is of great significance to formulate corresponding personalized management strategies for pregnant women based on the results of their health assessment, which is similar to the results of previous studies [35-37]. A research report pointed out that women were dissatisfied with the limited response of apps, and hoped that the apps could specifically analyze the characteristics of users and provide different solutions [38]. The biggest difference between personalized apps and other apps is that personalized apps can provide personalized services for pregnant women in different stages of pregnancy [39]. Studies have shown that personalized features in apps could improve users' compliance with lifestyle interventions, which is one of the most common intervention strategies in mHealth apps [40].

Interactive function is considered to be the most popular function in mobile medical apps [41]. The biggest advantage of interactive apps is that they can formulate targeted behavior change plans for users, continuously monitor the behavior characteristics of users, and then give targeted feedback according to the implementation of users' behaviors, which can motivate users to complete behavior changes [42]. Interactive design strengthens the connection between the user and the mobile app, providing psychological support for pregnant women [43]. In this study, pregnant women hope that the apps could compare the health information of pregnant women with the recommended behavioral goals, and provide feedback instantly for pregnant women to help them make timely behavioral adjustments, which is consistent with previous research results [44-46]. Some studies have used telephone or WeChat groups to intervene in pregnant women at high risk of GDM [47,48]. However, it is difficult to achieve instant feedback between doctors and patients, and the information of telephone and WeChat groups is also easily forgotten and covered. Therefore, it is important to design an app to prevent gestational diabetes. There is also a study using an app that counts low glycemic index function to provide dietary guidance to obese pregnant women, but this app has a single function and is designed for dietary guidance only, which cannot achieve the goal of interaction, resulting in an insignificant intervention effect [49].

Due to time constraints, most working pregnant women hoped apps would accurately push information related to themselves. Previous studies have found that providing users with personally relevant information was an important factor in promoting the use of mHealth programs [50]. In addition, Naughton et al [51] found that pushing personalized and relevant information could increase the value of contents and prevent users from falling out of mHealth management. In a qualitative study of Saudi women, it was found that due to social and cultural restrictions, some obese women could not do enough physical exercise outdoors [52]. Interviewees hoped that apps could recommend exercise methods that meet actual conditions for users and provide advice according to users' preferences.

Most pregnant women hoped that information could be in visual forms, such as illustrations and videos, to increase attractiveness. Presenting content in a visual way can help women obtain more information in a short period of time to meet women's needs for pregnancy knowledge. Previous studies have pointed out that multimedia could adapt to the learning styles of different individuals, enable pregnant women to take the initiative in learning, and encourage pregnant women to actively study [53,54]. Some studies have pointed out that the memory acquired by watching videos was stronger than that acquired by browsing text or pictures [55]. This study found that in terms of the video presentation form, pregnant women were more inclined to short videos or recorded videos. Therefore, when designing apps, developers should use a visual method as much as possible to present a variety of health information and educational resources for pregnant women to meet their needs for information support.

During the daily management of pregnancy, pregnant women will obtain decision-making information from different channels. Studies have shown [56] that pregnant women tend to trust professional decision support. Professional support provided by medical staff can meet the needs of pregnant women in pregnancy health care, which relieves anxiety and uncertainty of women during pregnancy [57,58] and helps pregnant women establish a healthy lifestyle [59]. When professionals cannot provide timely and effective feedback, pregnant women usually seek help from relatives and friends or through the internet. However, due to the poor quality of advice from various parties, pregnant women often get confused and make wrong decisions, which is not conducive to health management [60]. This study found that pregnant women hoped to obtain professional feedback from apps when they encountered problems or in the process of daily health management. This is similar to previous studies by Edwards et al [61] and Lau et al [43]. Previous studies have shown that when patients were unable to communicate effectively with experts, they felt bored and useless to use an app, resulting in a low utilization rate [36]. Therefore, when developing apps in the future, professional support functions should be added. In addition, when apps are actually used to manage the health of pregnant women, managers should coordinate the workload distribution of medical staff and include answering questions from pregnant women into their work scope, providing users with timely and effective professional support.

In addition to professional support, peer support also plays an important role in the formation and maintenance of patients'

self-management abilities. Peers can communicate with each other on their own experience, attitudes, and concepts, giving relevant suggestions to each other. Pregnant women in this study hoped that they could share confusion, emotion, and experience with peers who had similar needs and common goals through an app; thus, these women can obtain information and emotional support. This is similar to the findings of McDonald et al [62]. Studies have shown that peer support could overcome loneliness, powerlessness, and stress in pregnant women and increase maternal self-efficacy. Emotional support and exchange of experiences with peers are more acceptable to pregnant women. In addition, pregnant women can gain motivation from peer support to use an app [63]. Previous studies have found that providing a communication platform for users in a GDM prevention app could allow users to exchange their experiences and deal with problems during pregnancy, motivate users to achieve their expected goals, and increase the frequency of usage [64]. Therefore, apps should provide a module for pregnant women to communicate with their peers, allowing them to learn problem-solving skills through mutual assistance and to enhance psychological support.

Although some pregnant women in this interview study believed that the point reward function was of little significance, research has proven that the introduction of game mechanics could increase the user's sense of achievement and could cause the user to put more effort into accomplishing goals. Cafazzo et al [65] incorporated a gamification mechanism when designing a diabetes blood sugar management program. When young patients regularly checked their blood glucose, they could obtain corresponding points. These points could be redeemed for iTunes codes to purchase music and applications. The frequency of blood glucose monitoring increased by nearly 50% in patients using the apps. Ekezie et al [64] provided a virtual map for gestational diabetes mellitus patients in his app to improve the interest of patients' exercise, and set a step ranking list to encourage users to achieve exercise goals through competition, so as to increase user efficiency. It is suggested that when developing gestational diabetes prevention apps in the future, the gamification and reward mechanisms should be improved according to users' needs. Furthermore, more attractive methods should be set up to better assist patients in performing self-management, as well as raise their enthusiasm for using these apps.

Strengths and Limitations

Previous research has mostly focused on the user's experience and evaluation of developed apps after using them, and few users have participated in the design of apps, leading to poor user stickiness. In this study, we conducted deep interviews on the preferences and functional needs of mHealth apps among people at risk for gestational diabetes, and the findings are important to guide the development of future interventions and other pregnancy health apps. In addition, the strength of this study is that the data analysis was performed by two researchers and information was continuously validated during the data analysis process. One of the researchers was the interviewer and the other was not involved in the interview. Therefore, in the process of data analysis, the two researchers had a different understanding of the existing data. Then they discussed and finally reached a more objective and unified opinion, and provided the analysis results to the participants to make their opinions expressed fully and correctly.

This study has several limitations. First, this study only interviewed 15 pregnant women in the same hospital in Beijing, China. Regardless of the fact that the sample size of this study reached saturation, the method of maximum difference sampling was adopted to select the research objects as much as possible. However, no interviewees from rural groups were included; thus, the interview results may not reflect the diverse needs of all pregnant women, making it difficult to generalize the findings to women in rural areas or different cultural situations. Before the formal interview, interviewers trained and practiced to improve their interviewing skills. However, as interviewers were conducting interviews for the first time, their interviewing experience and skills were limited, which may have affected the comprehensiveness of data collection. In addition, interviewers conducted interviews as medical interns, which made some interviewees reluctant to express their true feelings and needs, leading to a potential impact on data collection.

Conclusion

mHealth apps can provide new tools for the health management of people at risk for gestational diabetes. Before the development of a gestational diabetes prevention app, this study deeply collected the views and functional requirements of mHealth app development from gestational diabetes risk groups. The results can provide valuable information for the future development of mHealth apps related to the prevention of gestational diabetes that meet users' needs, and can provide a reference for the design of intervention content in the future. In addition, the findings can provide a reference for the development of other types of apps during pregnancy.

Acknowledgments

The authors sincerely thank all the interviewees for their active participation in this study. This study was supported by the National Natural Science Foundation of China (project number 71904133).

Authors' Contributions

BD performed data collection, performed data analysis and interpretation, and drafted the manuscript. ZL assisted in qualitative data analysis and interpretation, and co-drafted the manuscript. BG participated in manuscript review. WL conceived and designed the study and was responsible for revising the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Department of Obstetrics and Gynecology, Chinese Medical Association. Guidelines for diagnosis and treatment of gestational diabetes mellitus. Chinese Journal of Perinatal Medicine 2014;17(08):537-545. [doi: 10.3760/cma.j.issn.1007-9408.2014.08.009]
2. Moyer VA, U.S. Preventive Services Task Force. Screening for gestational diabetes mellitus: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2014 Mar 18;160(6):414-420 [FREE Full text] [doi: 10.7326/M13-2905] [Medline: 24424622]
3. Gao C, Sun X, Lu L, Liu F, Yuan J. Prevalence of gestational diabetes mellitus in mainland China: A systematic review and meta-analysis. J Diabetes Investig 2019 Jan;10(1):154-162 [FREE Full text] [doi: 10.1111/jdi.12854] [Medline: 29683557]
4. Juan J, Yang H. Prevalence, prevention, and lifestyle intervention of gestational diabetes mellitus in China. Int J Environ Res Public Health 2020 Dec 18;17(24):9517 [FREE Full text] [doi: 10.3390/ijerph17249517] [Medline: 33353136]
5. Assaf-Balut C, García de la Torre N, Calle-Pascual AL. Detection, treatment and prevention programs for gestational diabetes mellitus: The St Carlos experience. Endocrinología, Diabetes y Nutrición (English ed.) 2020 May;67(5):342-350. [doi: 10.1016/j.endien.2020.06.009]
6. Artal R. The role of exercise in reducing the risks of gestational diabetes mellitus in obese women. Best Pract Res Clin Obstet Gynaecol 2015 Jan;29(1):123-132. [doi: 10.1016/j.bpobgyn.2014.05.013] [Medline: 25240421]
7. Juan J, Yang H, Tang X. Gestational diabetes: An opportunity to improve maternal and child health. Chinese Journal of Perinatal Medicine 2020;23(10):717-718. [doi: 10.3760/cma.j.issn.1007-9408.2020.10.104]
8. Seghieri G, Di Cianni G, Seghieri M, Lacaria E, Corsi E, Lencioni C, et al. Risk and adverse outcomes of gestational diabetes in migrants: A population cohort study. Diabetes Res Clin Pract 2020 May;163:108-128. [doi: 10.1016/j.diabres.2020.108128] [Medline: 32259610]
9. Yu G, Du Y. Risk factors of gestational diabetes mellitus and its influence on pregnancy outcome. Maternal & Child Health Care of China 2017;32(17):4070-4072. [doi: 10.7620/zgfybj.j.issn.1001-4411.2017.17.18]
10. Diabetes Canada Clinical Practice Guidelines Expert Committee, Feig DS, Berger H, Donovan L, Godbout A, Kader T, et al. Diabetes and pregnancy. Can J Diabetes 2018 Apr;42 Suppl 1:S255-S282. [doi: 10.1016/j.jcjd.2017.10.038] [Medline: 29650105]
11. WHO Global Observatory for eHealth. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. World Health Organization. 2011. URL: http://apps.who.int/iris/handle/10665/44607 [accessed 2022-06-13]
12. Mackillop L, Loerup L, Bartlett K, Farmer A, Gibson OJ, Hirst JE, et al. Development of a real-time smartphone solution for the management of women with or at high risk of gestational diabetes. J Diabetes Sci Technol 2014 Jul 07;8(6):1105-1114 [FREE Full text] [doi: 10.1177/1932296814542271] [Medline: 25004915]
13. Slater H, Campbell JM, Stinson JN, Burley MM, Briggs AM. End user and implementer experiences of mhealth technologies for noncommunicable chronic disease management in young adults: Systematic review. J Med Internet Res 2017 Oct 12;19(12):e406 [FREE Full text] [doi: 10.2196/jmir.8888] [Medline: 29233804]
14. Mallen JR, Shah MU, Drake R, Kreicher K, Falcone T, Karter N, et al. Utility of smartphone telemedical consultations for peritonsillar abscess diagnosis and triage. JAMA Otolaryngol Head Neck Surg 2020 Aug 20;146(10):909-913 [FREE Full text] [doi: 10.1001/jamaoto.2020.1972] [Medline: 32816011]
15. Athavale P, Thomas M, Delgadillo-Duenas AT, Leong K, Najmabadi A, Harleman E, et al. Linking high risk postpartum women with a technology enabled health coaching program to reduce diabetes risk and improve wellbeing: Program description, case studies, and recommendations for community health coaching programs. J Diabetes Res 2016;2016:4353956 [FREE Full text] [doi: 10.1155/2016/4353956] [Medline: 27830157]
16. Pal K, Dack C, Ross J, Michie S, May C, Stevenson F, et al. Digital health interventions for adults with type 2 diabetes: Qualitative study of patient perspectives on diabetes self-management education and support. J Med Internet Res 2018 Jan 29;20(2):e40 [FREE Full text] [doi: 10.2196/jmir.8439] [Medline: 29463488]
17. 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 30;19(8):e301 [FREE Full text] [doi: 10.2196/jmir.7814] [Medline: 28855146]
18. FDA Takes on Devicelike Mobile Apps. Medical Device + Diagnostic Industry. 2013 Sep 23. URL: https://www.mddionline.com/digital-health/fda-takes-devicelike-mobile-apps [accessed 2022-06-13]
19. Wang Q, Markopoulos P, Yu B, Chen W, Timmermans A. Interactive wearable systems for upper body rehabilitation: a systematic review. J Neuroeng Rehabil 2017 Mar 11;14(1):20 [FREE Full text] [doi: 10.1186/s12984-017-0229-y] [Medline: 28284228]
20. Nørgaard SK, Nichum VL, Barfred C, Juul HM, Secher AL, Ringholm L, et al. Use of the smartphone application "Pregnant with Diabetes". Dan Med J 2017 Nov;64(11):A5417 [FREE Full text] [Medline: 29115204]

21. Garcia-Zapirain B, de la Torre Díez I, Sainz de Abajo B, López-Coronado M. Development, technical, and user evaluation of a web mobile application for self-control of diabetes. Telemed J E Health 2016 Sep;22(9):778-785. [doi: 10.1089/tmj.2015.0233] [Medline: 26981852]
22. Rangraz Jeddi F, Nabovati E, Hamidi R, Sharif R. Mobile phone usage in patients with type II diabetes and their intention to use it for self-management: a cross-sectional study in Iran. BMC Med Inform Decis Mak 2020 Feb 07;20(1):24 [FREE Full text] [doi: 10.1186/s12911-020-1038-y] [Medline: 32033560]
23. Jimenez G, Lum E, Car J. Examining diabetes management apps recommended from a Google search: Content analysis. JMIR Mhealth Uhealth 2019 Jan 16;7(1):e11848 [FREE Full text] [doi: 10.2196/11848] [Medline: 30303485]
24. Krebs P, Duncan DT. Health app use among US mobile phone owners: A national survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
25. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251 [FREE Full text] [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
26. 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]
27. Saunders CH, Durand M, Scalia P, Kirkland KB, MacMartin MA, Barnato AE, et al. User-centered design of the consideRATE questions, a measure of people's experiences when they are seriously ill. J Pain Symptom Manage 2021 Mar;61(3):555-565.e5 [FREE Full text] [doi: 10.1016/j.jpainsymman.2020.08.002] [Medline: 32814165]
28. Graham AK, Wildes JE, Reddy M, Munson SA, Barr Taylor C, Mohr DC. User-centered design for technology-enabled services for eating disorders. Int J Eat Disord 2019 Oct 16;52(10):1095-1107 [FREE Full text] [doi: 10.1002/eat.23130] [Medline: 31313370]
29. Walden A, Garvin L, Smerek M, Johnson C. User-centered design principles in the development of clinical research tools. Clin Trials 2020 Dec 20;17(6):703-711. [doi: 10.1177/1740774520946314] [Medline: 32815381]
30. Taylor CB, Graham AK, Fitzsimmons-Craft EE, Sadeh-Sharvit S, Balantekin KN, Flatt RE, et al. Optimizing eating disorder treatment outcomes for individuals identified via screening: An idea worth researching. Int J Eat Disord 2019 Nov;52(11):1224-1228 [FREE Full text] [doi: 10.1002/eat.23169] [Medline: 31502312]
31. Gaynor M, Schneider D, Seltzer M, Crannage E, Barron ML, Waterman J, et al. A user-centered, learning asthma smartphone application for patients and providers. Learn Health Syst 2020 Jul 18;4(3):e10217 [FREE Full text] [doi: 10.1002/lrh2.10217] [Medline: 32685685]
32. Tumminia A, Vitacolonna E, Sciacca L, Dodesini AR, Festa C, Lencioni C, et al. "MySweetGestation": A novel smartphone application for women with or at risk of diabetes during pregnancy. Diabetes Res Clin Pract 2019 Dec;158:107896 [FREE Full text] [doi: 10.1016/j.diabres.2019.107896] [Medline: 31669627]
33. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
34. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
35. Berg M, Adolfsson A, Ranerup A, Sparud-Lundin C, University of Gothenburg Centre for Person-Centred Care. Person-centered Web support to women with type 1 diabetes in pregnancy and early motherhood--the development process. Diabetes Technol Ther 2013 Jan;15(1):20-25. [doi: 10.1089/dia.2012.0217] [Medline: 23297670]
36. Halili L, Liu R, Hutchinson KA, Semeniuk K, Redman LM, Adamo KB. Development and pilot evaluation of a pregnancy-specific mobile health tool: a qualitative investigation of SmartMoms Canada. BMC Med Inform Decis Mak 2018 Nov 12;18(95):95 [FREE Full text] [doi: 10.1186/s12911-018-0705-8] [Medline: 30419896]
37. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
38. O'Reilly SL, Laws R. Health-e mums: Evaluating a smartphone app design for diabetes prevention in women with previous gestational diabetes. Nutr Diet 2019 Nov 14;76(5):507-514. [doi: 10.1111/1747-0080.12461] [Medline: 30109762]
39. Lim JY, Kim JK, Kim Y, Ahn S, Yu J, Hwang JH. A Modular Mobile Health App for Personalized Rehabilitation Throughout the Breast Cancer Care Continuum: Development Study. JMIR Form Res 2021 Apr 13;5(4):e23304 [FREE Full text] [doi: 10.2196/23304] [Medline: 33847589]
40. Chen J, Lieffers J, Bauman A, Hanning R, Allman-Farinelli M. Designing health apps to support dietetic professional practice and their patients: Qualitative results from an international survey. JMIR Mhealth Uhealth 2017 Mar 31;5(3):e40 [FREE Full text] [doi: 10.2196/mhealth.6945] [Medline: 28363882]
41. Tripp N, Hainey K, Liu A, Poulton A, Peek M, Kim J, et al. An emerging model of maternity care: smartphone, midwife, doctor? Women Birth 2014 Mar;27(1):64-67. [doi: 10.1016/j.wombi.2013.11.001] [Medline: 24295598]
42. Sannino G, Forastiere M, De Pietro G. A wellness mobile application for smart health: Pilot study design and results. Sensors (Basel) 2017 Mar 17;17(3):611 [FREE Full text] [doi: 10.3390/s17030611] [Medline: 28304332]

43. Goetz M, Müller M, Matthies LM, Hansen J, Doster A, Szabo A, et al. Perceptions of patient engagement applications during pregnancy: A qualitative assessment of the patient's perspective. JMIR Mhealth Uhealth 2017 May 26;5(5):e73 [FREE Full text] [doi: 10.2196/mhealth.7040] [Medline: 28550005]
44. Lupton D, Pedersen S. An Australian survey of women's use of pregnancy and parenting apps. Women Birth 2016 Aug;29(4):368-375. [doi: 10.1016/j.wombi.2016.01.008] [Medline: 26874938]
45. Lau Y, Cheng LJ, Chi C, Tsai C, Ong KW, Ho-Lim SST, et al. Development of a healthy lifestyle mobile app for overweight pregnant women: Qualitative study. JMIR Mhealth Uhealth 2018 Apr 23;6(4):e91 [FREE Full text] [doi: 10.2196/mhealth.9718] [Medline: 29685868]
46. Lim K, Chan S, Lim SL, Tai BC, Tsai C, Wong SR, et al. A smartphone app to restore optimal weight (SPAROW) in women with recent gestational diabetes mellitus: Randomized controlled trial. JMIR Mhealth Uhealth 2021 Mar 16;9(3):e22147 [FREE Full text] [doi: 10.2196/22147] [Medline: 33724204]
47. Ferrara A, Hedderson MM, Brown SD, Ehrlich SF, Tsai A, Feng J, et al. A telehealth lifestyle intervention to reduce excess gestational weight gain in pregnant women with overweight or obesity (GLOW): a randomised, parallel-group, controlled trial. The Lancet Diabetes & Endocrinology 2020 Jun;8(6):490-500. [doi: 10.1016/s2213-8587(20)30107-8]
48. Meng C. Investigation on the influence of wechat platform pregnancy nutrition intervention on pregnancy outcome. Henan Journal of Preventive Medicine 2020;31(5):352-353,361. [doi: 10.13515/j.cnki.hnjpm.1006-8414.2020.05.010]
49. Zhang Y, Wang L, Yang W, Niu D, Li C, Wang L, et al. Effectiveness of low glycemic index diet consultations through a diet glycemic assessment app tool on maternal and neonatal insulin resistance: A randomized controlled trial. JMIR Mhealth Uhealth 2019 Apr 18;7(4):e12081 [FREE Full text] [doi: 10.2196/12081] [Medline: 30998227]
50. Shorten A, Fagerlin A, Illuzzi J, Kennedy HP, Lakehomer H, Pettker CM, et al. Developing an internet-based decision aid for women choosing between vaginal birth after cesarean and planned repeat cesarean. J Midwifery Womens Health 2015 Jun 08;60(4):390-400. [doi: 10.1111/jmwh.12298] [Medline: 26059075]
51. Naughton F, Prevost A, Gilbert H, Sutton S. Randomized controlled trial evaluation of a tailored leaflet and SMS text message self-help intervention for pregnant smokers (MiQuit). Nicotine Tob Res 2012 May;14(5):569-577. [doi: 10.1093/ntr/ntr254] [Medline: 22311960]
52. Alnasser AA, Alkhalifa AS, Sathiaseelan A, Marais D. What overweight women want from a weight loss app: a qualitative study on arabic women. JMIR Mhealth Uhealth 2015 May 20;3(2):e41 [FREE Full text] [doi: 10.2196/mhealth.4409] [Medline: 25993907]
53. Beach P. Self-directed online learning: A theoretical model for understanding elementary teachers' online learning experiences. Teaching and Teacher Education 2017 Jan;61:60-72. [doi: 10.1016/j.tate.2016.10.007]
54. Bert F, Gualano MR, Brusaferro S, De Vito E, de Waure C, La Torre G, et al. Pregnancy e-health: a multicenter Italian cross-sectional study on Internet use and decision-making among pregnant women. J Epidemiol Community Health 2013 Dec 01;67(12):1013-1018. [doi: 10.1136/jech-2013-202584] [Medline: 24072743]
55. Danbjørg DB, Wagner L, Kristensen B, Clemensen J. Intervention among new parents followed up by an interview study exploring their experiences of telemedicine after early postnatal discharge. Midwifery 2015 Jun;31(6):574-581. [doi: 10.1016/j.midw.2015.02.007] [Medline: 25765743]
56. Vogels-Broeke M, Daemers D, Budé L, de Vries R, Nieuwenhuijze M. Sources of information used by women during pregnancy and the perceived quality. BMC Pregnancy Childbirth 2022 Feb 08;22(1):109 [FREE Full text] [doi: 10.1186/s12884-022-04422-7] [Medline: 35135487]
57. Martis R, Brown J, McAra-Couper J, Crowther CA. Enablers and barriers for women with gestational diabetes mellitus to achieve optimal glycaemic control - a qualitative study using the theoretical domains framework. BMC Pregnancy Childbirth 2018 Apr 11;18(1):91 [FREE Full text] [doi: 10.1186/s12884-018-1710-8] [Medline: 29642898]
58. Lu Y, Barrett LA, Lin RZ, Amith M, Tao C, He Z. Understanding information needs and barriers to accessing health information across all stages of pregnancy: Systematic review. JMIR Pediatr Parent 2022 Feb 21;5(1):e32235 [FREE Full text] [doi: 10.2196/32235] [Medline: 35188477]
59. Harrison AL, Taylor NF, Frawley HC, Shields N. Women with gestational diabetes mellitus want clear and practical messages from credible sources about physical activity during pregnancy: a qualitative study. J Physiother 2019 Jan;65(1):37-42 [FREE Full text] [doi: 10.1016/j.jphys.2018.11.007] [Medline: 30573442]
60. Muhwava LS, Murphy K, Zarowsky C, Levitt N. Experiences of lifestyle change among women with gestational diabetes mellitus (GDM): A behavioural diagnosis using the COM-B model in a low-income setting. PLoS One 2019 Nov 25;14(11):e0225431 [FREE Full text] [doi: 10.1371/journal.pone.0225431] [Medline: 31765431]
61. Edwards KJ, Bradwell HL, Jones RB, Andrade J, Shawe JA. How do women with a history of gestational diabetes mellitus use mHealth during and after pregnancy? Qualitative exploration of women's views and experiences. Midwifery 2021 Jul;98:102995. [doi: 10.1016/j.midw.2021.102995] [Medline: 33784541]
62. McDonald SD, Sword W, Eryuzlu LE, Biringer AB. A qualitative descriptive study of the group prenatal care experience: perceptions of women with low-risk pregnancies and their midwives. BMC Pregnancy Childbirth 2014 Sep 26;14(1):334 [FREE Full text] [doi: 10.1186/1471-2393-14-334] [Medline: 25258167]

63. McLeish J, Redshaw M. Mothers' accounts of the impact on emotional wellbeing of organised peer support in pregnancy and early parenthood: a qualitative study. BMC Pregnancy Childbirth 2017 Jan 13;17(1):28 [FREE Full text] [doi: 10.1186/s12884-017-1220-0] [Medline: 28086827]
64. Ekezie W, Dallosso H, Saravanan P, Khunti K, Hadjiconstantinou M. Experiences of using a digital type 2 diabetes prevention application designed to support women with previous gestational diabetes. BMC Health Serv Res 2021 Aug 05;21(1):772 [FREE Full text] [doi: 10.1186/s12913-021-06791-9] [Medline: 34348719]
65. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012 May 08;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]

Abbreviations

GDM: gestational diabetes mellitus
mHealth: mobile health
UCD: user-centered design

Edited by A Mavragani; submitted 13.01.22; peer-reviewed by A Sarradon-Eck, NW Cheung; comments to author 03.05.22; revised version received 25.05.22; accepted 25.05.22; published 08.07.22.

Please cite as:
Duan B, Liu Z, Liu W, Gou B
Assessing the Views and Needs of People at High Risk of Gestational Diabetes Mellitus for the Development of Mobile Health Apps: Descriptive Qualitative Study
JMIR Form Res 2022;6(7):e36392
URL: https://formative.jmir.org/2022/7/e36392
doi:10.2196/36392
PMID:35802414

Original Paper

Developing an mHealth App for Empowering Cancer Survivors With Disabilities: Co-design Study

Rachel F Adler[1], PhD; Paulina Morales[1]; Jocelyn Sotelo[1]; Susan Magasi[2,3], PhD

[1]Department of Computer Science, Northeastern Illinois University, Chicago, IL, United States

[2]Department of Occupational Therapy, University of Illinois Chicago, Chicago, IL, United States

[3]Department of Disability and Human Development, University of Illinois Chicago, Chicago, IL, United States

Corresponding Author:
Rachel F Adler, PhD
Department of Computer Science
Northeastern Illinois University
5500 N Saint Louis Ave
Chicago, IL, 60625
United States
Phone: 1 7734424710
Email: r-adler@neiu.edu

Abstract

Background: The transition from active treatment to long-term cancer survivorship leaves the needs of many cancer survivors unaddressed as they struggle with physical, cognitive, psychological, and social consequences of cancer and its treatment. The lack of guidance after treatment has forced cancer survivors to manage long-term effects on their own, which has an impact on their overall health, quality of life, and social participation. Mobile health (mHealth) interventions can be used to promote self-management and evidence-informed education.

Objective: This study aims to design an mHealth app for cancer survivors with disabilities that will offer interventions to improve their quality of life and increase their self-efficacy to manage cancer as a chronic condition.

Methods: We organized 3 co-design workshops with cancer survivors (n=5). These workshops included persona development based on data from 25 interviews with cancer survivors with disabilities; prototype ideation, where we sketched ideas for the prototype; and prototype development, where participants critiqued, and suggested improvements for, the wireframes.

Results: These workshops helped us to define the challenges that cancer survivors with disabilities face as well as important considerations when designing an mHealth app for cancer survivors with disabilities, such as the need for including flexibility, engagement, socialization, and a minimalistic design. We also outline guidelines for other researchers to follow when planning their own co-design workshops, which include allowing more time for discussion among participants, having small participant groups, keeping workshops engaging and inclusive, and letting participants dream big.

Conclusions: Using a co-design process aided us in developing a prototype of an mHealth app for cancer survivors with disabilities as well as a list of guidelines that other researchers can use to develop their own co-design workshops and design their app. Furthermore, working together with cancer survivors ensured that the design team had a deeper sense of empathy toward the target users and kept the focus on our ultimate goal: creating something that cancer survivors would want to use and benefit from. Future work will include usability testing of a high-fidelity prototype based on the results of these workshops.

(JMIR Form Res 2022;6(7):e37706) doi:10.2196/37706

KEYWORDS

user-centered design; co-design; mobile health; mHealth; cancer survivors; disabilities

Introduction

Background

There are an estimated 16.9 million cancer survivors in the United States, and the number is projected to increase to 22.2 million by 2030 [1]. There is tremendous variability in cancer incidence and survivorship globally because of variation in the prevalence of risk factors as well as access to high-quality preventive, screening, and treatment options [2]. Approximately 40% of cancer survivors experience long-term physical,

cognitive, and psychological effects of cancer and its treatment [3,4]. Common long-term effects include pain, fatigue, cognitive effects (eg, cognitive dysfunction or forgetfulness), and psychosocial distress symptoms such as anxiety and depression that can in turn lead to activity limitations and participation restrictions. Long-term effects can negatively affect social participation and health-related quality of life [5]; yet, cancer survivors report that these issues are inadequately addressed within the cancer care system, leaving patients to figure out the impact and long-term management of cancer-related impairments on their own [6]. It is within the purview of the interdisciplinary field of cancer rehabilitation to address the "physical, psychological and cognitive impairments in an effort to maintain or restore function, reduce symptom burden, maximize independence and improve quality of life" [7] of patients with cancer, including those in long-term survivorship. An evidence-informed approach to cancer rehabilitation and survivorship support is the use of self-management interventions that help people learn to deal with the medical, social, and emotional impacts of cancer and its treatment. An emerging body of literature demonstrates that self-management interventions have a positive impact on cancer survivors' physical and psychological well-being and quality of life [8-10].

It should be noted that interdisciplinary cancer rehabilitation and survivorship services are significantly underused by cancer survivors [11]. Estimates suggest that <10% of the people with cancer-related impairments receive services [12,13]. The reason for this underuse is multifactorial, and in-person access to cancer rehabilitation and survivorship services has been identified as a contributing factor; for example, in the United States, there has been a decentralization of cancer care to favor high-quality cancer care in the community rather than in specialty hospitals. Diffusion of rehabilitation services in these community settings has been limited [14]. Indeed, cancer survivorship and rehabilitation programs are clustered in specialty care centers where fewer than half of the people with cancer receive treatment [15]. Access to rehabilitation services may be further exacerbated for cancer survivors with known disabilities because physical, cognitive, and emotional impairments can make care coordination, scheduling, and travel to and from appointments particularly onerous [7]. Furthermore, evidence suggests that oncologists receive limited education about cancer rehabilitation as well as its benefits and indications [16]. This can make service providers unlikely to initiate referrals for rehabilitation services. In addition, most National Cancer Institute–designated cancer centers do not provide information about cancer rehabilitation services on their websites, further restricting access and awareness of these potentially beneficial interventions [17]. This leads to a gap between the people who need support and those who receive it.

Mobile health (mHealth) interventions have been identified as a way to close the gap between the people who need rehabilitation services and those who are able to access them. mHealth interventions have a potentially democratizing impact on access to care because people are able to take in their own hands the tools to monitor and manage their health. mHealth interventions can provide people with tools to promote self-management, symptom monitoring, and evidence-informed education [18], as well as opportunities for peer support and information sharing [19]. Although a variety of mHealth symptom management apps have been developed for cancer survivors, the apps tend to focus on monitoring and managing individual symptoms such as pain and fatigue [10]. As a result, existing mHealth self-management interventions often fail to address the knowledge and skills that cancer survivors with known disabilities need to achieve their goals of creating a meaningful life in spite of the aftereffects of cancer and its treatment [5]; for example, our qualitative research with cancer survivors indicated that many people were uncertain of how and when to communicate about the impact that the aftereffects of cancer had on their abilities to fulfill their social roles and responsibilities [20]. These findings point to an unmet need for strategies to help survivors articulate the impact of cancer on their daily lives and to self-advocate for support and accommodations.

Objectives

To address the unmet needs of cancer survivors who can benefit from ongoing support, we are developing a self-management intervention for cancer survivors with known disabilities called WeCanManage. WeCanManage is conceptualized as a psychoeducational intervention theoretically grounded in Individual and Family Self-Management Theory [21]. The active ingredients of the intervention include instruction and structured practice in (1) the problem-solving–based self-management process of goal setting and action planning supplemented with evidence-informed strategies such as energy conservation and environmental modifications, (2) mindfulness practices [22], and (3) self-advocacy skills. These complementary approaches empower users to build their self-efficacy, the presumed mechanism of change [23], in medical management, role management, and emotional management of cancer as a chronic condition [24]. Consistent with emerging best practices for remote learning, content will be delivered through mobile microlearning modules [25] for a maximum of 10 minutes per day over a 4-week period. Engagement activities will be embedded across the intervention to promote deeper learning and integration into daily life and routines [26,27]. We plan to deliver WeCanManage as an mHealth app, because internet-based self-management interventions provide users with a practical, flexible, and cost-effective alternative to face-to-face interventions [28-32].

For mHealth and rehabilitation tools to be acceptable, accessible, and responsive to the needs of the intended users, Jones et al [33] highlight the importance of proactive engagement of stakeholders, including members of the disability community. However, patient input is rarely included in the development of self-management interventions [8]. To ensure that the WeCanManage platform and design meets the needs and preferences of our target end users, we engaged a cohort of cancer survivors in the design process.

Whereas user-centered design is the process of focusing on users and their needs throughout the stages of the design process [34], co-design, or participatory design, takes this process one step further where designers and users collaborate during the design process [35]. Co-design can be valuable when developing

health-related apps [36,37] and can incorporate engaging techniques such as scenarios or storytelling approaches [38], developing or discussing personas [39-41], reporting on likes or dislikes regarding apps [42], voting on features and solutions [43], sketching out prototypes [38,44,45], redesigning or critiquing prototypes [36,40,46], and answering questionnaires [41,45]. Supplies can be minimal, such as paper, Post-it notes, and posters [45,47,48].

In our work, we primarily use 2 common co-design techniques: persona development and prototyping. Personas are fictional representations of users that help designers to understand and empathize with users [49,50] by challenging them to think "beyond their personal experiences" [39]. Each persona can have a set of characteristics, such as their gender, age, profession, goals, personal history, health issues, technological skills, and hobbies; often, a photograph is included [50]. Personas can include disabilities, which can raise awareness of accessibility and ensure that a design can be used by all [51]. Cocreating personas with people with disabilities can provide even more insight into their experiences [52]. Despite the fact that personas are not real people, they are often developed based on analysis of common themes discovered during user interviews [51,53,54]. Co-design workshops can be effective alongside user interviews [55], particularly because they can be helpful for persona creation. Another common activity during co-design workshops is prototyping. These are often wireframes (low-fidelity sketches) of a potential design that can help lead to higher-fidelity prototypes that are functional and closer to the finished product [56].

To build our intervention and app design, we recruited cancer survivors with disabilities to work with us toward creating a persona (workshop 1) and a prototype of an mHealth app that would empower the community of survivors to self-manage the lifelong effects of their cancer treatment (workshops 2 and 3).

Our research questions (RQs) are as follows:

- RQ1: What are the important design features for an mHealth platform for cancer survivors with disabilities?
- RQ2: What is needed to create an effective co-design environment for this target group?

Methods

Recruitment

To design an mHealth app for cancer survivors with disabilities, we recruited a diverse group of cancer survivors with known disabilities (n=5), whom we call *survivor scientists*, using a citizen scientist approach, to collaborate with our interdisciplinary research and development team [57]. Our inclusion criteria included participants self-identifying as a cancer survivor living with long-term physical, cognitive, or social effects of cancer and its treatment and that they had an established relationship with a faculty member or an institutional or organizational partner. The survivor scientists had experienced breast cancer, head and neck cancer, sarcoma, brain cancer, and leukemia, as well as a range of long-term effects of cancer (including cognitive changes, visual impairments, communication challenges, and decreased functional mobility and fine motor control). All (5/5, 100%) of the survivor scientists were cancer-free and at least 5 years after diagnosis and completion of primary treatment. Several (3/5, 60%) of the survivor scientists are also active in cancer and disability advocacy organizations. The survivor scientists have a variety of professional backgrounds, including social work, graphic design, research support, and rehabilitation medicine with certification in cancer rehabilitation.

Together with the 5 survivor scientists (n=2, 40% men and n=3, 60% women), our team led 3 co-design workshops from July 2021 to October 2021. The first 2 workshops were conducted 2 weeks apart in July, and the third was conducted approximately 3 months later in October to provide enough time for our team to implement wireframes based on feedback from the previous 2 workshops. It should be noted that after the second workshop, we also provided the survivor scientists with the opportunity to continue to work with the research team in smaller groups and assigned them tasks related to their own interests, such as helping with content relating to cancer and disabilities. This work continued after the third workshop as well.

The participants received monetary compensation for their involvement in the co-design process. These workshops included (1) persona development, (2) prototype ideation, and (3) prototype development (wireframes). We developed a semistructured guide for each workshop based on the guide provided in Bradway et al [58]. Refer to Multimedia Appendices 1-3 for our guides for all 3 workshops. The workshops consisted of the design team (8 researchers, faculty, and students from the Departments of Computer Science and Occupational Therapy and Disability Studies from 2 universities) working in combination with our 5 survivor scientists. It should be noted that our last workshop had fewer participants because, of the 5 survivor scientists, 1 (20%) could not attend; in addition, of the 8 members of the design team, 1 (13%) researcher and 2 (25%) undergraduate students who had completed their research experience did not attend; therefore, the total number of participants went from 13 (n=8, 62%, design team members and n=5, 38%, survivor scientists) to 9 (n=5, 56%, design team members and n=4, 44%, survivor scientists). Because of COVID-19–related restrictions, all 3 workshops were held over Zoom, a videoconferencing platform. The workshops were video recorded. The first workshop lasted for 2 hours, and the remaining 2 workshops lasted for 2.5 hours each. This duration is consistent with previous research [40], and we chose an amount of time that would be long enough to accomplish our goals and still keep participants engaged but not too long, particularly because the workshops were conducted on the web, after work, and with participants who had long-term disabilities as a result of cancer.

Within a week after each workshop, the design team met and summarized the main findings from the workshop. These notes were used to inform the development of the workshops that followed; for example, based on the challenges discovered in workshop 1 and the prototype features that the participants mentioned wanting in workshop 2, we developed the low-fidelity prototype that was then shared with the participants in workshop 3 for their feedback.

Although our goal was to create an mHealth app for cancer survivors with disabilities that (1) normalized their experiences as a survivor, (2) taught problem-solving self-management skills, (3) introduced mindfulness-based practices, and (4) addressed self-advocacy skills and disability and survivor rights, the goal of the co-design workshops was to ensure that we would be designing these modules in a way that was usable and engaging to cancer survivors with disabilities.

Ethics Approval

We obtained institutional review board approval from the participating universities in the larger project (University of Illinois Chicago #2020-1067, Northeastern Illinois University #79, and Northwestern University #NUUIC21CC03). The survivor scientists functioned as members of the design team, and no personally identifying data were gathered about them. Ground rules for participation to ensure respect and confidentiality throughout the process were established and agreed to by all participants. All cancer survivors who participated in formative qualitative interviews provided written informed consent before data collection in compliance with approved institutional review board protocols at the collaborating institutions. More details on the interview process and findings from this phase of the study have been reported elsewhere [20].

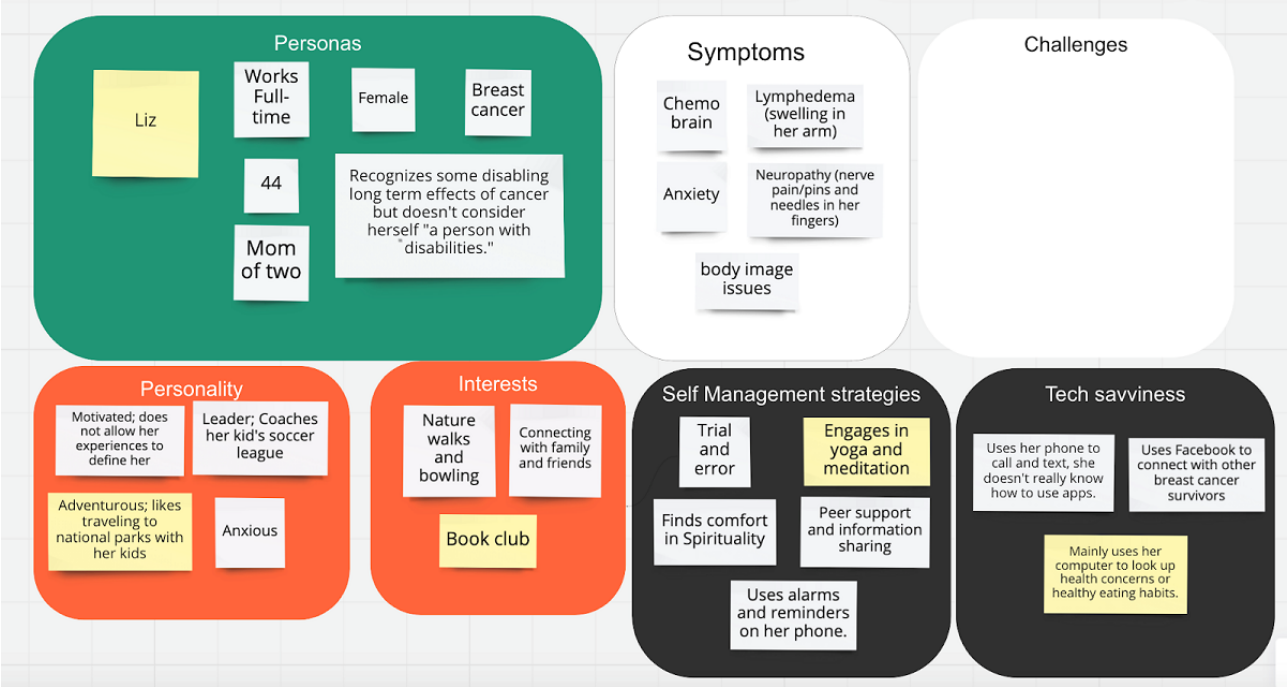
Results

Workshop 1: Persona Development Methodology

The aim of our first workshop was to complete the design of personas to help the design team understand the needs and challenges of cancer survivors with disabilities and build empathy through persona development. The survivor scientists (n=5) attended our first persona development workshop and worked alongside our team of researchers. In advance of this workshop, the research team created 2 personas of cancer survivors with disabilities, using Miro, a visual collaboration platform. These personas were developed based on preliminary analysis of 25 qualitative interviews of breast cancer, sarcoma, and head and neck cancer survivors to foster empathy for the struggles of survivors living with long-term disabilities. Specifically, demographic and clinical data were extracted from the data corpus (demographic surveys, interview transcripts, and field notes) to ground the preliminary personas in the lived experience of survivors of breast cancer, head and neck cancer, and sarcoma. Conceptually, we structured our personas in accordance with the World Health Organization's International Classification of Functioning, Disability, and Health (ICF) framework [59]. The ICF is an internationally recognized biopsychosocial framework that recognizes the dynamic interaction between impairments in body structures and function, activity limitations and participation restrictions, and environmental factors in people's lived experiences of disability. Extracted data included age, gender, race and ethnicity, cancer type, primary symptoms, and impact on social roles and participation. Although our work was informed by the literature [50,51] and the ICF, the categories chosen were based on common themes from our analysis that would make the personas more relatable. We input the data into a persona template provided by Miro to create 2 distinct personas. We decided not to include an image to avoid participants accepting what we provided as a given, but we wanted them to improve upon and complete the profiles. Figure 1 shows a sample of one of the personas developed for the workshop. The challenges section was intentionally left blank for the survivor scientists to fill in that category themselves.

We used 2 breakout rooms to discuss, modify, and complete each persona. Using a Zoom poll, the survivor scientists ranked the most relevant challenges that a cancer survivor was likely to face. Finally, we discussed the ranking results with the survivor scientists.

Figure 1. Sample persona before the workshop.



Workshop 1: Persona Development Results

During these breakout sessions, the survivor scientists critiqued the personas and described the needs and challenges; for example, in breakout room 1, the survivor scientists modified the persona for Liz, a breast cancer survivor, by providing more details making her even more relatable (Figure 2). Liz is now an African American woman, with a partner, and children aged 5 and 14 years. She works in retail, which requires prolonged standing and heavy lifting and has limited insurance tied to her job. The survivors discussed challenges, some of which include financial pressures, long-term symptoms interfering in her life, isolation because of lack of peer support, concern regarding the emotional toll on her children and relationship with her partner, and body image concerns. The survivor scientists felt that these details helped to encapsulate the complexity of living with the long-term effects of breast cancer while juggling multiple roles and responsibilities.

In breakout room 2, the survivor scientists discussed Solomon, a man aged 56 years, who is divorced and is on the fence regarding looking for a romantic partner because of body image issues as a result of head and neck cancer. His interests originally included karaoke and barbecues with his family most weekends. However, the survivor scientists discussed how he may have once enjoyed karaoke and eating with his friends before cancer, but now singing would be difficult, and eating with friends would cause anxiety because of his slower eating speed. The results of both breakout sessions showed that the survivor scientists deemed it important to acknowledge the high levels of anxiety caused by role loss, fear of recurrence, financial toxicity, and the strain that cancer continues to impose on relationships with family and friends. In addition, the survivor scientists wanted to highlight the challenges that cancer survivors encounter when re-establishing leisure roles and maintaining their employment because of physical and cognitive limitations, as well as the impact that role loss has on a person's identity and sense of self. Use of compensatory strategies and adaptation were also discussed.

After the breakout sessions, participants reconvened in the larger group and shared their enhanced personas. Using the challenges identified in both breakout sessions, we had participants select the 3 challenges that they felt were most important using Zoom's polling feature. The results, presented in Figure 3, reveal that isolation, financial pressures, and anxiety or depression were the biggest challenges. These were closely followed by participation in work and leisure, as well as social roles. It should be noted that 6 participants completed the Zoom poll because, of the 8 members of the design team, 1 (13%), a researcher, is also a breast cancer survivor and provided feedback alongside the survivor scientists based on her own experiences as a survivor.

Figure 2. Sample persona after the workshop.

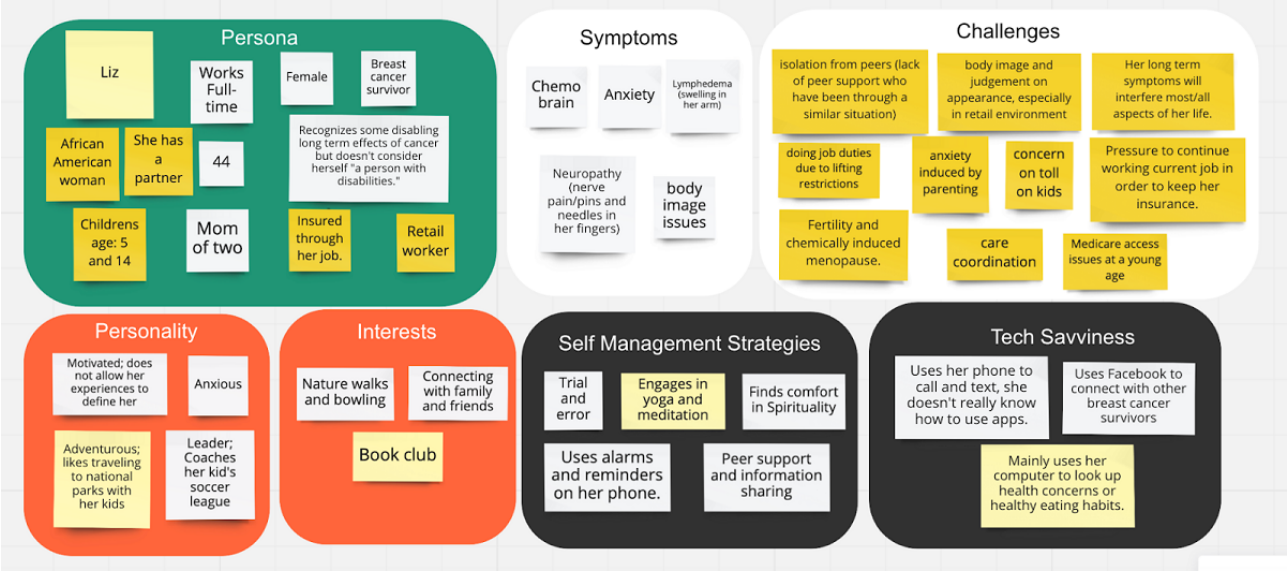
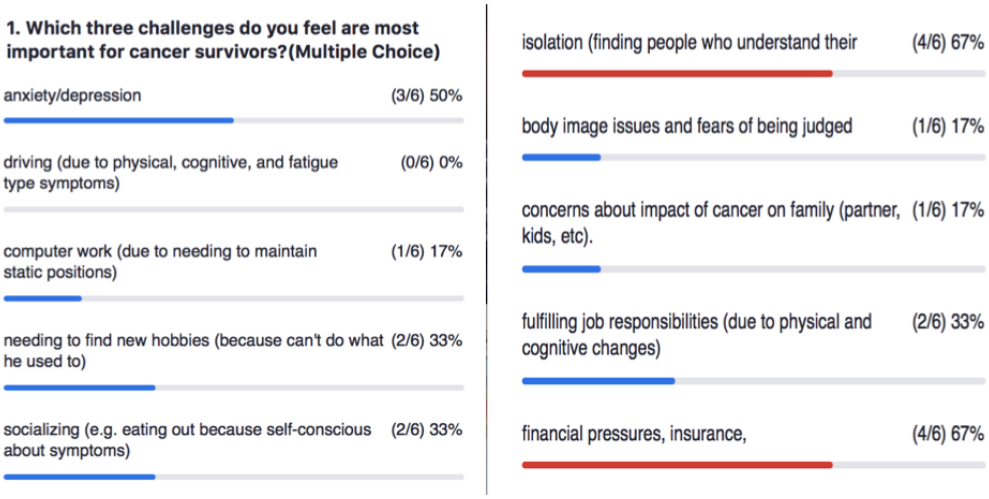


Figure 3. Ranking the results of the persona workshop.



Workshop 2: Prototype Ideation Methodology

The aim for the second workshop was to determine what features cancer survivors with long-term disabilities like in current apps. We first asked the survivor scientists to list (in the Zoom chat) the apps they like in general as well as the apps they dislike to encourage them to think about design. We then directed them to a shared Google Sheets document where they each had a column with their names to list the features they liked in apps and the features they did not like. We generated a word cloud using a free word cloud generator [60] with the results. Next, we divided participants into 2 breakout rooms (intentionally placing them with different people than at the previous workshop as well as keeping in mind diversity in terms of disability and gender). We showed each group different potential content and asked them how they would deliver this content so that it would be inclusive and engaging to cancer survivors. Using the Miro board, we also asked them to provide us with sketch ideas for how they envisioned the app. Although there were design limitations in terms of technical capabilities and budget on the development side, we told the survivor scientists to think in terms of an open universe where anything is possible because we did not want to limit their thinking. Once merged back into the larger group, each group presented their work to the other group. Using features that both groups felt were important in the design of the app, we had the participants rank the features that were most important to them using Slido, a live polling platform (Cisco Systems, Inc).

Workshop 2: Prototype Ideation Results

After the participants had listed their likes and dislikes in apps, we generated a word cloud (Figure 4). The features that the survivors liked included networking, ease of use, and simplicity, whereas the features that the survivors disliked included lack of user-friendliness, too many notifications and advertisements, and poor navigation.

Figure 4. (A) Participants' likes in apps. (B) Participants' dislikes in apps.

After we divided the survivors into 2 groups and asked them to think about ideas for sketches (using Miro) that they envisioned for the app, one of the groups asked the workshop leader to sketch a learning pathway to be able to provide ease of navigation and engagement (Figure 5A). The participants also mentioned that the app should include photos of people from different backgrounds in terms of race, disability, and age. In addition, they discussed including the use of animation, which can help those with attention disorders, icons instead of only text for those with cognitive disabilities, and customizable color choices and simpler layouts for those with visual impairments. In the second breakout room, of the 5 survivor scientists, 1 (20%) created a paper mockup that she felt would look engaging, and based on her drawing which she held up to the camera, along with feedback from the rest of the breakout room participants, workshop leaders sketched it using the Miro board (Figure 5B). The survivor scientists felt that the content should be engaging, with videos as well as content that was easy to access. Furthermore, they discussed that apps for people with disabilities tend to be poorly designed and "ugly" and that good design would alienate no one. They were very happy to be part of the process of designing the app.

After gathering all the information that the survivors added to the sticky notes in Miro, the word cloud, and the sketches, we asked the survivors to rank the features from most important to least important (Figure 6). The top 3 features that were most important to the survivor scientists were customizability and personalization, ease of navigation and searchable content, and accessibility. This was followed by minimal design.

Figure 5. Workshop sketches. (A) Breakout room 1. (B) Breakout room 2.

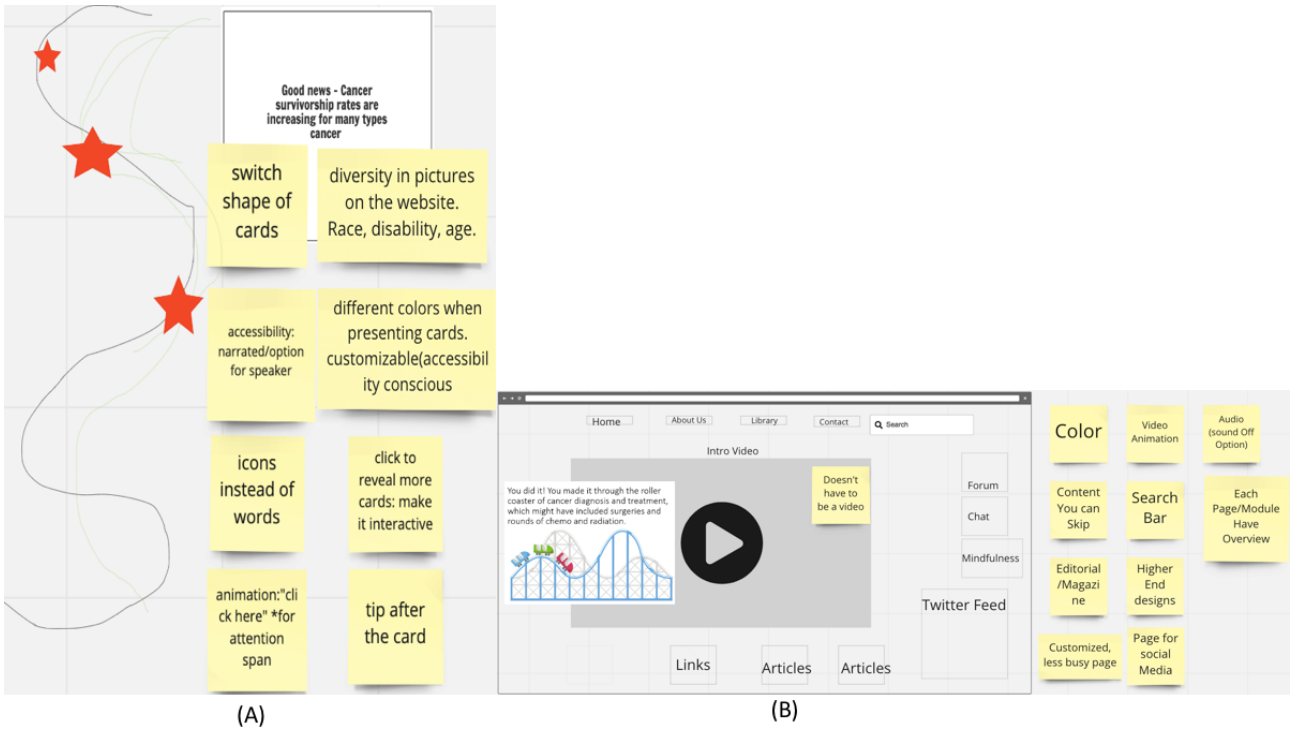
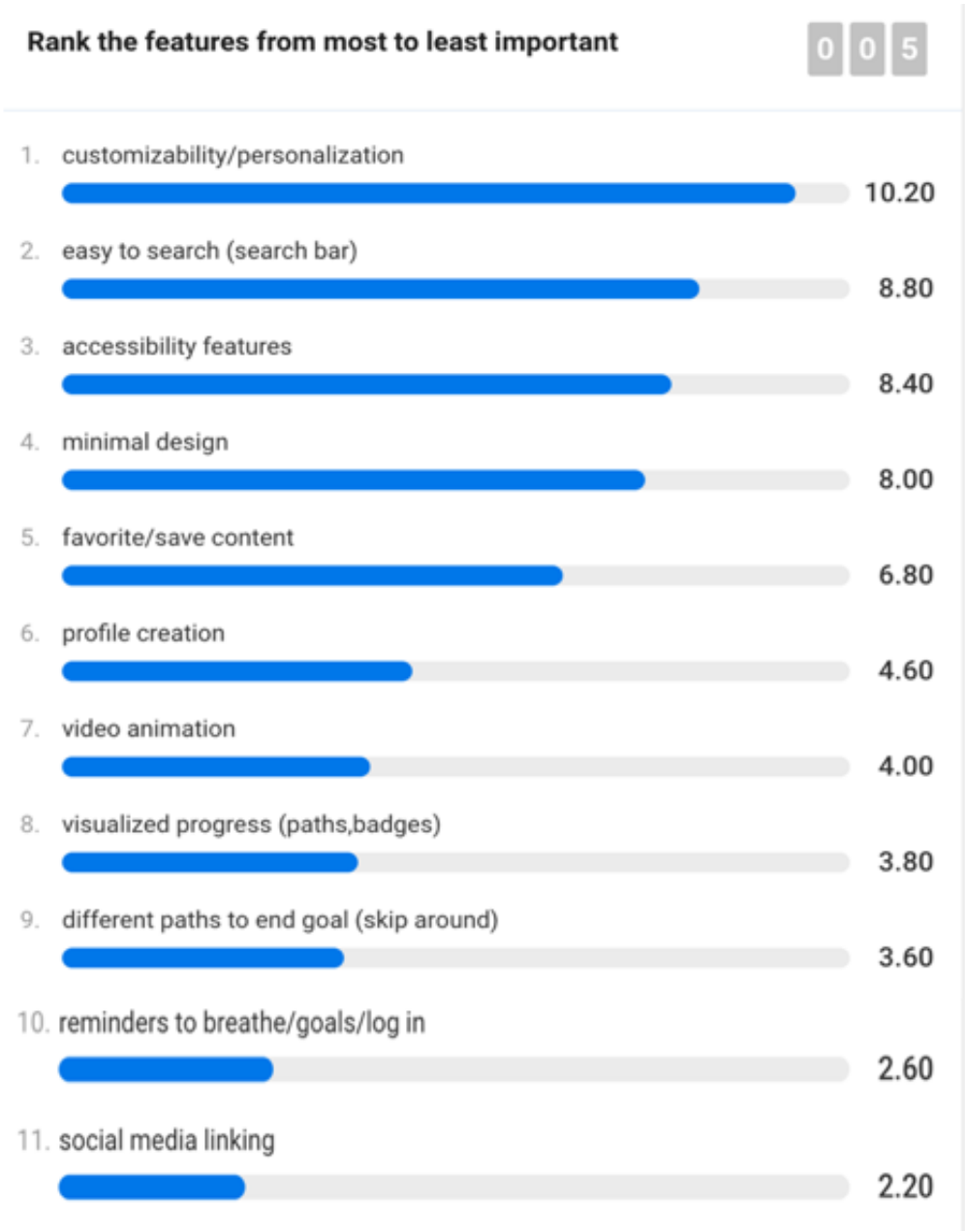


Figure 6. Features ranked from most important to least important.

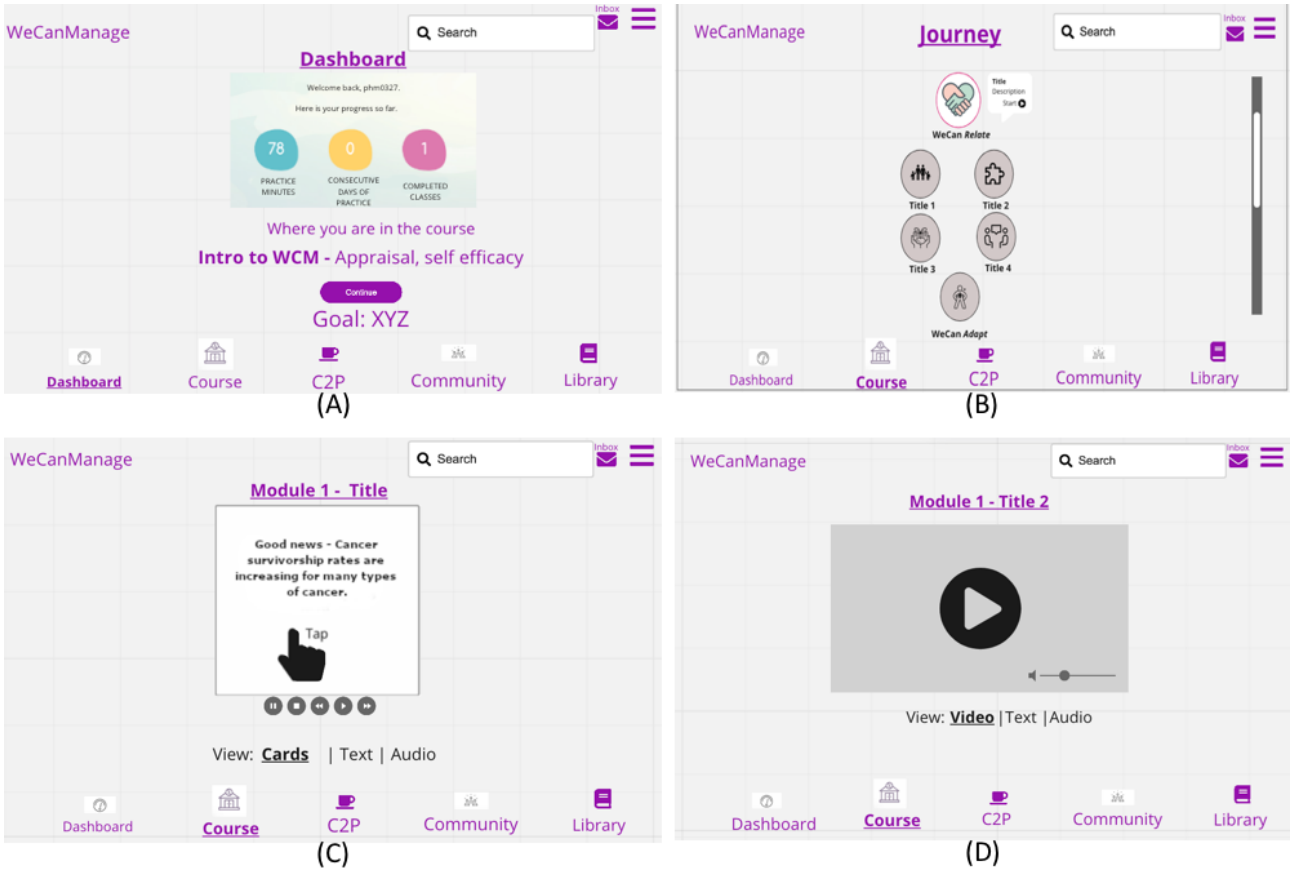


Workshop 3: Prototype Development Methodology

In our third workshop, we wanted to find out what the participants thought of the low-fidelity prototype that was created to address the challenges raised in workshop 1 (eg, isolation) and the features they mentioned wanting in workshop 2 (eg, customization). We also wanted to learn what improvements could be made. We divided the researchers and survivor scientists into 2 breakout rooms to critique the wireframes. After coming together in the larger Zoom room, both groups presented their thoughts on the design. We also led a conversation on whether to include a community forum or a feed-like aspect in the app. As shown in Figure 7A, the first screen presented to the survivor scientists was the Dashboard. To meet the survivor scientists' needs for personalization, we incorporated a section for a personalized goal that the user would type in and be working toward achieving. This would be visible on the landing page after logging in. We also matched the pathway that the survivors requested from the previous sketch (Figure 5A) in the Course tab (Figure 7B). Connect to Peers (C2P; the C2P tab) was provided as a space for users to build their network and connect with other cancer survivors through direct messaging, with the Community section being a place where users can share their experience and support one another through forums. The Library section would contain helpful resources. When logging into a course, the users will be offered a series of microlessons subdivided into 4 content modules: *WeCanRelate* seeks to normalize and validate their experiences as survivors, *WeCanAdapt* focuses on problem-solving–based self-management skill building to promote self-efficacy, *WeCanBreathe* introduces mindfulness-based practices, and *WeCanSpeakUp* addresses self-advocacy skills as well as disability and survivor rights. Users will also be given the opportunity to select the mode of presentation they prefer, which includes formatting the participative cards (or video) into text or audio (Figures 7C and Figures 7D). This will allow users to

comfortably access the content in the mode that best supports their learning style and access needs. We also provide knowledge checks throughout each course module.

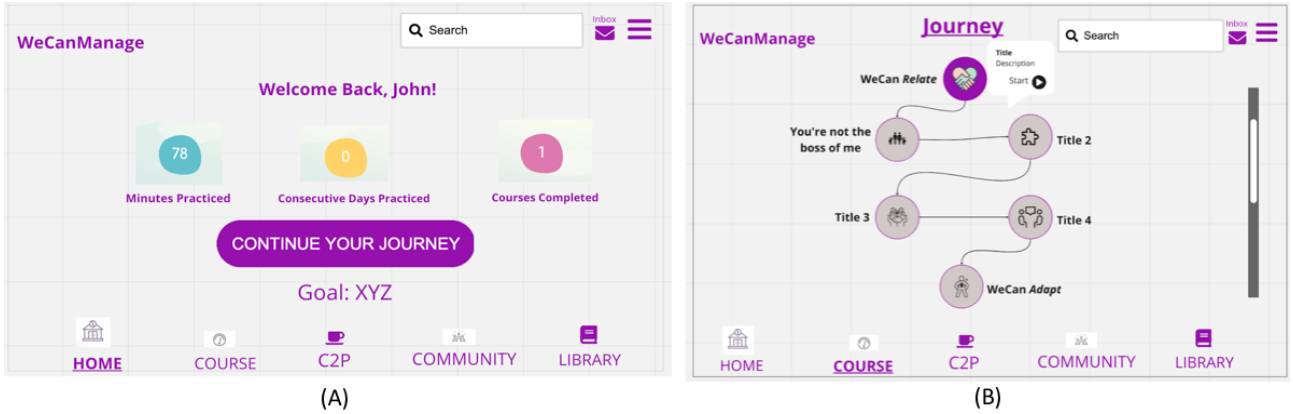
Figure 7. Wireframes. (A) Dashboard page. (B) Course navigation. (C) Cards module. (D) Video module. C2P: Connect to Peers; WCM: WeCanManage.



Workshop 3: Prototype Development Results

On the basis of the results of the workshop, the survivor scientists liked the wireframes but had concerns with the main Dashboard screen having too much information and the Journey screen having too much white space, and they did not like the terminology used, such as the words dashboard, course, and quiz. Figures 8A and Figures 8B show sketches of the home screen and Journey page, respectively, revised after discussions with the survivor scientists.

Figure 8. (A) Updated home screen. (B) Updated Journey page. C2P: Connect to Peers.



Although we had a larger group discussion on whether to include community forums or a feed-like design (as available on many social networking sites), the participants did not have a strong preference for one format over the other. The survivor scientists liked that there was a way to contact peers on the site either directly or through the forum or feed option.

Discussion

Principal Findings

We formulated our RQs to discover what design features are important for an mHealth platform for cancer survivors with disabilities (RQ1) and what would be needed to create an effective co-design environment for this target group (RQ2).

The 3 workshops provided us with a strong foundation to move forward with well-defined goals for designing an mHealth app for cancer survivors with disabilities. Our results indicate that participants wanted an mHealth app that encompasses (1) flexibility, (2) engagement, (3) socialization, and (4) a minimalistic design.

Our first finding is that participants wanted *flexibility* in the app. The top-ranked feature that the survivor scientists wanted in workshop 2 was customization and personalization. Participants wanted an app that can be easily customized to their preferences and their access needs. They also wanted accessibility features and an app that was easily searchable. Therefore, in our design for workshop 3, to suit the content to the learning needs of the individual we provided the option to view content (whether video or cards) in an audio-only or text-only format, providing flexibility to the user.

Furthermore, the survivor scientists, in workshop 2, reported liking features that were participative. Users are more receptive to mHealth apps when they are *engaging* [61]. The survivor scientists recommended the use of color, video, and interaction. They also liked the idea of a learning pathway for the course content. We incorporated this feedback in the prototype shown to the participants in workshop 3. As noted in the study by Jessen et al [40], participants liked receiving recognition from an mHealth app for finishing tasks and setting up their own goals and subgoals. Similarly, we found that participants liked being able to input their goals and view them on the main screen. In addition, we discovered that participants enjoyed ways to determine their progress through quiz-like activities and a journey section that kept track of their course pathway.

The participants reported in workshop 1 that one of the biggest challenges was isolation. Although engaging features can be beneficial in an mHealth app, communicating with others going through similar experiences has been found to be a key feature in supporting mHealth apps [40]. Therefore, we found that survivor scientists wanted *socialization* as well. In our design, we showed participants in workshop 3 the different ways in which users could interact with each other. Participants liked having a way to connect directly with peers going through similar experiences (the C2P section) and benefit from being able to ask questions or post comments to the community of cancer survivors (the Community section). This enhanced their socialization experience.

Although including goals, progress, and opportunities for socialization are effective techniques, similar to other studies we found that the app should be designed using a *minimalistic design* approach [62]. In workshop 2, our participants reported preferring minimal design. This concept was supported again in workshop 3, when participants reported not liking when there was too much content on the screen. They wanted a clean and simple interface that would engage them and allow them to easily comprehend where to go. This was most noted on the home screen, which they felt was initially too cluttered.

Comparison With Prior Work

The study was able to extend previous research by addressing RQ1 and providing tips on creating mHealth apps through what we learned in our co-design workshops (flexibility, engagement, socialization, and a minimalistic design). Furthermore, our work extends previous work on planning successful co-design workshops for health-related apps [36,37], particularly for internet-based co-design workshops. Näkki and Antikainen [63] found that using web-based tools can make it easier and cheaper to include users as co-designers. People with disabilities benefit from being able to work in remote settings, which therefore promotes a more inclusive environment [64]. The Zoom-based co-design workshops, despite their limitations, provided flexibility for the survivor scientists in terms of geographical limitations and participants with disabilities being comfortable and able to work out of their own homes. In our work, to address RQ2, we found that successful co-design workshops should be engaging, inclusive, provide more time for participants to speak up, use smaller participant groups, and let participants think big (provide a universe where anything is possible).

Researchers should keep the co-design process as *engaging* as possible while keeping in mind their target users, the technology, and any limitations [40,58]. Although keeping it engaging is true in face-to-face workshops as well, this can be even more pronounced in a web-based environment where “Zoom fatigue” [65] can be prevalent. Engaging techniques can be included in numerous ways. In our case, we incorporated many web-based tools such as Zoom’s breakout rooms, poll and chat features, Slido, word clouds, and Miro.

In addition, to help keep it engaging, a significant portion of time should be allotted *for participants to speak up*. Arsand and Demiris [38] discuss allocating sufficient time for several meetings with users to allow time for their creative ideas. In our first workshop we had many more activities planned that we did not get to complete because we did not allocate enough time for participation of the survivor scientists. We made adjustments by adding the needed time and not finishing everything as planned; for example, we had allocated 20 minutes for the breakout session in workshop 1, which we modified to 30 minutes while in the breakout rooms. We also provided 10 minutes for the survivor scientists to discuss their challenges, but this was not enough; therefore, we continued to let them talk and decided not to complete all the topics that came later in the list included in the workshop guide. Staying flexible allowed us to provide additional time to the participants. Additional time for conversation may be particularly important in a web-based format, which limits some of the informal communication and connections inherent in in-person workshops. Learning from this, we set the time allocated for the next 2 workshops to 2.5 hours (instead of 2 hours) and provided extra time for the breakout rooms and survivor scientist discussion (refer to Multimedia Appendices 2 and 3). Finally, we reduced the amount of time that we presented to make it more engaging and participatory. Providing survivors with additional time and opportunities for feedback and sharing of ideas in the workshops was crucial because many of their ideas helped to shape the design of the app.

For participants to be comfortable speaking up and participating, previous work has shown that breaking out into *smaller working groups* can be helpful [66,67]. Using Zoom breakout rooms in web-based workshops can foster engagement and reduce “Zoom

fatigue" [67]. We only included 5 survivor scientists, but, with the addition of the researchers, workshops 1 and 2 had 13 participants each. Although the breakout rooms were very helpful in facilitating the division of the team into smaller groups, after the second workshop we also recruited survivor scientists to work individually with a researcher on content development or design; for example, of the 5 survivor scientists, 1 (20%) came up with the idea of modifying WeCanManage course content into modules whose titles will all begin with WeCan; for example, WeCanRelate. Another survivor scientist offered to work on content development for specific cancer symptom management resources, whereas another will focus on more of the disability aspects. Having survivor scientists self-select into areas based on their own skills and interests enabled them to be even more efficient and feel useful in the co-design process.

Another approach to encourage participants to speak up is through creating an *inclusive environment*. Creating a comfortable and inclusive environment in co-design workshops is one of the lessons found in the study by Bradway et al [58]. Similarly, we wanted the environment to be welcoming to all. Therefore, we began the first workshop with time for rapport building and informal conversation before establishing ground rules aimed at creating an environment of trust and respect. The ground rules included emphasizing that everything discussed in the workshops was confidential, and all opinions were valid. We reminded participants of this at the beginning of the next 2 workshops. Although there may be differing opinions, it is important to show respect for all the different views, particularly because participants can be very passionate about an app that would be of direct benefit to them. Although there are a limited number of inclusive apps [68], ensuring respect for co-designers with their own particular backgrounds and disabilities can lead to the creation of apps that are more inclusive because the co-designers keep in mind their own personal experiences and disabilities in the design process.

In addition to making co-design participants feel comfortable sharing their ideas, it is important to allow them "the latitude to dream big and imagine a best-case scenario with no constraints" [69]. The concept of *dreaming big* and providing a universe where anything is possible can promote creativity and facilitate learning about what users find important in a design [69-72]. To encourage participants to speak up, we did not limit their ideas by mentioning technical or budgetary considerations but encouraged them to dream big. Particularly in a web-based co-design workshop, dreaming big could be even more difficult for participants with limited technical skills. However, we saw how important and useful this was throughout the design process; for example, although the learning pathway sketch (Figure 5A) was how the participants were picturing the design to be, after speaking with our development team we toned it down to a parallel tile-based design (Figure 7B). Later, this was modified after further conversations with the survivor scientists, using a compromise approach that worked for the development team as well as the design team (Figure 8B). Having co-designers who are not limited to the practical aspects of design can help to encourage creative thinking and lead to a more meaningful design.

Limitations

We encountered limitations in the co-design process because of technical difficulties and pandemic-related hurdles. Because of COVID-19–related restrictions, we conducted all workshops with survivor scientists over Zoom. Although this provided flexibility in terms of location and scheduling, we were limited in that we could not use Post-it notes, posters, and paper during the workshops. We conducted most of our designing of personas and prototypes through Miro. As the technology can be challenging for first-time users, we chose to have 1 person from our design team lead the Miro board activities in each breakout session.

Conclusions

The results from the co-design workshops provided our research team with a deeper level of empathy for our target users and a better understanding of long-term survivorship challenges and needs for an mHealth app. The collaborative development aided in creating a shared vision of target users among researchers and survivor scientists, while being an engaging co-design experience. Future work will continue to include survivor scientists in the design process as we create a high-fidelity prototype and conduct usability testing, which will be followed by the implementation of the app.

Acknowledgments

This research was supported by the National Cancer Institute of the National Institutes of Health (U54CA202995, U54202997, and U54CA203000). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Additional support was provided by Northeastern Illinois University's Student Center for Science Engagement, Committee on Organized Research, and College of Business and Technology; the US Department of Education Title III (P031C160209); and the Summer Research Opportunities Program of the University of Illinois Chicago. The authors would also like to acknowledge contributions from the following team members: David Victorson, Tamara Hamlish, Cassandra Winters, Beija Teolis, Lauren Dimayuga, Bruriah Horowitz, Hilary Marshall, Christina Sauer, John Abbate, Christine Adley, Linda Cassady, Peter Citro, and Deborah Friedman.

Conflicts of Interest

None declared.

Multimedia Appendix 1
A guide for workshop 1: persona development.
[DOCX File , 19 KB - formative_v6i7e37706_app1.docx]

Multimedia Appendix 2
A guide for workshop 2: prototype ideation.
[DOCX File , 18 KB - formative_v6i7e37706_app2.docx]

Multimedia Appendix 3
A guide for workshop 3: prototype development.
[DOCX File , 16 KB - formative_v6i7e37706_app3.docx]

References

1. Cancer Statistics. National Cancer Institute. URL: https://www.cancer.gov/about-cancer/understanding/statistics [accessed 2022-02-23]
2. American Cancer Society. 2018. URL: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/global-cancer-facts-and-figures/global-cancer-facts-and-figures-4th-edition.pdf [accessed 2022-05-25]
3. Institute of Medicine and National Research Council. From Cancer Patient to Cancer Survivor: Lost in Transition. Washington, DC, USA: National Academies Press (US); 2005.
4. Stein KD, Syrjala KL, Andrykowski MA. Physical and psychological long-term and late effects of cancer. Cancer 2008 Jun 01;112(11 Suppl):2577-2592 [FREE Full text] [doi: 10.1002/cncr.23448] [Medline: 18428205]
5. Newman RM, Alfano CM, Radomski MV, Pergolotti M, Wolf TJ, Sleight AG, et al. Catalyzing research to optimize cancer survivors' participation in work and life roles. OTJR (Thorofare N J) 2019 Oct;39(4):189-196. [doi: 10.1177/1539449219844749] [Medline: 31046601]
6. Boykoff N, Moieni M, Subramanian SK. Confronting chemobrain: an in-depth look at survivors' reports of impact on work, social networks, and health care response. J Cancer Surviv 2009 Dec;3(4):223-232 [FREE Full text] [doi: 10.1007/s11764-009-0098-x] [Medline: 19760150]
7. Silver JK, Raj VS, Fu JB, Wisotzky EM, Smith SR, Kirch RA. Cancer rehabilitation and palliative care: critical components in the delivery of high-quality oncology services. Support Care Cancer 2015 Dec;23(12):3633-3643. [doi: 10.1007/s00520-015-2916-1] [Medline: 26314705]
8. Cuthbert CA, Farragher JF, Hemmelgarn BR, Ding Q, McKinnon GP, Cheung WY. Self-management interventions for cancer survivors: a systematic review and evaluation of intervention content and theories. Psychooncology 2019 Nov;28(11):2119-2140. [doi: 10.1002/pon.5215] [Medline: 31475766]
9. Howell D, Harth T, Brown J, Bennett C, Boyko S. Self-management education interventions for patients with cancer: a systematic review. Support Care Cancer 2017 Apr;25(4):1323-1355. [doi: 10.1007/s00520-016-3500-z] [Medline: 28058570]
10. Hernandez Silva E, Lawler S, Langbecker D. The effectiveness of mHealth for self-management in improving pain, psychological distress, fatigue, and sleep in cancer survivors: a systematic review. J Cancer Surviv 2019 Feb;13(1):97-107. [doi: 10.1007/s11764-018-0730-8] [Medline: 30635865]
11. Raj VS, Pugh TM, Yaguda SI, Mitchell CH, Mullan SS, Garces NS. The who, what, why, when, where, and how of team-based interdisciplinary cancer rehabilitation. Semin Oncol Nurs 2020 Feb;36(1):150974. [doi: 10.1016/j.soncn.2019.150974] [Medline: 31955923]
12. Pergolotti M, Deal AM, Lavery J, Reeve BB, Muss HB. The prevalence of potentially modifiable functional deficits and the subsequent use of occupational and physical therapy by older adults with cancer. J Geriatr Oncol 2015 May;6(3):194-201 [FREE Full text] [doi: 10.1016/j.jgo.2015.01.004] [Medline: 25614296]
13. Cheville AL, Mustian K, Winters-Stone K, Zucker DS, Gamble GL, Alfano CM. Cancer rehabilitation: an overview of current need, delivery models, and levels of care. Phys Med Rehabil Clin N Am 2017 Feb;28(1):1-17. [doi: 10.1016/j.pmr.2016.08.001] [Medline: 27912990]
14. Alfano CM, Ganz PA, Rowland JH, Hahn EE. Cancer survivorship and cancer rehabilitation: revitalizing the link. J Clin Oncol 2012 Mar 20;30(9):904-906. [doi: 10.1200/JCO.2011.37.1674] [Medline: 22355063]
15. Frosch ZA, Illenberger N, Mitra N, Boffa DJ, Facktor MA, Nelson H, et al. Trends in patient volume by hospital type and the association of these trends with time to cancer treatment initiation. JAMA Netw Open 2021 Jul 01;4(7):e2115675 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.15675] [Medline: 34241630]
16. Rosario-Concepción RA, Calderín YB, Aponte CL, López-Acevedo CE, Sepúlveda-Irrizarry FL. Oncologists' attitude and knowledge about cancer rehabilitation. PM R 2021 Dec;13(12):1357-1361. [doi: 10.1002/pmrj.12547] [Medline: 33389793]
17. Silver JK, Raj VS, Fu JB, Wisotzky EM, Smith SR, Knowlton SE, et al. Most National Cancer Institute-designated cancer center websites do not provide survivors with information about cancer rehabilitation services. J Cancer Educ 2018 Oct;33(5):947-953. [doi: 10.1007/s13187-016-1157-4] [Medline: 28064402]

18. Ramey L, Osborne C, Kasitinon D, Juengst S. Apps and mobile health technology in rehabilitation: the good, the bad, and the unknown. Phys Med Rehabil Clin N Am 2019 May;30(2):485-497. [doi: 10.1016/j.pmr.2018.12.001] [Medline: 30954161]
19. Magasi S, Banas J, Horowitz B, Reis JP, The K, Wilson T, et al. WeCanConnect: development of a community-informed mHealth tool for people with disabilities and cancer. Prog Community Health Partnersh 2019;13(5):49-59 [FREE Full text] [doi: 10.1353/cpr.2019.0038] [Medline: 31378735]
20. Magasi S, Marshall HK, Winters C, Victorson D. Cancer survivors' disability experiences and identities: a qualitative exploration to advance cancer equity. Int J Environ Res Public Health 2022 Mar 06;19(5):3112 [FREE Full text] [doi: 10.3390/ijerph19053112] [Medline: 35270802]
21. Grey M, Schulman-Green D, Knafl K, Reynolds NR. A revised self- and family management framework. Nurs Outlook 2015;63(2):162-170. [doi: 10.1016/j.outlook.2014.10.003] [Medline: 25771190]
22. Bishop SR, Lau M, Shapiro S, Carlson L, Anderson ND, Carmody J, et al. Mindfulness: a proposed operational definition. Clin Psychol Sci Pract 2004 Sep;11(3):230-241. [doi: 10.1093/clipsy.bph077]
23. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/S15324796ABM2601_01] [Medline: 12867348]
24. Grady PA, Gough LL. Self-management: a comprehensive approach to management of chronic conditions. Am J Public Health 2014 Aug;104(8):e25-e31. [doi: 10.2105/AJPH.2014.302041] [Medline: 24922170]
25. Mohammed GS, Wakil K, Nawroly SS. The effectiveness of microlearning to improve students' learning ability. Int J Educ Res Rev 2018 Apr 16;3(3):32-38. [doi: 10.24331/ijere.415824]
26. Jahnke I, Lee YM, Pham M, He H, Austin L. Unpacking the inherent design principles of mobile microlearning. Tech Know Learn 2019 May 23;25(3):585-619. [doi: 10.1007/s10758-019-09413-w]
27. Khurgin A. Will the real microlearning please stand up? Association for Talent Development. 2022 Jun 19. URL: https://www.td.org/atd-blog/will-the-real-microlearning-please-stand-up [accessed 2016-11-29]
28. Arif MJ, El Emary IM, Koutsouris DD. A review on the technologies and services used in the self-management of health and independent living of elderly. Technol Health Care 2014;22(5):677-687. [doi: 10.3233/THC-140851] [Medline: 25134962]
29. Heapy AA, Higgins DM, Cervone D, Wandner L, Fenton BT, Kerns RD. A systematic review of technology-assisted self-management interventions for chronic pain: looking across treatment modalities. Clin J Pain 2015 Jun;31(6):470-492. [doi: 10.1097/AJP.0000000000000185] [Medline: 25411862]
30. Lyden JR, Zickmund SL, Bhargava TD, Bryce CL, Conroy MB, Fischer GS, et al. Implementing health information technology in a patient-centered manner: patient experiences with an online evidence-based lifestyle intervention. J Healthc Qual 2013;35(5):47-57. [doi: 10.1111/jhq.12026] [Medline: 24004039]
31. Pereira K, Phillips B, Johnson C, Vorderstrasse A. Internet delivered diabetes self-management education: a review. Diabetes Technol Ther 2015 Jan;17(1):55-63. [doi: 10.1089/dia.2014.0155] [Medline: 25238257]
32. Iacobelli F, Adler RF, Buitrago D, Buscemi J, Corden ME, Perez-Tamayo A, et al. Designing an mHealth application to bridge health disparities in Latina breast cancer survivors: a community-supported design approach. Design Health (Abingdon) 2018;2(1):58-76 [FREE Full text] [doi: 10.1080/24735132.2018.1452871] [Medline: 30506017]
33. Jones M, DeRuyter F, Morris J. The digital health revolution and people with disabilities: perspective from the United States. Int J Environ Res Public Health 2020 Jan 07;17(2):381 [FREE Full text] [doi: 10.3390/ijerph17020381] [Medline: 31936006]
34. Norman DA, Draper SW. User Centered System Design; New Perspectives on Human-Computer Interaction. Hillsdale, NJ, USA: L. Erlbaum Associates; 1986.
35. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008 Mar;4(1):5-18. [doi: 10.1080/15710880701875068]
36. LaMonica HM, Davenport TA, Burns J, Cross S, Hodson S, Veitch J, et al. Technology-enabled mental health service reform for open arms - veterans and families counselling: participatory design study. JMIR Form Res 2019 Sep 19;3(3):e13662 [FREE Full text] [doi: 10.2196/13662] [Medline: 31538937]
37. Woods L, Cummings E, Duff J, Walker K. Design thinking for mHealth application co-design to support heart failure self-management. Stud Health Technol Inform 2017;241:97-102. [Medline: 28809190]
38. Arsand E, Demiris G. User-centered methods for designing patient-centric self-help tools. Inform Health Soc Care 2008 Sep;33(3):158-169. [doi: 10.1080/17538150802457562] [Medline: 18850399]
39. Amann J, Fiordelli M, Brach M, Bertschy S, Scheel-Sailer A, Rubinelli S. Co-designing a self-management app prototype to support people with spinal cord injury in the prevention of pressure injuries: mixed methods study. JMIR Mhealth Uhealth 2020 Jul 09;8(7):e18018 [FREE Full text] [doi: 10.2196/18018] [Medline: 32673241]
40. Jessen S, Mirkovic J, Ruland CM. Creating gameful design in mHealth: a participatory co-design approach. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e11579 [FREE Full text] [doi: 10.2196/11579] [Medline: 30552080]
41. VanHeerwaarden N, Ferguson G, Abi-Jaoude A, Johnson A, Hollenberg E, Chaim G, et al. The optimization of an eHealth solution (thought spot) with transition-aged youth in postsecondary settings: participatory design research. J Med Internet Res 2018 Mar 06;20(3):e79 [FREE Full text] [doi: 10.2196/jmir.8102] [Medline: 29510970]

42. Schmitt Z, Yarosh S. Participatory design of technologies to support recovery from substance use disorders. Proc ACM Hum Comput Interact 2018 Nov;2(CSCW):1-27. [doi: 10.1145/3274425]
43. Groussard PY, Pigot H, Giroux S. From conception to evaluation of mobile services for people with head injury: a participatory design perspective. Neuropsychol Rehabil 2018 Jul;28(5):667-688. [doi: 10.1080/09602011.2015.1117499] [Medline: 26679473]
44. Ospina-Pinillos L, Davenport T, Mendoza Diaz A, Navarro-Mancilla A, Scott EM, Hickie IB. Using participatory design methodologies to co-design and culturally adapt the Spanish version of the mental health eClinic: qualitative study. J Med Internet Res 2019 Aug 02;21(8):e14127 [FREE Full text] [doi: 10.2196/14127] [Medline: 31376271]
45. Warren LR, Harrison M, Arora S, Darzi A. Working with patients and the public to design an electronic health record interface: a qualitative mixed-methods study. BMC Med Inform Decis Mak 2019 Dec 03;19(1):250 [FREE Full text] [doi: 10.1186/s12911-019-0993-7] [Medline: 31795998]
46. Trettin B, Danbjørg DB, Andersen F, Feldman S, Agerskov H. Development of an mHealth app for patients with psoriasis undergoing biological treatment: participatory design study. JMIR Dermatol 2021 May 10;4(1):e26673. [doi: 10.2196/26673]
47. Nielsen C, Agerskov H, Bistrup C, Clemensen J. User involvement in the development of a telehealth solution to improve the kidney transplantation process: a participatory design study. Health Informatics J 2020 Jun;26(2):1237-1252 [FREE Full text] [doi: 10.1177/1460458219876188] [Medline: 31566460]
48. O'Brien N, Heaven B, Teal G, Evans EH, Cleland C, Moffatt S, et al. Integrating evidence from systematic reviews, qualitative research, and expert knowledge using co-design techniques to develop a Web-based intervention for people in the retirement transition. J Med Internet Res 2016 Aug 03;18(8):e210 [FREE Full text] [doi: 10.2196/jmir.5790] [Medline: 27489143]
49. Sripathi V, Sandru V. Effective usability testing–knowledge of user centered design is a key requirement. Int J Emerg Technol Adv Eng 2013 Feb;3(1):627-635.
50. Williams I, Brereton M, Donovan J, McDonald K, Millard T, Tam A, et al. A collaborative rapid persona-building workshop: creating design personas with health researchers. Int J Sociotechnol Knowl Dev 2014;6(2):17-35. [doi: 10.4018/ijskd.2014040102]
51. Schulz T, Fuglerud KS. Creating personas with disabilities. In: Proceedings of the 13th International Conference on the Computers Helping People with Special Needs. 2012 Presented at: ICCHP '12; July 11-13, 2012; Linz, Austria p. 145-152. [doi: 10.1007/978-3-642-31534-3_22]
52. Fuglerud KS, Schulz T, Janson AL, Moen A. Co-creating persona scenarios with diverse users enriching inclusive design. In: Proceedings of the 14th International Conference on Universal Access in Human-Computer Interaction. Design Approaches and Supporting Technologies. 2020 Presented at: UAHCI '20; July 19–24, 2020; Copenhagen, Denmark p. 48-59. [doi: 10.1007/978-3-030-49282-3_4]
53. Pyae A, Yuen TB, Gossage M. Persona development for designing human-centered rehabilitation games for stroke patients. In: Proceedings of the 2nd Annual Global Healthcare Conference. 2013 Presented at: GHC '13; November 17-22, 2013; Lisbon, Portugal p. 154-159. [doi: 10.5176/2251-3833_ghc13.69]
54. Reeder B, Hills RA, Turner AM, Demiris G. Participatory design of an integrated information system design to support public health nurses and nurse managers. Public Health Nurs 2014;31(2):183-192 [FREE Full text] [doi: 10.1111/phn.12081] [Medline: 24117760]
55. Marent B, Henwood F, Darking M, EmERGE Consortium. Development of an mHealth platform for HIV care: gathering user perspectives through co-design workshops and interviews. JMIR Mhealth Uhealth 2018 Oct 19;6(10):e184 [FREE Full text] [doi: 10.2196/mhealth.9856] [Medline: 30339132]
56. Hanson-Smith V, Wimalasuriya D, Fortier A. NutriStat: tracking young child nutrition. In: Extended Abstracts on Human Factors in Computing Systems. 2006 Presented at: CHI EA '06: CHI '06; April 22-27, 2006; Montreal, Canada p. 1831-1836. [doi: 10.1145/1125451.1125798]
57. Watson KS, Henderson V, Murray M, Murphy AB, Levi JB, McDowell T, et al. Engaging African American men as citizen scientists to validate a prostate cancer biomarker: work-in-progress. Prog Community Health Partnersh 2019;13(5):103-112 [FREE Full text] [doi: 10.1353/cpr.2019.0043] [Medline: 31378740]
58. Bradway M, Morris RL, Giordanengo A, Årsand E. How mHealth can facilitate collaboration in diabetes care: qualitative analysis of co-design workshops. BMC Health Serv Res 2020 Nov 30;20(1):1104 [FREE Full text] [doi: 10.1186/s12913-020-05955-3] [Medline: 33256732]
59. International Classification of Functioning, Disability and Health (ICF). World Health Organization. Geneva, Switzerland: World Health Organization; 2001. URL: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health [accessed 2022-06-01]
60. Free Word Cloud Generator. URL: https://www.freewordcloudgenerator.com [accessed 2022-07-20]
61. Jessen S, Mirkovic J, Nes LS. MyStrengths, a strengths-focused mobile health tool: participatory design and development. JMIR Form Res 2020 Jul 24;4(7):e18049 [FREE Full text] [doi: 10.2196/18049] [Medline: 32706651]
62. Tasoudis S, Perry M. Participatory prototyping to inform the development of a remote UX design system in the automotive domain. Multimodal Technol Interact 2018 Oct 24;2(4):74. [doi: 10.3390/mti2040074]

63. Näkki P, Antikainen M. Online tools for co-design: user involvement through the innovation process. In: Proceedings of the NordiCHI 2008 Workshops. 2008 Presented at: NordiCHI '08; October 18-22, 2008; Trondheim, Norway p. 92-97.
64. Singh S. Disability ethics in the coronavirus crisis. J Family Med Prim Care 2020 May;9(5):2167-2171 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_588_20] [Medline: 32754466]
65. Nesher Shoshan H, Wehrt W. Understanding "Zoom fatigue": a mixed‐method approach. Appl Psychol 2021 Nov 14;71(3):827-852. [doi: 10.1111/apps.12360]
66. Lin P, Van Brummelen J. Engaging teachers to co-design integrated AI curriculum for K-12 classrooms. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan p. 1-12. [doi: 10.1145/3411764.3445377]
67. Kennedy A, Cosgrave C, Macdonald J, Gunn K, Dietrich T, Brumby S. Translating co-design from face-to-face to online: an Australian primary producer project conducted during COVID-19. Int J Environ Res Public Health 2021 Apr 14;18(8):4147 [FREE Full text] [doi: 10.3390/ijerph18084147] [Medline: 33919920]
68. Ramos G, Ponting C, Labao JP, Sobowale K. Considerations of diversity, equity, and inclusion in mental health apps: a scoping review of evaluation frameworks. Behav Res Ther 2021 Dec;147:103990 [FREE Full text] [doi: 10.1016/j.brat.2021.103990] [Medline: 34715396]
69. Etches A, Phetteplace E. Know thy users: user research techniques to build empathy and improve decision-making. Ref User Serv Q 2013;53(1):13-17. [doi: 10.5860/rusq.53n1.13]
70. Hoople G, Choi-Fitzpatrick A, Reddy E. Educating changemakers: cross disciplinary collaboration between a school of engineering and a school of peace. In: Proceedings of the 2018 IEEE Frontiers in Education Conference. 2018 Presented at: FIE '18; October 3-6, 2018; San Jose, CA, USA p. 1-5. [doi: 10.1109/fie.2018.8658611]
71. Sustar H, Bowen S, Dearden A, Fisher M, Wolstenholme D. Using popular culture to enable health service co-design with young people. In: European Academy of Design Conference. 2013 Presented at: EAD '13; April 17-19, 2013; Gothenburg, Sweden p. 1-17.
72. Webber R, Partridge R, Grindell C. The creative co-design of low back pain education resources. Evid Policy 2022 Feb 24;18(2):436-453. [doi: 10.1332/174426421x16437342906266]

Abbreviations

C2P: Connect to Peers
ICF: International Classification of Functioning, Disability, and Health
mHealth: mobile health
RQ: research question

Edited by A Mavragani; submitted 03.03.22; peer-reviewed by J Austin, A Lucas; comments to author 04.05.22; revised version received 02.06.22; accepted 23.06.22; published 26.07.22.

Please cite as:
Adler RF, Morales P, Sotelo J, Magasi S
Developing an mHealth App for Empowering Cancer Survivors With Disabilities: Co-design Study
JMIR Form Res 2022;6(7):e37706
URL: https://formative.jmir.org/2022/7/e37706
doi:10.2196/37706
PMID:35881439

Original Paper

Developing the Message Assessment Scale for Tobacco Prevention Campaigns: Cross-sectional Validation Study

Jessica M Rath[1,2,3], MPH, PhD; Siobhan N Perks[1], MPH; Kenneshia N Williams[1], MS, PhD; Tracy Budnik[4], MBA; John Geraci[5], MBA; Donna M Vallone[1,2,6], MPH, PhD; Elizabeth C Hair[1,2,6], PhD

[1]Truth Initiative Schroeder Institute, Washington, DC, United States

[2]Department of Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[3]Department of Behavioral and Community of Health, University of Maryland School of Public Health, College Park, MD, United States

[4]Bridge Analytics, Limited Liability Company, Mendham, NJ, United States

[5]Crux Research, Honeoye Falls, NY, United States

[6]School of Global Public Health, New York University, New York, NY, United States

Corresponding Author:
Jessica M Rath, MPH, PhD
Truth Initiative Schroeder Institute
900 G St. NW
Fourth Floor
Washington, DC, 20001
United States
Phone: 1 2404545555
Email: jrath@truthinitiative.org

Abstract

Background: Mass media campaigns are effective for influencing a broad range of health behaviors. Prior to launching a campaign, developers often conduct ad testing to help identify the strengths and weaknesses of the message executions among the campaign's target audience. This process allows for changes to be made to ads, making them more relevant to or better received by the target audience before they are finalized. To assess the effectiveness of an ad's message and execution, campaign ads are often rated using a single item or multiple items on a scale, and scores are calculated. Endorsement of a 6-item perceived message effectiveness (PME) scale, defined as the practice of using a target audience's evaluative ratings to inform message selection, is one approach commonly used to select messages for antitobacco campaigns; however, the 6-item PME scale often does not produce enough specificity to make important decisions on ad optimization. In addition, the PME scale is typically used with adult populations for smoking cessation messages.

Objective: This study includes the development of the Message Assessment Scale, a new tobacco prevention message testing scale for youth and young adults.

Methods: Data were derived from numerous cross-sectional surveys designed to test the relevance and potential efficacy of antitobacco truth campaign ads. Participants aged 15-24 years (N=6108) responded to a set of 12 core attitudinal items, including relevance (both personal and cultural) as well as comprehension of the ad's main message.

Results: Analyses were completed in two phases. In phase I, mean scores were calculated for each of the 12 attitudinal items by ad type, with higher scores indicating more endorsement of the item. Next, all items were submitted to exploratory factor analysis. A four-factor model fit was revealed and verified with confirmatory factor analysis, resulting in the following constructs: personally relevant, culturally relevant, the strength of messaging, and negative attributes. In phase II, ads were categorized by performance (high/medium/low), and constructs identified in phase I were correlated with key campaign outcomes (ie, main fact agreement and likelihood to vape). Phase II confirmed that the four constructs identified in phase I were all significantly correlated with main fact agreement and vape intentions.

Conclusions: Findings from this study advance the field by establishing an expanded set of validated items to comprehensively assess the potential effectiveness of advertising executions. This set of items expands the portfolio of ad testing measures for ads focused on tobacco use prevention. Findings can inform how best to optimize ad executions and message delivery for health behavior campaigns, particularly those focused on tobacco use prevention among youth and young adult populations.

(JMIR Form Res 2022;6(7):e38156) doi:10.2196/38156

KEYWORDS

communication; youth/young adults; scales; message; behavior; health; campaign; tobacco; smoking cessation; prevention; youth; young adults; data; data analysis

Introduction

Mass media public education campaigns are an effective population-level intervention for influencing a broad range of health behaviors. Prior to launching a campaign, developers often conduct message testing to assess creative executions that will be effective among the specific audience targeted by the campaign. For a campaign to be effective, messages must resonate with the target audience to influence changes in knowledge, attitudes, and intentions, and in turn, the specific health behavior of interest. According to the Centers for Disease Control and Prevention, when campaign messages are delivered to attain sufficient reach, frequency, and duration, one can expect changes in campaign-targeted attitudes after the campaign has aired for 12 to 18 months and can expect behavioral changes after the campaign has aired for 18 to 24 months [1]. To assess message effectiveness, campaign messages are often rated using a single item or multiple items on a scale, and scores are calculated [2].

Perceived message effectiveness (PME), defined as the practice of using a target audience's evaluative ratings to inform message selection, is one approach to assess message effectiveness. Measuring PME using six items (ie, the ad is powerful, the ad is meaningful, the ad captures my attention, the ad is informative, the ad is convincing, and the ad is worth remembering) on a 5-point agreement scale has been shown to be a valid predictor of ad effectiveness [3]. The assumption is that messages scoring higher on the PME scale would be more likely to affect actual message effectiveness, for example, changing knowledge, attitudes, intentions, and ultimately the behavior of interest [4].

Several studies demonstrate the predictive validity of PME for changes in attitudes, intentions, and behaviors related to substance use among young people, particularly related to antitobacco campaigns [3,5-8]. For example, a recent systematic review of longitudinal studies in the antitobacco campaign literature examined the use of PME as a validated indicator of message effectiveness; researchers found that across 6 studies, PME provided predictive validity in measuring the effectiveness of antitobacco-related messages. In particular, the review confirmed that PME was associated with a variety of beliefs (eg, beliefs about smoking and quitting smoking) and behaviors (eg, message recall, conversations about ads, quit intentions, and cessation behavior) [8]. Although PME has been widely used for message selection, the scale often does not produce enough specificity to effectively modify or optimize an ad execution to improve its efficacy. Ad optimization refers to the process of using data to guide modifications/changes to ads that make them more relevant to or better received by the target audience. In addition to PME not providing enough specificity, PME has also primarily been used to assess smoking cessation media messages designed for adult populations.

The purpose of this study was to develop and validate additional measures to assess message effectiveness while providing additional evidence for ad optimization for tobacco prevention campaigns designed for youth and young adults. Data related to ad optimization, prior to campaign launch, can increase the likelihood of maximizing campaign resources. Findings from this study can broaden the capacity of evaluators to conduct message testing for antitobacco campaigns, particularly with respect to younger populations.

Methods

Overview

This study was conducted in two phases. Phase I used data from 11 cross-sectional surveys, conducted using the Dynata online panel from December 2016 to September 2018. Surveys were identical (with the exception of minor customizations, specific to the ads) and were conducted to assess receptivity to truth campaign ads that messaged on tobacco use but varied in their strategic goals. Participants were randomized to view only one ad, to avoid positional bias, before completing the survey. Each survey included 12 items related to message receptivity, including perceptions of effectiveness and relevance. The final sample included 2577 participants, aged 15-24 years, across 5 antitobacco ads (n=1275) and 6 antivape ads (n=1302). During this phase, mean scores were calculated for each of the 12 items by ad type, and factor analyses were run. Factor analyses were run at this phase to identify how the 12 items fit together and mapped onto different constructs.

The phase II analysis was conducted to validate constructs identified at phase I. Together, these constructs would go on to be referred to as the Message Assessment Scale. Phase II included the constructs identified at phase I in 13 cross-sectional forced exposure surveys, conducted using the Dynata online panel from July 2014 to June 2019 to test truth campaign ads. The final sample included 3531 participants, aged 15-24 years, across 10 antitobacco ads (n=2633) and 3 antivape ads (n=898). Mean scores for the constructs were examined across ads to assess campaign-aligned performance, and correlations were run to examine the relationship between construct scores for antivape ads, main fact agreement, and vape intentions. The goal of the correlation analysis at this phase was to determine if the constructs identified at phase I impacted campaign outcomes in a campaign-aligned direction.

Ethical Considerations

All study protocols were reviewed and approved for human participation in research by the institutional review board of Advarra Inc (Pro00034056; formally Chesapeake IRB). Informed consent was included in the online survey and participants could check "Yes, I would like to continue" or "No, I do not wish to take part in this study." Parental permission was collected from minor participants younger than 18 years.

Measures

Attitudinal Items

Ad *likability* was assessed using the item "In general, what is your impression of this ad?" Response options were on a 5-point Likert scale including "dislike it a lot," "dislike it somewhat," "neither like nor dislike it," "like it somewhat," and "like it a lot." Average likability was reported, ranging from 1 to 5, with higher scores indicating a more likable ad.

All other attitudinal items were assessed on a 5-point Likert scale including "strongly disagree," "disagree," "neither agree nor disagree," "agree," and "strongly agree." The following base was used: "How strongly do you agree or disagree with each of the following statements about the ad?" Items included "It told me something I didn't already know," "It gave me good reasons not to [smoke; use e-cigarettes/vape]," "It really speaks to me," "I identify with what this message says," "It is for people like me," "It is relevant for my generation," "It feels modern/current," "It is an acceptable way to talk about the issue of [smoking; using e-cigarettes/vaping]," "It is motivating," "It is believable," "It makes me want to tell someone about it," "It is confusing," "It is too fast," and "It is offensive."

Main Fact Agreement

Each of the 13 cross-sectional surveys used in phase II included main facts for the ad being tested. Participants were asked "How strongly do you agree or disagree that the video communicates each of the following messages?" Examples of facts included were "Just because vaping is safer than cigarettes, doesn't make it safe" and "People who vape are being tested on." Response options were on a 5-point Likert scale ranging from "strongly disagree that the ad conveys this message" to "strongly agree that the ad conveys this message." Higher scores indicated greater fact agreement.

Vape Intentions

Self-reported change in likelihood to use or try vapes or e-cigarettes after seeing each ad was assessed by the item "Since having seen this ad are you...?" Answer options included "much more likely to use or try vapes or e-cigarettes including JUUL," "somewhat more likely to use or try vapes or e-cigarettes including JUUL," "no change in my likelihood to use or try vapes or e-cigarettes including JUUL," "somewhat less likely to use or try vapes or e-cigarettes including JUUL," and "much less likely to use or try vapes or e-cigarettes including JUUL." Items were categorized for analysis into "more likely," "no change," and "less likely" to use or try vapes or e-cigarettes including JUUL.

Demographics

Demographic items included age (grouped as 15-17 years, 18-21 years, and 22-24 years), gender (male and female), race/ethnicity (non-Hispanic White; Hispanic, Latino, or Spanish origin; non-Hispanic Black or African American; and non-Hispanic other/declined), and subjective financial situation (do not meet basic expenses, just meet basic expenses with nothing left over, meet needs with a little left over, and live comfortably).

Analysis Plan

Descriptive statistics were calculated to identify the phase I and phase II samples. In the phase I analysis, mean scores were calculated for each of the 12 attitudinal items by ad type, with higher scores indicating more endorsement of that item. The 12 attitudinal items were then entered into an exploratory factor analysis to determine possible factor structures. Factor loadings and subsequent interpretation of the relative component scores indicated that a four-factor solution best fit the data. Using the supported four-factor structure, a confirmatory factor analysis (CFA), using varimax rotation, was conducted to determine the validity of the scale, check model fit, and examine internal consistency for each identified factor using Cronbach alpha. Final factors and corresponding psychometric properties are presented to represent phase I results.

In phase II, an external marketing consultant, with a decade's long experience analyzing truth pre- and postmarket data, placed the 13 ads into high (n=6 ads), medium (n=4 ads), or low (n=3 ads) performance groups based on the original goals of each advertisement such as agreement with key targeted beliefs like "ending youth smoking is an achievable goal" and relative performance of each against strategic objectives. Mean scores were examined across all ads in each ad performance group to determine how these four constructs varied by performance. Mean scores were also used to determine if items performed in a campaign-aligned direction such that more agreement on items indicated the ads were more relevant and message comprehension was higher. Finally, correlations were examined between construct scores for the antivape ads and main fact agreement as well as between construct scores for the antivape ads and self-reported change in likelihood (ie, a decrease in intentions) to vape since having seen the ad. Data were analyzed using SPSS (IBM Corp) and MPlus (Muthén and Muthén) statistical modeling programs.

Results

Sample Description

Participant characteristics are summarized in Table 1. At phase I, the mean age was 19.4 (SD 2.9) years, while at phase II, the mean age was 18.8 (SD 2.5) years. Of the 2577 participants in phase I, 1315 (51%) were male and 1262 (49.9%) were female. The majority (n=1378, 53.5%) were non-Hispanic White. Of the 3531 participants in phase II, 1793 (50.8%) were male and 1738 (49.2%) were female. Non-Hispanic White respondents also represented the largest racial category (n=1829, 51.8%). Finally, at phase I, most (n=1649, 64%) respondents reported meeting their financial needs either "comfortably" or with "a little left over," and at phase II, most (n=2317, 65.6%) respondents also reported meeting their financial needs either "comfortably" or with "a little left over."

Table 1. Sample characteristics for the phase I and phase II samples.

Characteristic	Phase I (n=2577), n (%)	Phase II (n=3531), n (%)
Age (years)		
15-17	834 (32.4)	1193 (33.8)
18-21	1030 (40.0)	1833 (51.9)
22-24	713 (27.7)	505 (14.3)
Gender		
Male	1315 (51.0)	1793 (50.8)
Female	1262 (49.0)	1738 (49.2)
Race/ethnicity		
Non-Hispanic White	1378 (53.5)	1829 (51.8)
Hispanic, Latino, or Spanish origin	518 (15.3)	504 (14.3)
Non-Hispanic Black or African American	393 (20.1)	746 (21.1)
Non-Hispanic, Other/Declined	288 (11.2)	452 (12.8)
Subjective financial situation		
Do not meet basic expenses	191 (7.4)	248 (7.0)
Just meet basic expenses with nothing left over	737 (28.6)	966 (27.4)
Meet needs with a little left over	1071 (41.6)	1418 (40.2)
Live comfortably	578 (22.4)	899 (25.5)

Phase I

CFA results are summarized in Table 2. The factor analysis revealed four new constructs (personally relevant, culturally relevant, strength of message, and negative attributes) with a total of 12 items. Fit statistics indicated that the four-factor model fit the data well (comparative fit index 0.97, root mean square error of approximation 0.05, standardized root mean squared residual 0.03; χ^2_{66}=9182.5; P<.001) [9].

Mean scores for each of the 12 items tested in phase I are listed in Table 3 by ad type (antitobacco and antivape). Mean score analyses indicated that the items performed similarly across antitobacco and antivape ads. The personal relevance, cultural relevance, and strength of message items had higher overall mean scores (scores closer to 5 indicated more agreement that the ads were personally relevant, were culturally relevant, and had strong messaging). The negative attribute items, however, had lower overall mean scores, indicating less confusion, offensiveness, or pacing concerns.

Table 2. Confirmatory factor analysis results for the phase I sample.

	Estimate	SE	Est/SE
Personally relevant			
It really speaks to me	0.83	0.01	83.50
I identify with what this message says	0.73	0.01	55.50
It is for people like me	0.72	0.02	49.24
Culturally relevant			
It is relevant for my generation	0.75	0.02	48.31
It feels modern/current	0.72	0.02	45.76
It is an acceptable way to talk about the issue	0.75	0.02	50.76
Strength of message			
It is motivating	0.81	0.01	74.20
It is believable	0.73	0.01	51.19
It makes me want to tell someone about it	0.74	0.01	56.79
Negative attributes			
It is confusing	0.81	0.02	39.06
It is offensive	0.57	0.03	22.97
The pace was too fast	0.56	0.02	25.23

Table 3. Mean scores for the items tested in phase I, by ad type.

	Antitobacco ads (n=1275), mean (SD)	Antivape ads (n=1302), mean (SD)
Personally relevant (range 1-5)		
It really speaks to me	3.36 (1.12)	2.99 (1.17)
I identify with what this message says	3.41 (1.12)	3.28 (1.09)
It is for people like me	3.64 (1.08)	3.17 (1.16)
Culturally relevant (range 1-5)		
It is relevant for my generation	3.97 (0.91)	3.78 (1.08)
It feels modern/current	3.92 (0.90)	3.55 (1.07)
It is an acceptable way to talk about the issue	3.93 (0.94)	3.65 (1.06)
Strength of message (range 1-5)		
It is motivating	3.69 (0.99)	3.25 (1.11)
It is believable	4.02 (0.88)	3.69 (1.01)
It makes me want to tell someone about it	3.49 (1.12)	3.21 (1.13)
Negative attributes (range 1-5)		
It is confusing	2.02 (1.05)	2.23 (1.07)
It is offensive	2.04 (1.09)	2.01 (0.99)
It is too fast	2.34 (1.02)	2.54 (1.06)

Phase II

Internal consistency for the constructs identified at phase I was assessed using Cronbach alpha. The four new constructs had alpha scores above .69, indicating an acceptable level of internal consistency (Table 4). Mean scores for the personal relevance, cultural relevance, and strength of message constructs were the highest for the high-performance ads. Accordingly, the mean scores for the negative attributes constructs were the lowest for the high-performance ads (Table 4).

Correlations with main fact agreement and vape intentions for the antivape ads are summarized in Tables 5 and 6. Overall, results revealed that the personal relevance, cultural relevance, and strength of message constructs were significantly positively correlated with main fact agreement and vape intentions for most of the antivape ads. In other words, as personal relevance,

cultural relevance, and strength of message scores increased, participants were more likely to agree with the ads' main fact and report a decrease in intentions to vape. Results also revealed that higher scores on the negative attributes construct showed an overall significant inverse relationship with main fact agreement and vape intentions.

Table 4. Mean scores for the four constructs tested in phase II, by ad performance level.

	Alpha	High performance (n=6 ads), mean (SD)	Medium performance (n=4 ads), mean (SD)	Low performance (n=3 ads), mean (SD)
Personally relevant[a]	.77	3.57 (0.21)[b]	3.45 (0.07)	3.08 (0.19)
Culturally relevant[a]	.79	3.80 (0.15)[c]	3.89 (N/A[d])[e]	3.56 (0.25)
Strength of message[a]	.69	3.69 (0.23)	3.49 (0.08)	3.34 (0.21)
Negative attributes[a]	.75	2.18 (0.10)	2.24 (0.13)	2.55 (0.19)

[a]Range of possible scores for constructs was 1-5.

[b]n=5 ads in the "high performance" group for this item.

[c]n=2 ads in the "high performance" group for this item.

[d]N/A: not applicable.

[e]n=1 ad in the "medium performance" group for this item.

Table 5. Correlations with main fact agreement for the phase II antivape ads.

Ad tested	Antivape ad 1 (n=298), *r* (*P* value)	Antivape ad 2 (n=299), *r* (*P* value)	Antivape ad 3 (n=301), *r* (*P* value)
Personally relevant	0.20 (.001)	0.32 (<.001)	0.24 (<.001)
Culturally relevant	0.23 (<.001)	0.43 (<.001)	0.35 (<.001)
Strength of message	0.21 (<.001)	0.45 (<.001)	0.28 (<.001)
Negative attributes	–0.19 (.001)	–0.22 (<.001)	–0.23 (<.001)

Table 6. Correlations with vape intentions for the phase II antivape ads.

Ad tested	Antivape ad 1 (n=298), *r* (*P* value)	Antivape ad 2 (n=299), *r* (*P* value)	Antivape ad 3 (n=301), *r* (*P* value)
Personally relevant	0.21 (.004)	0.20 (.008)	0.13 (.11)
Culturally relevant	0.28 (<.001)	0.30 (<.001)	0.19 (.02)
Strength of message	0.23 (.002)	0.32 (<.001)	0.12 (.14)
Negative attributes	–0.02 (.78)	–0.32 (<.001)	–0.13 (.10)

Discussion

Principal Findings

Study findings demonstrate the utility of an expanded set of ad testing items to aid in message selection and optimization (four constructs). The set of validated constructs, when used together, are referred to as the Message Assessment Scale. The constructs can be used to assess each individual item or by calculating construct scores. These constructs provide useful specific data to inform how best to increase an ad's relevance and effectiveness, specific to a youth and young adult audience. For example, an ad may perform low on cultural relevance, in which case qualitative responses to the overall advertisement "likes" and "dislikes" would be coded. The initial low score on this scale would tip researchers off to the need to dive deeper and look for comments in the qualitative questions, which could inform the low score and provide insight on how to improve it. The ads in this study that best met their goals were more personally and culturally relevant, and had stronger messaging than the lower performing ads. This is especially important because the validated constructs were significantly correlated with intentions not to vape for antivape ads. The better an ad performs on these constructs, the more potential it has to decrease vape intentions among a youth and young adult audience. Ad optimization, based on construct results, demonstrates the utility of maximizing message effectiveness for health behavior campaigns.

In Table 3, we see higher mean scores on the constructs for antitobacco (combustible products) versus antivape ads. These findings were not surprising given how long antitobacco (cigarette) ads have been on air, rates of smoking at that time, and the consensus around their impact on health. Constructs were developed from antivape ads airing early in the vaping epidemic, a period of time in which we were trying to stop a behavior that youth were really enjoying. An antivape message at that time felt less relevant to our audience because there was no data to indicate health or other negative implications from vaping yet.

The Message Assessment Scale can help establish benchmarks for modifying aspects of the execution, the placement of the

execution, and the frequency of airing the execution. Ad testing, thus, informs placement and frequency decisions. For example, ads scoring higher on the measures may receive more or less frequent rotation on television or digital platforms, an ad on television or digital platforms may be coupled with another ad that increases scores, or a digital ad may be selected to be elevated to a television spot. Additionally, qualitative items related to ad receptivity and disapproval can be coupled with data from quantitative items to comprehensively assess likeability, level of novel information, and whether the execution provides motivation. For example, asking "what did you like most about the ad" and "what did you dislike about the ad" with open-ended responses can provide meaningful data that may not have been received from quantitative items with predetermined answer options. Moreover, the measures also provide insight into the possible reasons for low testing scores, including issues related to pacing and ad characteristics. If an ad is seen as offensive or confusing, this may interfere with the ad's ability to effectively deliver a message to the target audience.

Limitations

Although this study has many strengths, there are some limitations. First, the study used panel data, which does not reflect a probabilistic sample. As such, responses may not be generalizable to a broader population. However, sample quotas were set to yield approximately equal proportions by gender and age group. Additionally, this work did not examine correlations to in-market performance, which would ultimately test the ability of the constructs to predict how an ad may perform in the real-world.

Conclusions

Findings advance the field by establishing an expanded set of validated items to comprehensively assess the potential effectiveness of advertising executions. The constructs can provide critical information for message optimization and message selection, particularly among a youth and young adult audience. Ensuring ads meet testing benchmarks before airing them across media platforms helps ensure cost efficiency while providing critical empirical evidence that messages will effectively shift knowledge, attitudes, and beliefs among the target audience. Future studies should explore whether the constructs perform similarly across demographic subgroups and correlate to actual campaign performance.

Conflicts of Interest

None declared.

References

1. Schar E, Gutierrez K, Murphy-Hoefer R, Nelson DE. Tobacco use prevention media campaigns: lessons learned from youth in nine countries. Centers for Disease Control and Prevention. 2006. URL: https://stacks.cdc.gov/view/cdc/11400 [accessed 2022-06-29]
2. Baig SA, Noar SM, Gottfredson NC, Boynton MH, Ribisl KM, Brewer NT. UNC perceived message effectiveness: validation of a brief scale. Ann Behav Med 2019 Jul 17;53(8):732-742 [FREE Full text] [doi: 10.1093/abm/kay080] [Medline: 30321252]
3. Davis KC, Duke J, Shafer P, Patel D, Rodes R, Beistle D. Perceived effectiveness of antismoking ads and association with quit attempts among smokers: evidence from the tips from former smokers campaign. Health Commun 2017 Aug;32(8):931-938. [doi: 10.1080/10410236.2016.1196413] [Medline: 27435919]
4. Noar SM, Barker J, Yzer M. Measurement and design heterogeneity in perceived message effectiveness studies: a call for research. J Commun 2018 Oct;68(5):990-993 [FREE Full text] [doi: 10.1093/joc/jqy047] [Medline: 30479402]
5. Bigsby E, Cappella JN, Seitz HH. Efficiently and effectively evaluating public service announcements: additional evidence for the utility of perceived effectiveness. Commun Monogr 2013 Mar;80(1):1-23 [FREE Full text] [doi: 10.1080/03637751.2012.739706] [Medline: 25568588]
6. Brennan E, Durkin SJ, Wakefield MA, Kashima Y. Assessing the effectiveness of antismoking television advertisements: do audience ratings of perceived effectiveness predict changes in quitting intentions and smoking behaviours? Tob Control 2014 Sep;23(5):412-418. [doi: 10.1136/tobaccocontrol-2012-050949] [Medline: 23604496]
7. Davis KC, Nonnemaker J, Duke J, Farrelly MC. Perceived effectiveness of cessation advertisements: the importance of audience reactions and practical implications for media campaign planning. Health Commun 2013;28(5):461-472. [doi: 10.1080/10410236.2012.696535] [Medline: 22812702]
8. Noar SM, Barker J, Bell T, Yzer M. Does perceived message effectiveness predict the actual effectiveness of tobacco education messages? A systematic review and meta-analysis. Health Commun 2020 Feb;35(2):148-157 [FREE Full text] [doi: 10.1080/10410236.2018.1547675] [Medline: 30482058]
9. Browne MW, Cudeck R. Alternative ways of assessing model fit. Sociol Methods Res 2016 Jun 29;21(2):230-258. [doi: 10.1177/0049124192021002005]

Abbreviations

CFA: confirmatory factor analysis

PME: perceived message effectiveness

Edited by A Mavragani; submitted 22.03.22; peer-reviewed by T Turk, N Martin, S Beasley; comments to author 30.05.22; revised version received 17.06.22; accepted 21.06.22; published 26.07.22.

Please cite as:
Rath JM, Perks SN, Williams KN, Budnik T, Geraci J, Vallone DM, Hair EC
Developing the Message Assessment Scale for Tobacco Prevention Campaigns: Cross-sectional Validation Study
JMIR Form Res 2022;6(7):e38156
URL: https://formative.jmir.org/2022/7/e38156
doi:10.2196/38156
PMID:35881429

Original Paper

Implementing a Health Utility Assessment Platform to Acquire Health Utilities in a Hemodialysis Outpatient Setting: Feasibility Study

Adeboye A Adejare[1], PhD; Heather J Duncan[2], PhD; R Geoffrey Motz[2], MD; Silvi Shah[2], MD; Charuhas V Thakar[2], MD; Mark H Eckman[3], MD

[1]Department of Biomedical Informatics, University of Cincinnati, Cincinnati, OH, United States

[2]Division of Nephrology and Hypertension, University of Cincinnati, Cincinnati, OH, United States

[3]Division of General Internal Medicine and the Center for Clinical Effectiveness, University of Cincinnati, Cincinnati, OH, United States

Corresponding Author:
Adeboye A Adejare, PhD
Department of Biomedical Informatics
University of Cincinnati
CARE/Crawley Building
Suite E-870
Cincinnati, OH, 45220
United States
Phone: 1 8566252214
Email: adejaraa@mail.uc.edu

Abstract

Background: Patients with end-stage kidney disease (ESKD) wait roughly 4 years for a kidney transplant. A potential way to reduce wait times is using hepatitis C virus (HCV)–viremic kidneys.

Objective: As preparation for developing a shared decision-making tool to assist patients with ESKD with the decision to accept an HCV-viremic kidney transplant, our initial goal was to assess the feasibility of using The Gambler II, a health utility assessment tool, in an ambulatory dialysis clinic setting. Our secondary goals were to collect health utilities for patients with ESKD and to explore whether the use of race-matched versus race-mismatched exemplars impacted the knowledge gained during the assessment process.

Methods: We used The Gambler II to elicit utilities for the following ESKD-related health states: hemodialysis, kidney transplant with HCV-unexposed kidney, and transplantation with HCV-viremic kidney. We created race exemplar video clips describing these health states and randomly assigned patients into the race-matched or race-mismatched video arms. We obtained utilities for these 3 health states from each patient, and we evaluated knowledge about ESKD and HCV-associated health conditions with pre- and postintervention knowledge assessments.

Results: A total of 63 patients with hemodialysis from 4 outpatient Dialysis Center Inc sites completed the study. Mean adjusted standard gamble utilities for hemodialysis, transplant with HCV-unexposed kidney, and transplantation with HCV-viremic kidney were 82.5, 89, and 75.5, respectively. General group knowledge assessment scores improved by 10 points (P<.05) following utility assessment process. The use of race-matched exemplars had little effect on the results of the knowledge assessment of patients.

Conclusions: Using The Gambler II to collect utilities for patients with ESKD in an ambulatory dialysis clinic setting proved feasible. In addition, educational information about health states provided as part of the utility assessment process tool improved patients' knowledge and understanding about ESKD-related health states and implications of organ transplantation with HCV-viremic kidneys. A wide variation in patient health state utilities reinforces the importance of incorporating patients' preferences into decisions regarding use of HCV-viremic kidneys for transplantation.

(JMIR Form Res 2022;6(7):e33562) doi:10.2196/33562

KEYWORDS

health utility assessment; patient reported outcomes; end-stage kidney disease; hemodialysis; hepatitis C

Introduction

Chronic kidney disease disproportionally impacts African Americans, making it a prototypical disease in which to investigate disparities in health utility assessments (HUAs) [1]. Due to the limited availability of organs and the large number patients with end-stage kidney disease (ESKD) on waiting lists for transplantation, the average patient waits roughly 4 years before receiving a kidney transplant [2]. Patients with African American racial background who have ESKD wait even longer at an average of 4.5 years [2-4]. One path toward increasing the availability of organs and reducing waiting times is to use hepatitis C virus (HCV)–viremic kidneys for transplantation. The decision to accept transplantation with an HCV-viremic kidney hinges on the balance between the decreased waiting time afforded by accepting such an organ and each patient's values and preferences about receiving such an organ. These trade-offs make this decision an ideal setting for shared decision-making (SDM) [5]. SDM clarifies patients' values and preferences, can improve self-efficacy, and engages patients in conversations with their clinicians about treatment choices [5,6].

As of 2017, a total of 746,557 patients in the United States have ESKD [3]. A total of 101,337 patients with ESKD were wait-listed for kidney transplantation in 2019, while the number of patients receiving kidney transplants in 2019 was 24,273 [7]. Studies show that receiving a kidney transplant, even with an HCV-viremic organ, can improve the survival and quality of life and reduce lifetime costs for patients with ESKD who are currently on dialysis [8]. Importantly, using HCV-viremic organs can increase the availability of otherwise high-quality organs and potentially reduce the waiting time for some patients to receive kidney transplants [9]. However, some patients, even after undergoing successful transplantation with an HCV-viremic kidney, have worries and concerns that continue to impact their quality of life. Thus, including individual patient's values and preferences regarding the relevant health states is a critical component of SDM. In anticipation of the future development of an SDM tool, the primary goal of this study was to assess the feasibility of using The Gambler II, a health utility assessment tool, in an ambulatory dialysis clinic setting. Our secondary goals were to collect utilities for patients with ESKD and to explore whether the use of race-matched versus race-mismatched exemplars impacted the knowledge gained during the assessment process.

The Gambler II is a health utility assessment software platform that can use a variety of assessment techniques to gather patients' utilities [10,11]. These include the visual analog scale (VAS), standard gamble (SG), and time trade-off (TTO) [12] (Multimedia Appendix 1 A). The utility assessment process itself provides an opportunity to educate patients about the health states for which values and preferences are being sought [13,14]. The Gambler II uses video clips of patient actors to describe health states. In addition, The Gambler II can match the demographics of patient actors in the video clips to those of the patient whose utilities are being assessed. This latter feature provided us with an opportunity to explore a hypothesis, based on exemplification theory, that information communicated by people who mostly resemble the participant is more likely to engage and inform the participant [15,16].

Our longer-term goal is to develop an SDM tool that can be used by clinicians to facilitate discussions about acceptance of HCV-viremic kidneys, particularly for patients who may have longer predicted waiting times on transplant lists. The tool would assess patients' utilities for hemodialysis and transplantation with either an HCV-unexposed or an HCV-viremic kidney and would use that information along with patient-specific demographic information and predictions of organ availability based on factors including age, sex, blood type, dialysis vintage, calculated panel reactive antibodies, comorbidities, and the region in which the transplant is being carried out to make a recommendation for the best transplantation strategy for that patient. As many patients on dialysis, particularly African Americans living in less affluent urban areas, are impacted by the digital divide and have less access to home computers and the internet [6], we envision using this tool in ambulatory dialysis clinic settings with readily available computing platforms such as laptop and tablet computers.

Methods

Ethics Approval

This study was reviewed and approved by the University of Cincinnati Institutional Review Board (UC IRB ID – 2019-0792) as well as by the Dialysis Clinic Inc Administrative Review Office. Both boards approved the study before patient recruitment began. No compensation was provided to the patients for this research.

Study Design

Patient Recruitment

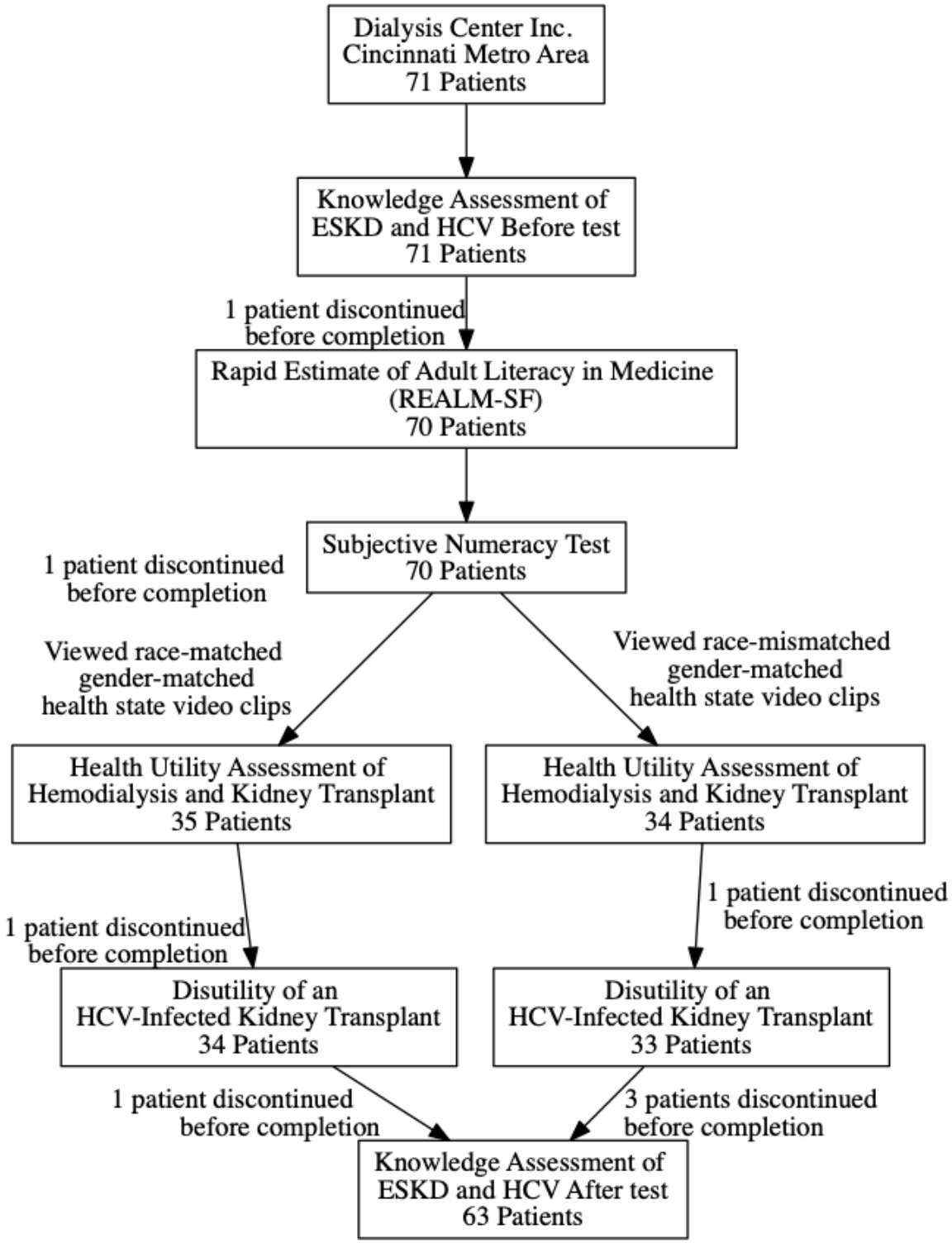
We worked with physicians in the Division of Nephrology and Hypertension to help recruit patients with chronic hemodialysis receiving treatment at any of the 4 outpatient dialysis centers in the greater Cincinnati metropolitan area managed by Dialysis Center Inc (DCI) [17]. Inclusion criteria were a diagnosis of ESKD, receiving intermittent facility hemodialysis at the designated outpatient center (DCI), and the ability to understand the English language. We included adults between the ages of 21 and 80 years. Patients with significant cognitive or reading deficits were excluded from the study. Patients who did not opt out of being contacted were given further explanation on the purpose of the study to assess their interest and willingness to participate.

Study Flow

Patients were recruited from 4 DCI sites in the Cincinnati metropolitan region. Prior to participating in the utility assessment process, they underwent a knowledge assessment survey. We collected demographic information and assessed health literacy using the Rapid Estimate of Adult Literacy in Medicine and subjective numeracy. We randomized patients to 1 of 2 study arms in which they viewed either race-matched or race-mismatched video clips describing each of the 3 health states. The patients' utilities for these health states were obtained, and then a postinterview knowledge assessment was

repeated (Figure 1). Those patients who indicated interest were met by the study principal investigator (AAA) in their local dialysis center, underwent a formal consent process, and answered on a laptop computer a minimal amount of demographic and clinical information, completing a short survey on educational status, time on dialysis, history of prior kidney transplant, and interest in a future kidney transplant. The patients then completed a Research Electronic Data Capture (REDCap) survey evaluating subjective numeracy and health literacy along with a short previsit knowledge assessment [18-20]. At any time, patients could withdraw from the study and did not have to provide a reason for withdrawal. To understand preintervention state of knowledge about hemodialysis, HCV infection, and kidney transplantation, we developed a 10-item multiple-choice questionnaire (Multimedia Appendix 1 B).

Figure 1. Study flow. ESKD: end-stage kidney disease; HCV: hepatitis C virus.



One of our secondary goals was to evaluate the impact of patients viewing race-matched versus race-mismatched video clips describing the health states being assessed. As mentioned, exemplification theory suggests that people are more engaged and receptive when they are presented with visual information by people who look most like themselves [16]. We hypothesized that learning about relevant health states would be greater among patients who receive video clip information from patient actors who are similar to them as opposed to receiving information from those who are different (eg, race, gender, and age category). We evaluated this by measuring the change score on the knowledge survey given before and after the HUA sessions. We did not provide any training materials prior to the HUA sessions and did not assume patients had knowledge about hepatitis C or any other knowledge besides their current understanding of dialysis and kidney transplant. We determined that 30 patients would be required in each study arm to detect a difference as small as 12 points (on a scale of 0 to 100) in the change score between pre- and postknowledge survey (power of 0.80 with an alpha of .05), using a 2-sided *t* test. We randomized patients into 1 of 2 study arms (Multimedia Appendix 1 C). The first arm presented patients with video clips of race- and gender-matched health state descriptions, while the second arm presented patients with video clips purposely mismatched for race (Figure 2 and Figure 3). The entire interview process lasted roughly 1 hour and took place while patients were receiving their 3- to 4-hour long dialysis treatment.

Another secondary goal of our study was to collect utilities for the 3 ESKD-related health states described above. While studies have assessed patient utilities for hemodialysis and kidney transplantation, to our knowledge, none have explored health utilities for transplantation with HCV-viremic kidneys [21,22]. We collected these data using the 3 HUA methods of visual analog scale ("feeling thermometer"), standard gamble, and time trade-off. We broke the assessment process into 2 parts. First, we assessed health states of intermittent hemodialysis and transplantation with an HCV-unexposed kidney, using anchor states of "Well" (without ESKD) and "Dead." We next assessed the health state of transplantation with an HCV-viremic kidney using transplantation with an HCV-unexposed kidney as the best outcome and Dead as anchor states. Patients entered one of these two HUA groups to explicitly assess how much of a risk they were willing to take (in the standard gamble) to avoid receiving an HCV-viremic kidney given the opportunity of a noninfected kidney transplant (Multimedia Appendix 1 D).

To control for possible confounding, we also collected demographic information (age, sex, and race), highest educational level attained, dialysis vintage, history of prior kidney transplant, interest in receiving a transplanted kidney, health literacy using the Rapid Estimate of Adult Literacy in Medicine Short Form, and subjective numeracy [20-22].

Figure 2. Screenshots of different health utility assessments. Top: patient evaluating the hemodialysis health state through the standard gamble utility assessment. Bottom: patients evaluating kidney transplant using the time trade-off; The Gambler uses life tables to determine the duration of life expectancy for the time trade-off. The video clip in the figure demonstrates how a user can watch a demographically matched patient actor describe the health state being assessed. We did not incorporate the image of the patient actor to protect their privacy. ESKD: end-stage kidney disease.

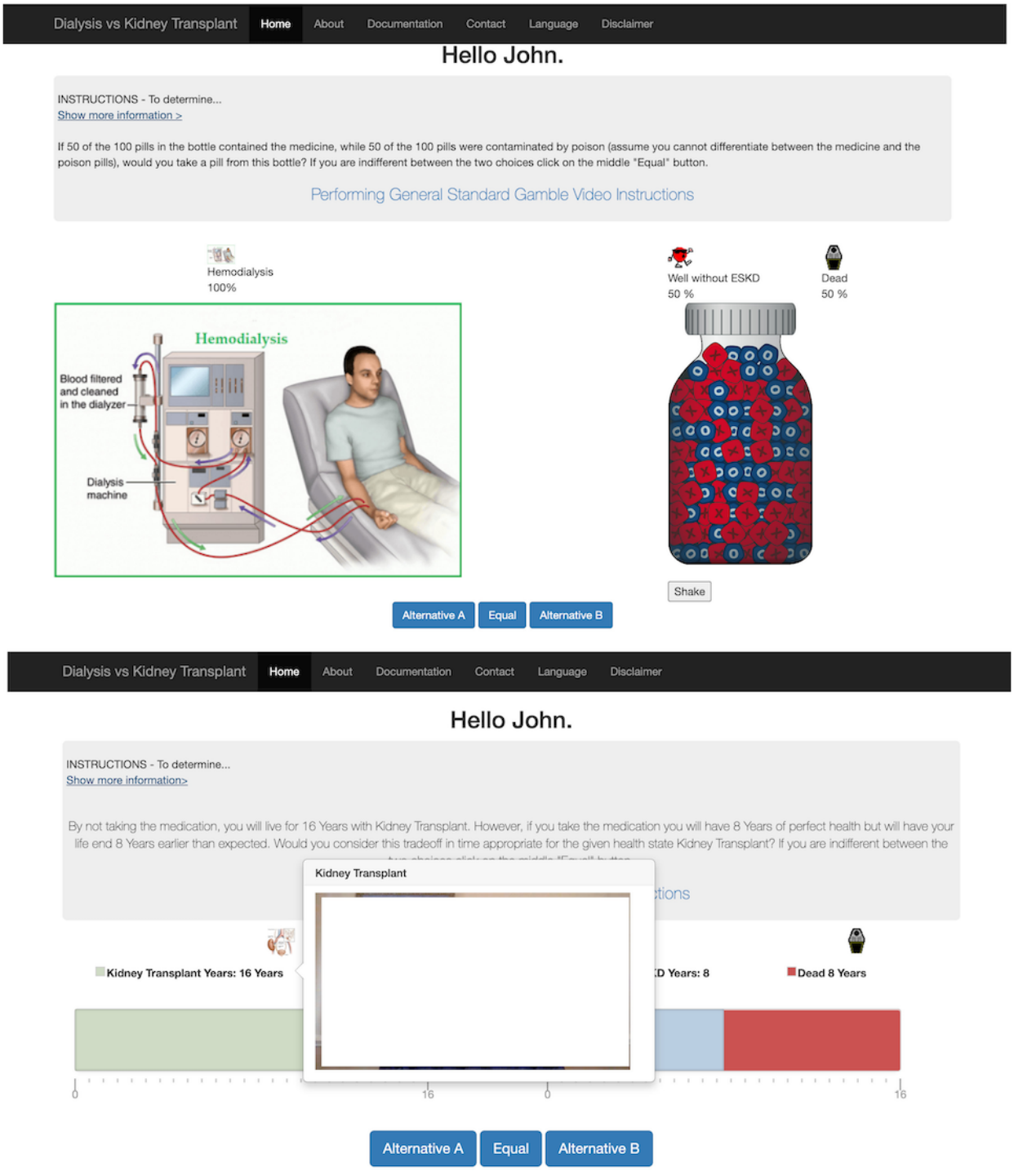
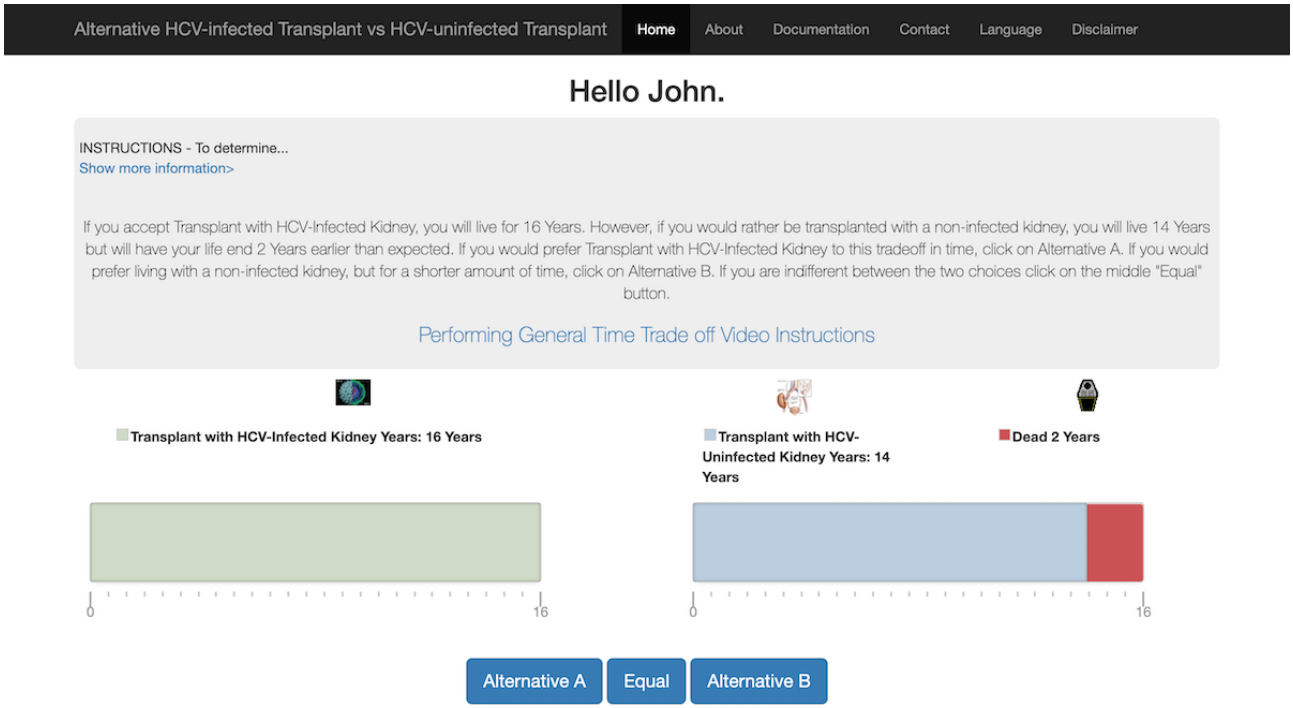


Figure 3. Screenshot of patient going through the time trade-off health utility assessment for transplantation with a hepatitis C virus (HCV)–infected kidney. This assessment differs from others in that the best anchor health state is a transplantation with an uninfected kidney.



Tools

We conducted surveys and performed HUAs using a laptop computer (2014 Apple MacBook Pro running Mac OS 10.14) while we met with most patients in their community dialysis center during their usual intermittent hemodialysis sessions. Patients used either headphones, earbuds, or audio output from the laptop to hear the video clips that were part of The Gambler II. We used Google Chrome version 80 as the web browser to access The Gambler II and REDCap (9.1.10). Patient actors used a standard script for each health state, which was developed and vetted by team members (MHE, HJD, SS, CVT, and GRM). We recorded 12 different video clips describing the 3 health states for all combinations of gender and race (ie, men, women, African American, and White) using Apple's QuickTime X 10.5 with video mastering done on Adobe Premiere and Adobe Audition 2019. Due to restrictions of social distancing during the COVID-19 pandemic, the institutional review board granted an amendment to the protocol that allowed us to interview some patients remotely [23]. For these patients, we used several teleconferencing applications including Microsoft Teams, Microsoft Skype, and Cisco's WebEx to administer surveys and conduct HUAs.

We used The Gambler II to perform health utility assessments [11]. For survey data collection, we used REDCap, a web-based survey platform designed to capture users' data with web forms [18]. All data were collected and stored on a secure server at the University of Cincinnati. Access to data was restricted to institutional review board–authorized users with login credentials. We analyzed data using Python (Python Software Foundation) and R (The R foundation) [24-26].

Results

Demographic Data

We recruited 71 patients from 4 DCI sites in the Cincinnati Metropolitan area. A total of 63 patients consented and completed the study. Of those who consented, 62 (98%) enrolled in the study during dialysis clinic visits, while 1 (2%) enrolled through teleconferencing software at their residence. The cohort's age range was from 24 to 80 years, with a median age of 59 years and mean age of 58 years (Multimedia Appendix 1). Moreover, 44 (70%) patients had African American racial backgrounds, and 19 (30%) were European American. A high school diploma was the most frequent highest level of educational attainment (Table 1). The median Rapid Estimate of Adult Literacy in Medicine Short Form and subjective numeracy scores were 7.0 and 4.0, respectively. In addition, of the 63 patients, 54 (86%) had not received a previous kidney transplant, and 47 (75%) had an interest in receiving a kidney transplant. The participants spent an average of 5.9 years on dialysis (African Americans: 7.3 years, European Americans: 2.79 years; *P*=.07; Multimedia Appendix 1).

Table 1. Summary statistics of the population broken into race-matched versus race-mismatched study arms.

Characteristics	Values			
	Overall (n=63)	Race-matched videos (n=33)	Race-mismatched videos (n=33)	*P* value
Age (years), mean (SD)	57.8 (12.3)	56.1 (13.5)	59.8 (10.8)	.26
Race or ethnicity, n (%)				.80
African American	44 (69.8)	24 (72.7)	20 (66.7)	
European American	19 (30.2)	9 (27.3)	10 (33.3)	
Gender, n (%)				.54
Female	30 (47.6)	14 (42.4)	16 (53.3)	
Male	33 (52.4)	19 (57.6)	14 (46.7)	
Highest education level attained, n (%)				.57
Less than high school diploma	10 (15.9)	4 (12.1)	6 (20.0)	
High school or general education diploma	25 (39.7)	16 (48.5)	9 (30.0)	
Some college, no degree	11 (17.5)	6 (18.2)	5 (16.7)	
Associate degree	5 (7.9)	3 (9.1)	2 (6.7)	
Bachelor’s degree	9 (14.3)	3 (9.1)	6 (20.0)	
Master's degree	3 (4.8)	1 (3.0)	2 (6.7)	
Years on dialysis, mean (SD)	5.9 (8.1)	6.79 (7.8)	5.0 (8.3)	.07
History of previous kidney transplant, n (%)	9 (14.3)	6 (18.2)	3 (10.0)	.48
Interested in receiving a kidney transplant, n (%)				.64
Maybe	3 (4.8)	1 (3.0)	2 (6.7)	
No	13 (20.6)	8 (24.2)	5 (16.7)	
Yes	47 (74.6)	24 (72.7)	23 (76.7)	

Knowledge Assessments

We conducted a knowledge assessment using a 10-item questionnaire administered before the utility assessment as a pretest and after as a posttest (Multimedia Appendix 1). The Cronbach alpha for the pretest was .990, and it was .994 for the posttest. As shown in Tables 2 and 3, for the cohort, the improvement in test scores (mean and median 10.0 points) following the utility assessment process and viewing of health state videos were clinically and statistically significant ($P<.001$, paired Wilcoxon *t* test). However, there was not a statistically significant difference in pretest versus posttest change scores between the 2 study arms, or between African Americans compared with European Americans, with *P* values of .95 and .96 (Mann-Whitney test), respectively.

Table 2. Evaluation of patients’ health literacy, numeracy, and their knowledge of end-stage kidney disease and hepatitis C stratified by study arm.

Tests	Overall, mean (SD)	Race-matched videos, mean (SD)	Raced-mismatched videos, mean (SD)	*P* value[a]
Before test	79.0 (17.3)	80.2 (17.2)	77.7 (17.7)	.59
After test	89.0 (15.9)	89.6 (15.8)	88.2 (16.1)	.77
Numeracy	3.9 (1.1)	4.1 (1.0)	3.7 (1.1)	.14
REALM-SF[b]	6.1 (1.6)	6.1 (1.5)	6.1 (1.8)	.98
Change score	10.0 (13.8)	9.5 (12.7)	10.5 (15.0)	.09

[a]*P* value denotes comparison between race-matched and race-mismatched population health utilities. This assumes that if $P \leq .05$, there was a significant difference in health utilities between race-matched (same race) and race-mismatched (different race) health utilities.

[b]REALM-SF: Rapid Estimate of Adult Literacy in Medicine Short Form.

Table 3. Evaluation of patients' health literacy, numeracy, and their knowledge of end-stage kidney disease and hepatitis C stratified by race.

Tests	Overall, mean (SD)	African American, mean (SD)	European American, mean (SD)	*P* value[a]
Before test	79.0 (17.3)	78.4 (16.0)	80.4 (20.5)	.42
After test	89.0 (15.9)	88.3 (15.6)	90.4 (16.7)	.45
Numeracy	3.9 (17.3)	3.8 (1.1)	4.2 (1.0)	.18
REALM-SF[b]	6.1 (1.1)	5.9 (1.9)	6.6 (0.77)	.26
Change score	10.0 (13.8)	9.9 (13.2)	10.1 (15.4)	.96

[a]*P* value denotes comparison between African and European American population health utilities. This assumes that if *P*≤.05, there was a significant difference in health utilities between European and African American health utilities.

[b]REALM-SF: Rapid Estimate of Adult Literacy in Medicine Short Form.

End-Stage Kidney Disease Health Utilities

We assessed utilities from patients on chronic intermittent hemodialysis for 3 health states relevant to decision-making about kidney transplantation. The results are reported in Tables 4 and 5 and are depicted pictorially in Figure 4. With all 3 assessment methods, means utilities were highest for transplantation with an HCV-unexposed kidney. Hemodialysis had the lowest average utility assessed with the VAS, while transplantation with an HCV-viremic kidney had the lowest utility using the TTO and SG. Looking at ranking of utilities within each patient, 47 (75%) and 39 (62%) of the patients rated transplantation with an HCV-viremic kidney lower than hemodialysis with SG utilities and TTO utilities, respectively.

Table 4. Health utilities evaluation of race-matched vs race-mismatched video patient cohort.

Health utilities	Overall, mean (SD)	Race-matched videos, mean (SD)	Race-mismatched videos, mean (SD)	*P* value[a]
Visual analog scale				
Hemodialysis	57.9 (25.9)	63.2 (26.7)	52.0 (25.3)	.11
Kidney transplant	88.2 (17.8)	85.2 (22.3)	91.5 (10.2)	.68
Hepatitis C–viremic kidney transplant[b]	66.30 (27.3)	65.6 (27.4)	67.1 (27.7)	.94
Standard gamble				
Hemodialysis	82.5 (23.1)	85.3 (20.5)	79.4 (25.6)	.26
Kidney transplant	89.0 (18.0)	87.4 (19.4)	90.7 (15.8)	.67
Hepatitis C–viremic kidney transplant[b]	75.5 (28.2)	75.5 (29.7)	75.5 (26.9)	.68
Time trade-off				
Hemodialysis	80.3 (20.5)	78.2 (21.0)	82.6 (20.0)	.39
Kidney transplant	84.8 (22.0)	85.4 (22.6)	84.2 (21.7)	.74
Hepatitis C-viremic kidney transplant[b]	73.8 (28.1)	75.1 (27.0)	72.3 (29.6)	.76

[a]*P* value denotes comparison between race-matched and race-mismatch population health utilities. This assumes that if *P*≤.05, there was a significant difference in health utilities between race-matched (same race) and race-mismatched (different race) health utilities.

[b]Health utility normalization equation: Raw hepatitis C virus utility * Kidney transplant utility = hepatitis C virus utility.

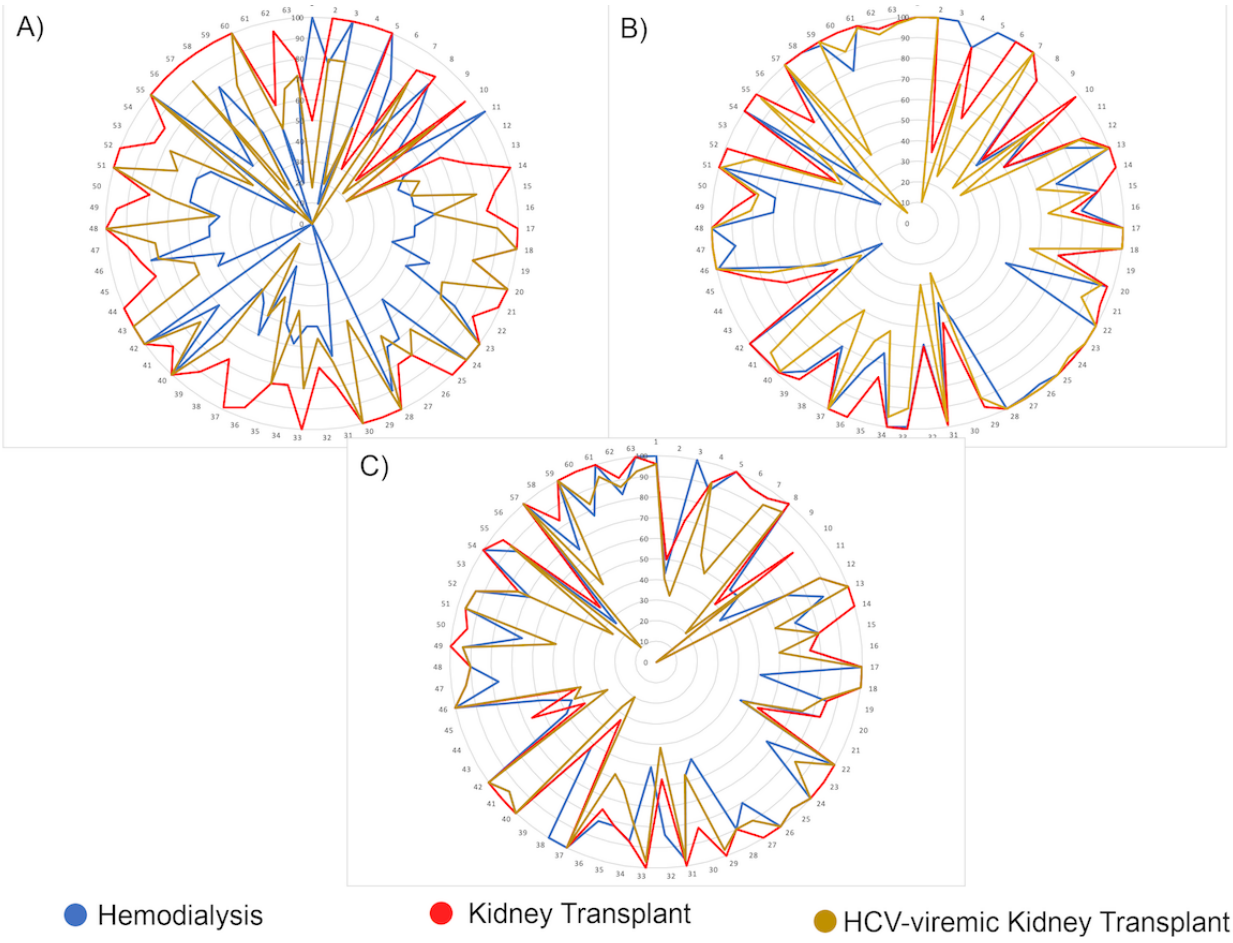
Table 5. Health utilities evaluation of the African American versus European American patient cohort.

Health utilities	Overall, mean (SD)	African American, mean (SD)	European American, mean (SD)	*P* value[a]
Visual analog scale, mean (SD)				
Hemodialysis	57.9 (25.9)	59.1 (27.4)	55.1 (22.4)	.38
Kidney transplant	88.2 (17.8)	86.7 (20.3)	91.8 (9.3)	.73
Hepatitis C–viremic kidney transplant[b]	66.30 (27.3)	65.6 (27.4)	68.0 (27.8)	.70
Standard gamble, mean (SD)				
Hemodialysis	82.5 (23.1)	83.0 (25.2)	81.3 (17.6)	.22
Kidney transplant	89.0 (18.0)	88.8 (19.6)	89.4 (14.3)	.52
Hepatitis C–viremic kidney transplant[b]	75.5 (28.2)	77.5 (28.5)	70.9 (27.5)	.24
Time trade-off				
Hemodialysis, mean (IQR; SD)	80.3 (66.0-100.0; 20.5)	83.2 (75.0-100.0; 21.0)	73.6 (66.0-83.5; 18.2)	.04
Kidney transplant, mean (SD)	84.8 (22.0)	85.2 (23.8)	84.0 (17.7	.16
Hepatitis C–viremic kidney transplant[b], mean (SD)	73.8 (28.1)	75.4 (29.7)	70.0 (21.0)	.20

[a]*P* value denotes comparison between race-matched and race-mismatch population health utilities. This assumes that if $P \leq .05$, there was a significant difference in health utilities between race-matched (same race) and race-mismatched (different race) health utilities.

[b]Health utility normalization equation: Raw hepatitis C virus utility * Kidney transplant utility = hepatitis C virus utility.

Figure 4. Radar plot of the health state utilities. (A) Visual analog scale; (B) standard gamble; and (C) time trade-off. The 3 figure panels show the utilities assessed for each of the 63 patients in the study. Each panel summarizes results for the 3 different utility assessment methods. Different colors are used to represent each of the 3 health states: hemodialysis, transplantation with a hepatitis C (HCV)–unexposed kidney, and transplantation with an HCV-viremic kidney. Each number on the outside circle represents a single patient's utility scores.



Utility assessments for transplantation with an HCV-infected kidney were carried out with anchor states of transplant with HCV-unexposed kidney as the best outcome and death as the worst outcome. We normalized these utilities to the same 0-100 scale used for the other health states by multiplying the raw utility for transplant with HCV-exposed kidney times the utility of transplant with an HCV-unexposed kidney. Normalized utility means using VAS, SG, and TTO assessments were 66.3, 75.5, and 73.8, respectively. Looking at the raw SG utility weights for transplantation with an HCV-viremic kidney, given a choice between transplantation with an HCV-unexposed kidney and an HCV-viremic kidney, patients were willing to take a 24.5% chance of dying to avoid receiving an HCV-viremic kidney. We did not find statistically significant differences in utilities between race-matched and race-mismatched study arms. However, we did find that African Americans on average had higher utility weights than European Americans for hemodialysis evaluated with the TTO (83.2 versus 73.6, respectively; *P*=.04 [Mann-Whitney]).

Discussion

Principal Results

Regarding our primary study goal, we found that using our health utility assessment tool in an ambulatory community dialysis clinic setting was feasible. The entire process took approximately 1 hour per patient, with health utility assessments taking approximately 30 minutes on average. Barriers regarding home internet access may have increased the acceptability of utility assessment in the community dialysis clinic setting [27]. Many patients expressed pleasure at the opportunity to engage in a value-added activity while they were "captive" during their 3- to 4-hour dialysis session. This bodes well for plans to conduct shared decision-making visits using The Gambler II platform in these same community dialysis clinics.

Secondary Results

Our second goal was to collect health utilities from patients with dialysis for several ESKD health states relevant to the decision about kidney transplantation. We found a wide variation in health utilities from patient to patient. Consistent with prior studies, utilities assessed using the SG technique were higher than those determined with the VAS or TTO techniques. SG holistically incorporates risk attitude into its assessment, and since most people are risk averse, SG utilities tend to be higher [28]. Of note, there was a wide variation in utilities for transplantation with an HCV-viremic kidney, and in a few instances, patients had lower utilities for this health state than for continued hemodialysis. While mean SG utilities for transplant with an HCV-viremic kidney were lower than hemodialysis (75.5 vs 82.5), standard deviations were large (28 and 23, respectively), which further demonstrated significant differences in utility values across the patients. To this point, 30 (48%) patients had higher SG utilities for transplant with an HCV-viremic kidney compared with hemodialysis. We also explored whether there were racial differences in health state utilities between patients with African American and European American backgrounds. We found African Americans had higher TTO utilities for hemodialysis compared to their European American counterparts. One possible explanation is the phenomenon of accommodation. While quality of life may diminish markedly when patients move from a better state of health to a health state marred by chronic disease or disability, studies have shown that, over time, many patients accommodate to the new health state with an accompanying improvement in assessments of quality of life [29,30]. Indeed, in subanalyses stratified by race, African Americans had a significantly longer dialysis vintage (7.3 years) compared with European Americans (2.8 years; *P*=.04).

During the testing of The Gambler II in a clinical setting, the COVID-19 pandemic interrupted patient recruitment. We performed a subanalysis on patients who enrolled after the COVID-19 interruption. We found that gender and study arm did not have a statistically significant impact on knowledge scores or utilities. Stratifying by race, we found that African Americans had higher SG utilities for all but hemodialysis (Multimedia Appendix 1). Furthermore, African Americans had higher TTO utilities for all 3 health states (Multimedia Appendix 1). African Americans had a mean utility of 87.6 for transplantation with an HCV-viremic kidney compared to a mean utility of 69.3 for European Americans (*P*=.004) during the pandemic. We did not see similar differences among patients in the prepandemic cohort given the small sample size of European Americas (n≤5).

To our knowledge, this is the first paper directly eliciting health utilities from patients with dialysis for transplant with an HCV-viremic kidney. Our major finding is that patients' utilities for this outcome vary dramatically, and that for some patients, the worry and concern associated with even a successful transplant of an HCV-viremic kidney may result in a utility weight lower than their current health state of chronic intermittent hemodialysis. We found a slight negative correlation between the difference between the SG utility for transplant with an HCV-viremic kidney and hemodialysis and dialysis vintage. For patients who have a longer dialysis vintage, SG utility for transplant with an HCV-viremic kidney and hemodialysis had a negative correlation (–0.16). This means that there is a trend toward a more negative view of transplant with an HCV-viremic kidney compared with hemodialysis among patients who have been on dialysis for a longer time. We also examined how patients' history of a failed kidney transplant may affect health utilities. We compared this patient population hemodialysis and transplant with an HCV-viremic kidney SG utilities. There was a compelling but not statistically significant trend toward a larger decrement in this value among patients with prior failed kidney transplant (–15.6 versus –5.5; *P*=.29). However, our cohort only had 9 (14%) patients with a prior transplant.

Regarding our third goal, our study showed that health state videos narrated by patient actors, viewed as part of the process of assessing patients' values and preferences for health outcomes of ESKD, can improve knowledge about these health states. On average, knowledge scores improved by 10 points between the pretest and posttest, demonstrating that these videos can educate patients about the relevant health states as a beneficial side effect of the utility assessment process. While we expected to find a positive impact of using race-matched exemplars compared

with race-mismatched exemplars, there were no significant differences between study arms on knowledge gain. Whether our study simply lacked the power to detect a difference and whether matching patient demographics truly matters is a question for further investigation. Of note, patients watching videos narrated by African American patient actors had greater knowledge gains (median 10.0) than those viewing videos narrated by European American patient actors (median 5.0), although this was not statistically significant (P=.25).

Regardless, patients found the videos informative and impactful. Following the utility assessment process, several patients voiced their appreciation for the video descriptions of health states. Some patients even reported being moved to tears.

Comparison With Prior Work

Prior studies noted the potential benefits of using HCV-viremic kidneys to expand the pool of organs available to hemodialysis patients, thus reducing waiting times for transplantation and providing opportunities for transplantation in patients who might not have received a kidney otherwise. Cost-effectiveness analyses have shown the strategy to be cost-effective at a policy level for the general population of patients with ESKD [13,22]. However, none of these analyses used utility assessments from actual patients for the outcome transplant with an HCV-viremic kidney [31,32]. Given the marked patient-to-patient variability in utilities for the 3 health states we studied, patient preferences for health states must be considered in the shared decision-making process about transplantation with an HCV-viremic versus an HCV-uninfected kidney.

The Gambler II provides a consistent and efficient platform to elicit patient utilities and could be integrated into a tool to facilitate shared decision-making. One could envision such a tool that performs personalized decision analyses, using a combination of individual patient's utilities along with clinical and demographic information needed to estimate organ waiting list times, to provide estimates of quality-adjusted survival or life expectancy with each strategy for that individual patient.

Limitations

Our study had some limitations. Given the demographics of the Cincinnati metropolitan area dialysis clinics in our study, it was difficult to recruit patients with European American racial backgrounds in equal numbers to those with African American backgrounds. In addition, the COVID-19 pandemic negatively impacted our total recruitment due to a pause in study activities between March and October of 2020 [23]. A possible contributing factor to utility weights for transplant with an HCV-viremic kidney being lower than hemodialysis for some patients was the way we assessed utilities for this health state separately from hemodialysis and transplant with an HCV-unexposed kidney. We wanted an explicit measure of the utility differences between transplant with an HCV-unexposed and an HCV-viremic kidney. This measure would include in the SG assessment an estimate of the willingness of patients to risk death to avoid such a transplant. However, the normalized utility values for this health state were not directly assessed compared with transplant with an HCV-unexposed kidney.

Conclusions

In conclusion, we found that it is feasible to use a computer platform to assess patients' utilities for health states related to ESKD in ambulatory community dialysis clinics. Furthermore, we demonstrated that utility weights vary dramatically from person to person. These findings have implications for the future development of shared decision-making tools to aid clinicians and their patients with the challenging question of whether to accept transplantation with HCV-viremic kidneys in patients to reduce waiting times and decrease time spent on hemodialysis. This decision will depend upon both expected organ–waiting list times for HCV-unexposed and HCV-exposed kidneys and individual patient's values and preferences for these relevant health states.

Acknowledgments

We wish to thank the patients who entered as participants in the study as well as Dialysis Center Inc for allowing us access to patients in their facilities. We thank our colleagues in the University of Cincinnati's Department of Biomedical Informatics, Division of Nephrology and Hypertension, Center for Health Informatics, and Cincinnati Children's Hospital Medical Center for their assistance with patient recruitment and technical support.

MHE and AAA have investigator-initiated grant support from Bristol Meyers-Squibb (CV-185-764). MHE also has grant support from the National Institute of Child Health and Human Development (R01HD094213) and the National Center for Advancing Translational Sciences (UL1TR001425). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

Data related to this study and The Gambler software are available to use upon request.

Authors' Contributions

All Members contributed to writing—review and editing conceptualization; AAA conducted formal analysis, and was responsible for methodology, investigation, software, validation, visualization, and preparation of the original draft. HJD was responsible for resources, methodology, supervision, and validation; GRM and CVT conducted supervision of the study; SS provided supervision

and validation; and MHE contributed to supervision, investigation, project administration, validation, as well as writing and preparation of the original draft.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Manuscript appendix containing select demographic information about the patients, the questions asked of patients before and after the assessment, and information given to patients throughout the assessment. It also contains the explanation for the health utility assessments, screenshots of the assessments, and utility information generated from the patient assessment.
[DOCX File , 858 KB - formative_v6i7e33562_app1.docx]

References

1. Nally JV. Chronic kidney disease in African Americans: Puzzle pieces are falling into place. Cleve Clin J Med 2017 Nov 01;84(11):855-862. [doi: 10.3949/ccjm.84gr.17007] [Medline: 29173252]
2. Hall YN, Choi AI, Xu P, O'Hare AM, Chertow GM. Racial Ethnic Differences in Rates and Determinants of Deceased Donor Kidney Transplantation. JASN 2011 Mar 03;22(4):743-751. [doi: 10.1681/asn.2010080819]
3. United SRDS. United States Renal Data System (USRDS) Agreement for Release of Data. American Journal of Kidney Diseases 2006 Jan;47:269-270 [FREE Full text] [doi: 10.1053/j.ajkd.2005.12.019]
4. Arce CM, Goldstein BA, Mitani AA, Lenihan CR, Winkelmayer WC. Differences in Access to Kidney Transplantation between Hispanic and Non-Hispanic Whites by Geographic Location in the United States. CJASN 2013 Oct 10;8(12):2149-2157. [doi: 10.2215/cjn.01560213]
5. Peek ME, Gorawara-Bhat R, Quinn MT, Odoms-Young A, Wilson SC, Chin MH. Patient trust in physicians and shared decision-making among African-Americans with diabetes. Health Commun 2013 Aug;28(6):616-623 [FREE Full text] [doi: 10.1080/10410236.2012.710873] [Medline: 23050731]
6. Barry MJ, Edgman-Levitan S. Shared Decision Making — The Pinnacle of Patient-Centered Care. N Engl J Med 2012 Mar;366(9):780-781. [doi: 10.1056/nejmp1109283]
7. Hart A, Lentine K, Smith J, Miller J, Skeans M, Prentice M, et al. OPTN/SRTR 2019 Annual Data Report: Kidney. Am J Transplant 2021 Feb;21 Suppl 2:21-137 [FREE Full text] [doi: 10.1111/ajt.16502] [Medline: 33595191]
8. Eckman M, Ward J, Sherman K. Cost Effectiveness of Universal Screening for Hepatitis C Virus Infection in the Era of Direct-Acting, Pangenotypic Treatment Regimens. Clin Gastroenterol Hepatol 2019 Apr;17(5):930-939.e9 [FREE Full text] [doi: 10.1016/j.cgh.2018.08.080] [Medline: 30201597]
9. Li A, Cholankeril G, Cheng X, Tan J, Kim D, Toll A, et al. Underutilization of Hepatitis C Virus Seropositive Donor Kidneys in the United States in the Current Opioid Epidemic and Direct-Acting Antiviral Era. Diseases 2018 Jul 10;6(3):62 [FREE Full text] [doi: 10.3390/diseases6030062] [Medline: 29996536]
10. Tolley K. What are health utilities? Bandolier. URL: http://www.bandolier.org.uk/painres/download/What%20is%202009/What_are_health_util.pdf [accessed 2022-07-11]
11. Adejare AA, Eckman MH. Automated Tool for Health Utility Assessments: The Gambler II. MDM Policy Pract 2020 Mar 18;5(1):2381468320914307 [FREE Full text] [doi: 10.1177/2381468320914307] [Medline: 32215320]
12. Szende A, Schaefer C. A taxonomy of health utility assessment methods and the role for uncertainty analysis. Eur J Health Econ 2006 Jun;7(2):147-151. [doi: 10.1007/s10198-005-0334-x] [Medline: 16404619]
13. George N, Liapakis A, Korenblat KM, Li T, Roth D, Yee J, et al. A Patient Decision Support Tool for Hepatitis C Virus and CKD Treatment. Kidney Med 2019 Jul;1(4):200-206 [FREE Full text] [doi: 10.1016/j.xkme.2019.06.003] [Medline: 32734200]
14. Matza LS, Sapra SJ, Dillon JF, Kalsekar A, Davies EW, Devine MK, et al. Health state utilities associated with attributes of treatments for hepatitis C. Eur J Health Econ 2015 Dec 7;16(9):1005-1018 [FREE Full text] [doi: 10.1007/s10198-014-0649-6] [Medline: 25481796]
15. Lenert LA, Ziegler J, Lee T, Unfred C, Mahmoud R. The risks of multimedia methods: effects of actor's race and gender on preferences for health states. J Am Med Inform Assoc 2000 Mar 01;7(2):177-185 [FREE Full text] [doi: 10.1136/jamia.2000.0070177] [Medline: 10730601]
16. Brosius H, Peter C. Exemplification Theory. Int Encycl Media Eff 2017;1:9. [doi: 10.1002/9781118783764.wbieme0062]
17. Dialysis Clinic, Inc. URL: https://www.dciinc.org [accessed 2022-07-19]
18. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
19. Arozullah A, Yarnold P, Bennett C, Soltysik R, Wolf M, Ferreira R, et al. Development and validation of a short-form, rapid estimate of adult literacy in medicine. Med Care 2007 Nov;45(11):1026-1033. [doi: 10.1097/MLR.0b013e3180616c1b] [Medline: 18049342]

20. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM. Measuring Numeracy without a Math Test: Development of the Subjective Numeracy Scale. Med Decis Making 2007 Sep 14;27(5):672-680. [doi: 10.1177/0272989x07304449]
21. Liem YS, Bosch JL, Hunink MGM. Preference-based quality of life of patients on renal replacement therapy: a systematic review and meta-analysis. Value Health 2008 Jul;11(4):733-741 [FREE Full text] [doi: 10.1111/j.1524-4733.2007.00308.x] [Medline: 18194399]
22. McCauley M, Mussell A, Goldberg D, Sawinski D, Molina R, Tomlin R, et al. Race, Risk, and Willingness of End-Stage Renal Disease Patients Without Hepatitis C Virus to Accept an HCV-Infected Kidney Transplant. Transplantation 2018 Apr;102(4):e163-e170. [doi: 10.1097/TP.0000000000002099] [Medline: 29346260]
23. Ohio Issues "Stay at Home" Order; New Restrictions Placed on Day Cares for Children. Innovate Ohio Platform. URL: https://governor.ohio.gov/wps/portal/gov/governor/media/news-and-media/ohio-issues-stay-at-home-order-and-new-restrictions-placed-on-day-cares-for-children [accessed 2020-12-12]
24. R: A Language and Environment for Statistical Computing Internet. R Foundation for Statistical Computing. 2018. URL: https://www.R-project.org [accessed 2022-07-08]
25. Perez F, Granger BE. IPython: A System for Interactive Scientific Computing. Comput. Sci. Eng 2007 May;9(3):21-29. [doi: 10.1109/mcse.2007.53]
26. Pollard T, Johnson A, Raffa J, Mark R. An open source Python package for producing summary statistics for research papers. JAMIA Open 2018 Jul;1(1):26-31 [FREE Full text] [doi: 10.1093/jamiaopen/ooy012] [Medline: 31984317]
27. Lockwood MB, Dunn-Lopez K, Burke L, Becker YT, Saunders M. Frequency of In-Home Internet Use Among Prekidney and Postkidney Transplant Patients—Facilitators and Barriers to Use and Trends Over Time. Transplantation Direct 2017 Nov;3(11):e216. [doi: 10.1097/txd.0000000000000735]
28. Rosen AB, Tsai JS, Downs SM. Variations in Risk Attitude across Race, Gender, and Education. Med Decis Making 2016 Jul 01;23(6):511-517. [doi: 10.1177/0272989x03258431]
29. Sung L, Young N, Greenberg M, McLimont M, Samanta T, Wong J, et al. Health-related quality of life (HRQL) scores reported from parents and their children with chronic illness differed depending on utility elicitation method. J Clin Epidemiol 2004 Nov;57(11):1161-1166. [doi: 10.1016/j.jclinepi.2004.05.003] [Medline: 15567632]
30. Wachterman M, McCarthy E, Marcantonio E, Ersek M. Mistrust, misperceptions, and miscommunication: a qualitative study of preferences about kidney transplantation among African Americans. Transplant Proc 2015 Mar;47(2):240-246 [FREE Full text] [doi: 10.1016/j.transproceed.2015.01.016] [Medline: 25769556]
31. Eckman MH, Woodle ES, Thakar CV, Alloway RR, Sherman KE. Cost-effectiveness of Using Kidneys From HCV-Viremic Donors for Transplantation Into HCV-Uninfected Recipients. Am J Kidney Dis 2020 Jun;75(6):857-867. [doi: 10.1053/j.ajkd.2019.11.005] [Medline: 32081494]
32. Scott N, Snell G, Westall G, Pilcher D, Raggatt M, Walker RG, et al. Cost-effectiveness of transplanting lungs and kidneys from donors with potential hepatitis C exposure or infection. Sci Rep 2020 Jan 29;10(1):1459 [FREE Full text] [doi: 10.1038/s41598-020-58215-z] [Medline: 31996734]

Abbreviations

DCI: Dialysis Center Inc
ESKD: end-stage kidney disease
HCV: hepatitis C virus
HUA: health utility assessment
REDCap: Research Electronic Data Capture
SDM: shared decision-making
SG: standard gamble
TTO: time trade-off
VAS: visual analog scale

Edited by A Mavragani; submitted 13.09.21; peer-reviewed by J Edwards, R Armstrong Junior; comments to author 29.05.22; revised version received 14.06.22; accepted 24.06.22; published 28.07.22.

Please cite as:
Adejare AA, Duncan HJ, Motz RG, Shah S, Thakar CV, Eckman MH
Implementing a Health Utility Assessment Platform to Acquire Health Utilities in a Hemodialysis Outpatient Setting: Feasibility Study
JMIR Form Res 2022;6(7):e33562
URL: https://formative.jmir.org/2022/7/e33562
doi:10.2196/33562
PMID:35900828

Original Paper

An mHealth Intervention to Improve Pre-Exposure Prophylaxis Knowledge Among Young Black Women in Family Planning Clinics: Development and Usability Study

Amy K Johnson[1,2], MSW, PhD; Sadia Haider[3], MD; Katie Nikolajuk[1], MSW; Lisa M Kuhns[1,2], MPH, PhD; Emily Ott[3], MA; Darnell Motley[4], PhD; Brandon Hill[5], PhD; Lisa Hirschhorn[2], MD

[1]Division of Adolescent and Young Adult Medicine, Ann & Robert H Lurie Children's Hospital of Chicago, Chicago, IL, United States

[2]Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[3]Department of Obstetrics & Gynecology, Rush University, Chicago, IL, United States

[4]Department of Medicine, Chicago Center for HIV Elimination, University of Chicago, Chicago, IL, United States

[5]Howard Brown Health, Chicago, IL, United States

Corresponding Author:
Amy K Johnson, MSW, PhD
Division of Adolescent and Young Adult Medicine
Ann & Robert H Lurie Children's Hospital of Chicago
225 E Chicago Box 161
Chicago, IL, 60611
United States
Phone: 1 3122277733
Email: akjohnson@luriechildrens.org

Abstract

Background: Young Black women between the ages of 18 and 24 years are disproportionately impacted by HIV, yet they have a low self-perception of HIV risk and limited exposure to prevention strategies. Pre-exposure prophylaxis (PrEP) is a safe and effective biomedical HIV prevention strategy for those at risk for HIV infection, but uptake has been slow among cisgender women. Family planning clinics are a primary source of health care access for young women, providing an ideal opportunity to integrate PrEP information and care into existing clinic practices.

Objective: The aim of this study was to use a multistage, community-engaged process to develop a mobile health app and to evaluate the feasibility and acceptability of the app.

Methods: Using user-centered design, the *In the Loop* app was developed in collaboration with a community advisory board of young Black women. This study employed a multistage design, which included community-engaged app development, user testing, and evaluation of the app's feasibility and acceptability. A pre- and postdesign was used to assess the impact of the app on PrEP knowledge immediately after app use. Descriptive statistics (eg, mean, SD, and percentage values) were used to describe the sample, and Wilcoxon matched-pairs signed-ranks test was used to detect changes in PrEP knowledge before and immediately after using the app.

Results: A total of 50 sexually active, young Black women, aged 18-24 (mean 21, SD 1.9) years, were enrolled in this study. Analysis comparing scores before and immediately after use of the app revealed a significant increase in PrEP content knowledge scores on a 7-item true or false scale (z=–6.04, P<.001). Overall, participants considered the *In the Loop* app feasible and acceptable to use while waiting for a family planning visit. The majority of participants (n=46, 92%) agreed that they would recommend *In the Loop* to friends to learn more about PrEP. Participants rated the overall quality of the app 4.3 on a 1-5 scale (1=very poor and 5=very good). Of 50 participants, 40 (80%) agreed that the app was easy to use, and 48 (96%) agreed that they found the information in the app easy to understand. Finally, 40 (80%) agreed that they had enjoyed using the app while waiting for their family planning visit.

Conclusions: Our findings suggest that young Black women waiting for family planning visits found the *In the Loop* app to be feasible and acceptable. This study demonstrates the value of engaging young Black women in the app design process. As family planning clinics are a primary source of health care access for young women, they provide an ideal setting to integrate PrEP information and care into existing clinic practices. Next steps in the development of the *In the Loop* app include implementing

user-suggested improvements and conducting efficacy testing in a randomized controlled trial to determine the app's impact on PrEP uptake.

(JMIR Form Res 2022;6(7):e37738) doi:10.2196/37738

KEYWORDS

mHealth; adolescent health; young Black women; pre-exposure prophylaxis; HIV; mobile health; PrEP; mobile app

Introduction

Approximately 23% of all people living with HIV in the United States are cisgender women, and women accounted for 19% of new infections in 2018 [1]. The majority (87%) of HIV infections among women are due to heterosexual sex [1]. Black women are disproportionately impacted by HIV, with an 18-fold higher risk of acquiring HIV compared to White women [2]. In Chicago, 85.9% of new HIV infections among women occur among non-Hispanic Black women [3]. In particular, young Black women aged 18-24 years often have multiple co-occurring risk factors (eg, high HIV density in sexual networks, concurrent sexual partners, experiences of violence, substance use, and sex partners with incarceration history), yet they have a low self-perception of HIV risk, complicating HIV prevention efforts [3-7].

Pre-exposure prophylaxis (PrEP) is a safe and effective biomedical HIV prevention strategy for those at risk for HIV infection, including young Black women [8,9]. PrEP has the advantage over other prevention interventions of allowing women to have autonomy and control over their sexual health because it does not require a sexual partner's permission or cooperation for use [10]. However, women make up a disproportionately low percentage of PrEP users in the United States [11]. The Centers for Disease Control and Prevention estimates that 468,000 women in the United States are eligible for PrEP, but only 19,000 women have ever been prescribed PrEP [11-13]. From 2014 to 2016, only 2% of women who have indication for PrEP received a prescription, and women accounted for less than 5% of all PrEP users in 2016 [10]. Chicago is a leading city in PrEP implementation, but uptake of PrEP among women is extremely low. For example, of the patients who have initiated PrEP in sexually transmitted infections (STI) clinics of the Chicago Department of Public Health, only 1.5% are women [2,14]. Known individual-level barriers to PrEP uptake among women include low rates of PrEP knowledge, low perceived risk of HIV acquisition, and lack of awareness of how and where to access PrEP [2,14]. Increasing knowledge, risk perception, awareness, and uptake of PrEP among young Black women who may be at an increased risk for HIV are key public health interventions that have the potential to reduce new HIV infections [15].

Family planning clinics are primary sources of health care access for young women, providing an ideal opportunity to integrate PrEP information and care into existing clinic practices [16,17]. A total of 60% of women, including young Black women, consider family planning clinics their primary source of medical care, with 40% indicating it is their only source of health care [17]. In 2018, Title X family planning clinics served over 3.9 million patients, most of whom were young (63%), female (87%), and had low income (65%) [18]. Over 1 million HIV tests were performed in Title X clinics in 2018, establishing a precedent for receiving HIV prevention services at family planning clinics [18]. Optimizing PrEP awareness via integration with family planning services has the potential to increase PrEP uptake and impact incidence on a population level, but it has not yet been fully explored or implemented.

The ubiquitous nature of new digital media, characterized by its adaptability and interactivity, offers opportunities to disseminate confidential information to adolescents in a relevant, youth-friendly format. Mobile health (mHealth) approaches have been proven to be most impactful when integrated into existing health system functions, rather than as stand-alone solutions [19,20]. Given the limited provider time per patient, brief theory-driven mHealth intervention models delivered in family planning clinics waiting rooms have the potential to aid patients in increasing knowledge, informing decision-making, and framing questions to providers [21-23].

Researchers have identified the lack of tailored apps for minority communities as an ongoing challenge in digital media intervention [24,25]. To date, no studies have been conducted to design, target, and promote an mHealth intervention to enhance HIV risk assessment, increase PrEP knowledge, and influence PrEP uptake among young Black women within a family planning setting. Patients prefer health information to be interactive [26]; a waiting room–based mHealth intervention uses a current missed opportunity to enhance HIV risk assessment and provide patient education about PrEP, and it has the potential for high impact [22,23]. This pilot study is the first step in designing and implementing an mHealth HIV prevention and PrEP promotion intervention tailored for young Black women.

Methods

This pilot study employed a multistage design that included community-engaged app development, user testing, and evaluation of the app's feasibility and acceptability.

Ethics Approval

All procedures were approved by the Institutional Review Board at Lurie Children's Hospital and the University of Chicago (2018-2189).

Setting

The 2017 rate (27.9 per 100,000 population) of HIV infection diagnoses in Chicago is approximately 2.5 times higher than the national rate, and the prevalence rate for Chicago (827.9 per 100,000 population) is nearly 3 times the national rate [27]. Community areas on the West and South sides of Chicago represent a concentrated HIV epidemic, with HIV prevalence

consistently over 5% [3,27]. The study recruitment site was a Planned Parenthood Illinois clinic that is located centrally within the city and serves predominantly patients of color.

In the Loop App Prototype Development

A community advisory board (CAB) of 9 young Black women was formed to guide and tailor content of the app prototype. CAB participants were recruited through partner agencies, email listservs, and flyers posted throughout the community. Recruitment flyers included information about the project, participation expectations, and learning opportunities (ie, the app design process). Our goal was to have a group of 8-10 young women who met the inclusion criteria for the larger pilot study to participate in the CAB. CAB members were provided with a US $50 incentive for each 2-hour meeting they attended.

Guided by user-centered design and the information systems research frameworks, the study team conducted relevance, rigor, and design cycles to refine and adapt an existing prototype app called *miPrEP* [28]. The *miPrEP* app was developed to increase PrEP knowledge and engagement among young Black men who have sex with men [29]. The CAB conducted 4 design sprints over a 3-month period; sessions were guided by generating app content and approving mock-ups (ie, iterative design process). In the rigor cycle, the CAB members met and reviewed the available components of the *miPrEP* app. The *miPrEP* app contained basic PrEP information, as well as a video about how PrEP works in the body. The study team identified and presented additional health-based apps to the CAB for review. Existing content and desired features of an adapted sexual health app were discussed and ranked. Finally, features and content were identified for inclusion and tailoring for the *In the Loop* prototype. In the design cycle, a low fidelity prototype was developed. Paper and pencil mock-ups were used to design the visual assets, video content, and text materials. The prototype of *In the Loop* was created using paper models to represent each intervention component and a Google slide deck to demonstrate app navigation and flow. All elements were iteratively revised with the CAB until there was consensus on the content, quantities, and the form of the app.

User Testing and Pilot Evaluation

Potential participants were recruited from a Planned Parenthood clinic waiting room and, if interested and eligible, they completed informed consent and the study procedures. Participants were eligible if they (1) reported being cisgender women; (2) identified as African American or Black; (3) were English speakers; (4) were aged 18-24 years (inclusive); (5) reported vaginal or anal sex with a male partner within the past 6 months; (6) were seeking testing or treatment for a sexually transmitted infection, or seeking contraception or abortion services; (7) self-reported being HIV-negative; (8) were not currently pregnant; and (9) were neither currently taking PrEP as HIV prevention nor intending to initiate PrEP at the current clinic appointment.

A pre- and postdesign assessment was used in this pilot study to evaluate the impact of the app on PrEP knowledge immediately after using the app. PrEP knowledge was measured via 7 true or false items that were selected based on use in previous studies with cisgender women [14,30]. Other study aims were to assess usability, including basic feasibility and acceptability of the app prototype among young Black women. Usability was assessed through open-ended questions that asked participants their opinion about the app (eg, “what did you like about using *In the Loop* while waiting for your appointment?” and “what information would you add to *In the Loop* app?”) and ways to improve the app (eg, text or graphic display, navigation, and content). Acceptability of *In the Loop* was evaluated using a series of 9 items derived from an abbreviated acceptability rating scale used in prior research [31]. Participants used a 5-point Likert scale (1=strongly disagree and 5=strongly agree) to rate the extent to which they agreed with each acceptability statement (eg, “I found the information in the app easy to understand and comprehend”). The measure was administered at the immediate postintervention assessment. Overall interest in using mHealth for sexual health was measured via a 7-item scale based on previous research with adolescents (eg, “how likely are you to download a sexual health mobile phone app?”) [32-34]. Finally, participants were asked about preferences toward emerging modalities of PrEP, specifically their interest in a vaginal ring, implant, or injection, or PrEP combined with birth control [35].

Analysis

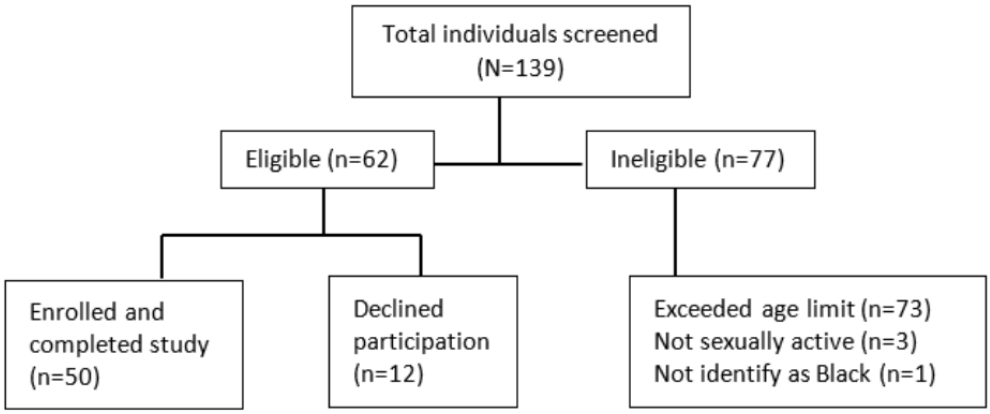
Descriptive statistics (eg, mean, SD, and percentage values) were used to describe the sample, and Wilcoxon matched-pairs signed-ranks test was used to detect changes in PrEP knowledge before and immediately after using the app [36]. Open-ended responses were recorded verbatim and were thematically coded by the first author (AKJ), following a standardized procedure [37]. Each open-ended response was read twice, themes were then created based on response content, and each response was coded into one of 4 main themes. If the response contained more than one theme, it was coded as having both themes.

Results

Participant Characteristics

The *In the Loop* prototype was evaluated from July 2019 through August 2019. A total of 50 sexually active, young Black women aged 18-24 (mean 21, SD 1.9) years were enrolled in the study. Initially, a total of 139 individuals were screened, with 62 being eligible; of those eligible, 50 (81%) enrolled and completed the study visit (Figure 1).

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



Of the 50 participants, 23 (46%) reported being current students, with 21 (42%) participants reporting their highest level of education as some college education, 19 (38%) as high school graduate, and 9 (18%) of them reported having completed college (Table 1). The majority of participants (n=37, 74%) were employed either full- or part-time. To characterize potential social determinants of health that impact HIV transmission, participants reported whether they had ever received government assistance (n=19, 38%) and whether they had ever been homeless (n=11, 22%).

A total of 44 (88%) participants reported identifying as straight or heterosexual, and 6 (12%) reported their sexual orientation as bisexual. In terms of sexual behavior, the majority (n=42, 84%) reported having 1 sex partner in the past 3 months. Inconsistent condom use was common, with 43 (86%) participants reporting sometimes or never using condoms during vaginal sex; 9 (18%) participants reported having sex with a person whose HIV status they did not know; 6 (12%) of them had a recent STI diagnosis, and 24 (48%) had experienced an unplanned pregnancy.

Table 1. Participant characteristics (N=50).

Characteristics	Values
Age, mean (SD)	21 (1.9)
Sexual orientation, n (%)	
Straight	44 (88)
Bisexual	6 (12)
Hispanic or Latina ethnicity, n (%)	
No	47 (94)
Yes	3 (6)
Education level, n (%)	
Some high school	1 (2)
High school or General Education Diploma	19 (38)
Some college or Associate of Arts degree	21 (42)
Completed college or Bachelor of Arts degree	9 (18)
Current student, n (%)	
Yes	23 (46)
No	27 (54)
Received government aid, n (%)	
Yes	19 (38)
No	29 (58)
Unsure	2 (4)
Employment, n (%)	
No	13 (26)
Part-time	23 (46)
Full-time	14 (28)
Ever been homeless, n (%)	
Yes	11 (22)
No	39 (78)
Sexual behavior in the past 3 months, n (%)	
1 sex partner	42 (84)
2 or more sex partners with no overlap	6 (12)
2 or more sex partners with overlap	2 (4)
Condom use in vaginal sex, n (%)	
Always	6 (12)
Sometimes	25 (50)
Never	18 (36)
Condom use in anal sex, n (%)	
Always	1 (2)
Sometimes	4 (8)
Never	5 (10)
Did not have anal sex	40 (80)
Had sex with a person whose HIV status you did not know, n (%)	
Yes	9 (18)
No	41 (82)

Characteristics	Values
Had a sexually transmitted infection, n (%)	
Yes	6 (12)
No	44 (88)
Unplanned pregnancies, n (%)	
Yes	24 (48)
No	26 (52)
How likely are you to become infected with HIV based on current behavior, n (%)	
Very unlikely	22 (44)
Unlikely	22 (44)
Somewhat likely	6 (12)

PrEP Knowledge

Prior to using the app, 30 (60%) participants reported having "never heard of PrEP" and 44 (88%) considered themselves as "very unlikely or unlikely" at risk for HIV based on their current sexual behaviors. Analysis comparing scores before and immediately after using the app revealed a significant increase in PrEP content knowledge scores on a 7-item true or false scale. Total mean scores improved from 2.8 (SD 1.2) prior to using the app to 4.9 (SD 1.4) immediately after using the app (z=–6.04, P<.001). In item-by-item analysis, 4 of 7 items showed statistically significant improvements from before to immediately after app use (Table 2). After using the app, 13 (26%) participants reported they were "very likely" to start PrEP in the next 3 months, and 46 (92%) participants agreed with the statement "PrEP is a good option for HIV prevention for young women like me."

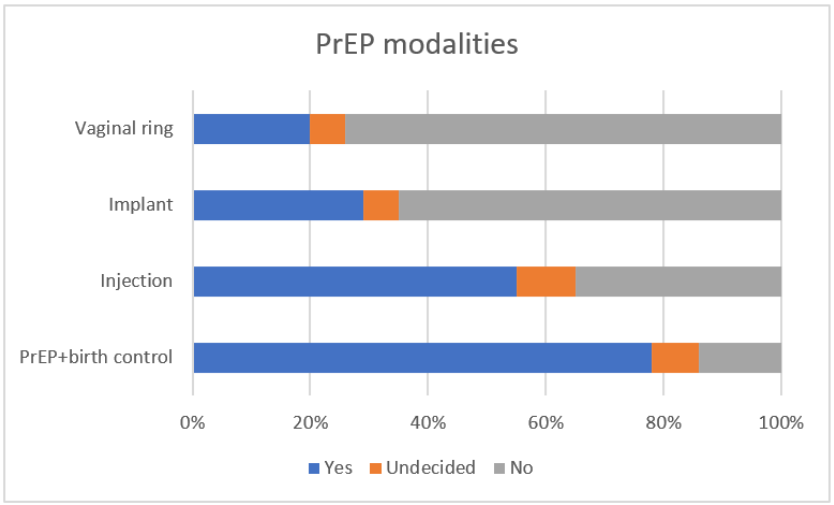
In the immediate postsurvey, participants were asked whether they were interested in other PrEP modalities (Figure 2), including vaginal ring, implant, or injection, or PrEP combined with birth control. The most frequently endorsed mode was a dual method option with PrEP and birth control combined; 39 (78%) participants endorsed this option.

Table 2. Pre-exposure prophylaxis (PrEP) knowledge before and immediately after use of the app (N=50).

True or false items	Negative ranks, n (%)	Ties, n (%)	*Z* scores
PrEP can be bought over the counter without a prescription	6 (12)	39 (78)	–0.63
PrEP needs to be taken daily in order to be effective	13 (26)	36 (72)	–3.60[a]
PrEP is only recommended for gay and bisexual men	42 (84)	7 (14)	–6.48[a]
PrEP can be used to prevent infection after being exposed to HIV	29 (58)	20 (40)	–5.39[a]
PrEP can be taken by both men and women to prevent HIV infection	4 (8)	44 (88)	–1.34
You need to be at least 18 years of age to get PrEP	11 (22)	37 (74)	–2.89[a]
Only HIV doctors can prescribe PrEP	9 (18)	37 (74)	–1.73

[a]P<.001.

Figure 2. Interest in alternate PrEP modalities (N=50). PrEP: pre-exposure prophylaxis.



Technology Use for Sexual Health

When participants were asked about interest levels in using technology as part of sexual health care, 22 (44%) indicated they were very interested, and 21 (42%) reported being somewhat interested. A total of 31 (63%) indicated they were very interested in using a mobile app for sexual health education, 10 (21%) indicated they were somewhat interested, and 8 (16%) indicated no interest. Just over half (n=26, 52%) of the sample reported being very interested in receiving sexual health reminders via email (n=29, 57%) or text message (n=25, 51%). Finally, only 17 (35%) participants reported being very interested in learning about sexual health topics via internet games.

In the Loop App Acceptability and Usability

Overall, participants considered the *In the Loop* app feasible and acceptable to use while waiting for a family planning visit. More specifically, 46 (92%) participants agreed that they would recommend *In the Loop* to friends to learn more about PrEP. Participants also reported that they liked using the app—rated 4.6 on a scale from 1 (strongly disagree) to 5 (strongly agree). Participants rated the overall quality of the app 4.3 on a scale from 1 (very poor) to 5 (very good); 40 (80%) participants agreed that the app was easy to use, and 48 (96%) agreed that they found the information in the app easy to understand. Finally, 40 (80%) agreed that they had enjoyed using the app while waiting for their family planning visit.

Participants were also asked to provide open-ended feedback on improvements that could be made to the app, which were summarized into four different themes to inform future refinement: increasing interactive components, refining app navigation, refining app aesthetics, and increasing sexual health and HIV prevention information (Table 3). Finally, participants responded to an open-ended question providing insight to their favorite parts of the app, including liking the fact that the main character is a young Black woman (“I liked that it was relatable. I liked that there was a character that looked like me” [participant ID, ie, PID 66]), that it did not require a lot of time to use (“It was quick. Wasn’t long. Wasn’t boring” [PID 15]), and that it was private or confidential (“Private on the iPad” [PID 50]).

Two functional bugs were identified in testing the app with users. The first bug caused a video to stall if the navigate back button was pushed; the second bug caused an incorrect follow-up prompt to be displayed. Both bugs, once identified, were remedied.

Table 3. Suggested app improvements.

Theme	Illustrative quote
Increase interactive components	• "It was kind of wordy. They should highlight words, so it won't look like a whole paragraph. The paragraph is intimidating. Should be more interactive." [PID 71]
Refine app navigation	• "Pointing out things, like for each section, so they know what they can press and where to move on, or how to leave the video." [PID 60]
Refine app esthetics	• "The blue is okay, but add more bright colors…like when you are navigating change the colors up." [PID 82]
Increase sexual health and HIV prevention information	• "Add place that people can go to get STI[a] testing or on PrEP[b]. Like if you want to go somewhere else you put in your zipcode and go." [PID 28] • "More about mechanisms or each birth control option; how some of these birth controls are inserted; how does it work?" [PID 16]

[a]STI: sexually transmitted infection.

[b]PrEP: pre-exposure prophylaxis.

Discussion

Principal Findings

This study demonstrates that using an app in a family planning clinic waiting room is feasible and acceptable to young Black women. As family planning clinics are primary sources of health care for young Black women, integrating PrEP services into their practices is a critical component of comprehensive reproductive health care. Further, providing high-quality, confidential information via an mHealth app while waiting for clinic visits makes the most of a current missed opportunity. The *In the Loop* app is an innovative strategy to engage young Black women in HIV prevention.

In the Loop followed key tenets of developing effective mHealth interventions, including incorporating ease of modification and tailoring to the target population. Our team, in collaboration with young Black women, modified the content of an existing PrEP knowledge app and tailored it to be specific to the needs of the target population. Through user testing, we identified and addressed 2 bugs in the app. Finally, we gathered feedback on further refinements such as increasing the interactive components and refining the app navigation. The majority of participants were interested in using technology as part of sexual health care, with an app being the most popular mode of technology selected.

Comparison With Prior Work

Cordova et al [38] designed and evaluated an mHealth intervention—*Storytelling 4 Empowerment*—to provide HIV, STI, and drug abuse prevention information to adolescents in primary care waiting rooms. In the acceptability pilot, participants reported approval of using down time while waiting for clinic visits; however, they also highlighted the need for privacy in shared spaces. Participants recommended privacy covers over tablets. In our pilot study, participants noted that the app felt private and did not raise concerns over privacy of participating in an mHealth app in a family planning clinic waiting room; nevertheless, it should be considered best practice to use privacy screens on tablets as recommended by Cordova et al [38].

In a systematic review, Chávez et al [39] highlighted the fact that apps tailored to adolescents' specific race, ethnicity, and gender identity demonstrated stronger effects on health behavior compared to apps that were not tailored. Results from our community-engaged app development demonstrated high levels of app acceptability among the target population. Not only is the tailoring of mHealth important to success, but also the inclusion of the target population in mHealth development is critical for ensuring salience and acceptability.

In a 2020 systematic review of mHealth strategies to promote PrEP uptake, only 1 app was identified that was tailored specifically for young women; it was designed for those aged 11-14 years in Western Kenya [40,41]. No apps were identified that were tailored for young Black women in the United States. The results of the review highlight the need for mHealth strategies designed specifically to meet the needs of young Black women. We believe that our app will help to fill this gap.

Limitations

Results of this pilot study must be interpreted in light of several limitations. First, the app was tailored exclusively for young Black women in a single urban setting, and thus our results cannot be generalized to other populations of young Black women. Further, the intervention was tailored for young Black women who have sex with men; further research is needed to tailor the intervention for young Black women who have sex with women or transgender and gender-diverse individuals. Second, the high positive rating of the app may be biased due to social desirability. We attempted to mitigate the effect of this bias by using computer-assisted questionnaires. Finally, although *In the Loop* was found to be acceptable and feasible, and to have an immediate impact on PrEP knowledge, future studies are needed to determine the sustained effect of the app on PrEP knowledge and to determine whether the app can improve PrEP uptake. Our results are limited, as they reflect prescores and immediate postscores without randomization; future studies should use randomization and assess durable effects of the intervention. Despite these limitations, this formative research may help guide future design and implementation of mHealth

interventions in family planning settings, optimizing their chances for success.

Conclusions

Overall, our findings suggest that young Black women waiting for family planning visits found *In the Loop* to be feasible and acceptable. Further, this study demonstrates the value of engaging young Black women in the design process. As family planning clinics are primary sources of health care access for young women, they provide an ideal setting to integrate PrEP information and care into existing clinic practices. mHealth approaches can be scaled rapidly with fidelity. Future studies should evaluate the efficacy and durability of the *In the Loop* app in improving PrEP knowledge and uptake among PrEP-eligible young Black women.

Acknowledgments

The authors would like to thank staff at University of Chicago's Ci3 Lab who created the digital esthetics and the app prototype. The authors would also like to thank the community advisory board and study participants for their time and expertise. Finally, the authors would like to acknowledge the partnership of Planned Parenthood, Illinois, where this pilot study was conducted.

This research was made possible with support from the Third Coast Center for AIDS Research, a center funded by National Institutes of Health (P30 AI117943).

Conflicts of Interest

None declared.

References

1. HIV surveillance report. Centers for Disease Control and Prevention. 2017. URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-2017-vol-29.pdf [accessed 2022-06-16]
2. Goparaju L, Praschan NC, Warren-Jeanpiere L, Experton LS, Young MA, Kassaye S. Stigma, partners, providers and costs: potential barriers to PrEP Uptake among US Women. J AIDS Clin Res 2017 Sep;8(9):730 [FREE Full text] [doi: 10.4172/2155-6113.1000730] [Medline: 29201531]
3. HIV/STI surveillance report. Chicago Department of Public Health. 2017. URL: https://www.chicago.gov/content/dam/city/depts/cdph/HIV_STI/HIV_STISurveillanceReport2016_12012017.pdf [accessed 2022-06-16]
4. Grieb SMD, Davey-Rothwell M, Latkin CA. Social and sexual network characteristics and concurrent sexual partnerships among urban African American high-risk women with main sex partners. AIDS Behav 2012 May 23;16(4):882-889 [FREE Full text] [doi: 10.1007/s10461-011-0030-z] [Medline: 21861193]
5. Ivy W, Miles I, Le B, Paz-Bailey G. Correlates of HIV infection among African American women from 20 cities in the United States. AIDS Behav 2014 Apr 28;18 Suppl 3(S3):266-275 [FREE Full text] [doi: 10.1007/s10461-013-0614-x] [Medline: 24077972]
6. HIV among African Americans. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchhstp/newsroom/docs/factsheets/cdc-hiv-aa-508.pdf [accessed 2022-06-16]
7. Hodder SJ, Justman J, Haley DF, Adimora AA, Fogel CI, Golin CE, HIV Prevention Trials Network Domestic Prevention in Women Working Group. Challenges of a hidden epidemic: HIV prevention among women in the United States. J Acquir Immune Defic Syndr 2010 Dec;55 Suppl 2:S69-S73 [FREE Full text] [doi: 10.1097/QAI.0b013e3181fbbdf9] [Medline: 21406990]
8. Murnane PC, Celum C, Mugo N, Campbell JD, Donnell D, Bukusi E, Partners PrEP Study Team. Efficacy of preexposure prophylaxis for HIV-1 prevention among high-risk heterosexuals: subgroup analyses from a randomized trial. AIDS 2013 Aug 24;27(13):2155-2160 [FREE Full text] [doi: 10.1097/QAD.0b013e3283629037] [Medline: 24384592]
9. Abdool Karim Q, Abdool Karim SS, Frohlich JA, Grobler AC, Baxter C, Mansoor LE, CAPRISA 004 Trial Group. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. Science 2010 Sep 03;329(5996):1168-1174 [FREE Full text] [doi: 10.1126/science.1193748] [Medline: 20643915]
10. Seidman DW, Weber S. Integrating preexposure prophylaxis for human immunodeficiency virus prevention into women's health care in the United States. Obstet Gynecol 2016 Jul;128(1):37-43. [doi: 10.1097/AOG.0000000000001455] [Medline: 27275793]
11. Siegler AJ, Mouhanna F, Giler RM, Weiss K, Pembleton E, Guest J, et al. The prevalence of pre-exposure prophylaxis use and the pre-exposure prophylaxis-to-need ratio in the fourth quarter of 2017, United States. Ann Epidemiol 2018 Dec;28(12):841-849 [FREE Full text] [doi: 10.1016/j.annepidem.2018.06.005] [Medline: 29983236]
12. Smith DK, Van Handel M, Wolitski RJ, Stryker JE, Hall HI, Prejean J, et al. Vital signs: estimated percentages and numbers of adults with indications for preexposure prophylaxis to prevent HIV acquisition--United States, 2015. MMWR Morb Mortal Wkly Rep 2015 Nov 27;64(46):1291-1295 [FREE Full text] [doi: 10.15585/mmwr.mm6446a4] [Medline: 26606148]

13. Smith DK, Van Handel M, Grey J. Estimates of adults with indications for HIV pre-exposure prophylaxis by jurisdiction, transmission risk group, and race/ethnicity, United States, 2015. Annals of Epidemiology 2018 Dec;28(12):850-857.e9. [doi: 10.1016/j.annepidem.2018.05.003] [Medline: 29941379]
14. Hirschhorn LB, Brown RN, Friedman EE, Greene G, Bender A, Christeller C, et al. Black cisgender women's PrEP knowledge, attitudes, preferences, and experience in Chicago. J Acquir Immune Defic Syndr 2020 Aug 15;84(5):497-507 [FREE Full text] [doi: 10.1097/QAI.0000000000002377] [Medline: 32692108]
15. Auerbach JD, Kinsky S, Brown G, Charles V. Knowledge, attitudes, and likelihood of pre-exposure prophylaxis (PrEP) use among US women at risk of acquiring HIV. AIDS Patient Care STDS 2015 Feb;29(2):102-110 [FREE Full text] [doi: 10.1089/apc.2014.0142] [Medline: 25513954]
16. Fuentes L, Ingerick M, Jones R, Lindberg L. Adolescents' and young adults' reports of barriers to confidential health care and receipt of contraceptive services. J Adolesc Health 2018 Jan;62(1):36-43 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.10.011] [Medline: 29157859]
17. Kavanaugh ML, Jerman J. Contraceptive method use in the United States: trends and characteristics between 2008, 2012 and 2014. Contraception 2018 Jan;97(1):14-21 [FREE Full text] [doi: 10.1016/j.contraception.2017.10.003] [Medline: 29038071]
18. Family planning annual report summary. US Department of Health & Human Services. URL: https://www.hhs.gov/opa/title-x-family-planning/fp-annual-report/fpar-infographic/index.html [accessed 2019-08-30]
19. Kay MS, Takane M. mHealth: new horizons for health through mobile technologies. World Health Organization. 2011. URL: http://apps.who.int/iris/bitstream/handle/10665/44607/9789241564250_eng.pdf [accessed 2022-06-16]
20. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? PLoS Med 2013 Feb 12;10(2):e1001382 [FREE Full text] [doi: 10.1371/journal.pmed.1001382] [Medline: 23424286]
21. Gignon M, Idris H, Manaouil C, Ganry O. The waiting room: vector for health education? The general practitioner's point of view. BMC Res Notes 2012 Sep 18;5(1):511 [FREE Full text] [doi: 10.1186/1756-0500-5-511] [Medline: 22988947]
22. Gilliam ML, Martins SL, Bartlett E, Mistretta SQ, Holl JL. Development and testing of an iOS waiting room "app" for contraceptive counseling in a Title X family planning clinic. Am J Obstet Gynecol 2014 Nov;211(5):481.e1-481.e8. [doi: 10.1016/j.ajog.2014.05.034] [Medline: 24881829]
23. Lyles CR, Altschuler A, Chawla N, Kowalski C, McQuillan D, Bayliss E, et al. User-centered design of a tablet waiting room tool for complex patients to prioritize discussion topics for primary care visits. JMIR Mhealth Uhealth 2016 Sep 14;4(3):e108 [FREE Full text] [doi: 10.2196/mhealth.6187] [Medline: 27627965]
24. Gilliam M, Hebert L, Brown R, Akinola M, Hill B, Whitaker A, et al. Exploring the feasibility and effectiveness of a contraceptive counseling waiting room app. Contraception 2016 Oct;94(4):412. [doi: 10.1016/j.contraception.2016.07.110]
25. Guse K, Levine D, Martins S, Lira A, Gaarde J, Westmorland W, et al. Interventions using new digital media to improve adolescent sexual health: a systematic review. J Adolesc Health 2012 Dec;51(6):535-543. [doi: 10.1016/j.jadohealth.2012.03.014] [Medline: 23174462]
26. Akinola M, Hebert LE, Hill BJ, Quinn M, Holl JL, Whitaker AK, et al. Development of a mobile app on contraceptive options for young African American and Latina women. Health Educ Behav 2019 Feb 13;46(1):89-96. [doi: 10.1177/1090198118775476] [Medline: 29896969]
27. HIV/STI surveillance report. Chicago Department of Public Health. 2018. URL: https://www.chicago.gov/content/dam/city/depts/cdph/infectious_disease/STI_HIV_AIDS/HIVSTI_SURVEILLANCE_REPORT_2018_1272018.pdf [accessed 2022-06-16]
28. Cronholm S, Göbel H. Evaluation of the information systems research frameworkmpirical evidence from a design science research project. EJISE 2016;19:157-167 [FREE Full text]
29. Hill BB, Schneider J, Rosebush J, Richards M. Evaluating the feasibility, acceptability, and initial efficacy of a pre-exposure prophylaxis knowledge mobile app for young men who have sex with men (YMSM). In: APHA. 2017 Presented at: American Public Health Association Annual Meeting; Nov 7; Atlanta, GA URL: https://apha.confex.com/apha/2017/meetingapp.cgi/Paper/387585
30. Johnson AK, Fletcher FE, Ott E, Wishart M, Friedman EE, Terlikowski J, et al. Awareness and intent to use pre-exposure prophylaxis (PrEP) among African American women in a family planning clinic. J Racial Ethn Health Disparities 2020 Jun 17;7(3):550-554. [doi: 10.1007/s40615-019-00683-9] [Medline: 31848943]
31. Mustanski B, Garofalo R, Monahan C, Gratzer B, Andrews R. Feasibility, acceptability, and preliminary efficacy of an online HIV prevention program for diverse young men who have sex with men: the keep it up! intervention. AIDS Behav 2013 Nov 15;17(9):2999-3012 [FREE Full text] [doi: 10.1007/s10461-013-0507-z] [Medline: 23673793]
32. Muessig KE, Bien CH, Wei C, Lo EJ, Yang M, Tucker JD, et al. A mixed-methods study on the acceptability of using eHealth for HIV prevention and sexual health care among men who have sex with men in China. J Med Internet Res 2015 Apr 21;17(4):e100 [FREE Full text] [doi: 10.2196/jmir.3370] [Medline: 25900881]
33. Ramanathan N, Swendeman D, Comulada WS, Estrin D, Rotheram-Borus MJ. Identifying preferences for mobile health applications for self-monitoring and self-management: focus group findings from HIV-positive persons and young mothers. Int J Med Inform 2013 Apr;82(4):e38-e46. [doi: 10.1016/j.ijmedinf.2012.05.009] [Medline: 22704234]

34. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251 [FREE Full text] [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
35. Beymer MR, Holloway IW, Pulsipher C, Landovitz RJ. Current and future PrEP medications and modalities: on-demand, injectables, and topicals. Curr HIV/AIDS Rep 2019 Aug 20;16(4):349-358 [FREE Full text] [doi: 10.1007/s11904-019-00450-9] [Medline: 31222499]
36. Lam FC, Longnecker MT. A modified Wilcoxon rank sum test for paired data. Biometrika 1983;70(2):510-513. [doi: 10.1093/biomet/70.2.510]
37. Gibbs G. Thematic coding and categorizing. In: Analyzing qualitative data. Los Angeles, California: Sage; 2007:38-56.
38. Cordova D, Alers-Rojas F, Lua FM, Bauermeister J, Nurenberg R, Ovadje L, et al. The usability and acceptability of an adolescent mHealth HIV/STI and drug abuse preventive intervention in primary care. Behav Med 2018 Jul 15;44(1):36-47 [FREE Full text] [doi: 10.1080/08964289.2016.1189396] [Medline: 27223646]
39. Chávez NR, Shearer LS, Rosenthal SL. Use of digital media technology for primary prevention of STIs/HIV in youth. J Pediatr Adolesc Gynecol 2014 Oct;27(5):244-257. [doi: 10.1016/j.jpag.2013.07.008] [Medline: 24332613]
40. LaBelle M, Strong C, Tseng YC. mHealth strategies to promote uptake and adherence to PrEP: a systematic review. 2020 Presented at: International Conference on Human-Computer Interaction; 19-24 July; Copenhagen, Denmark. [doi: 10.1007/978-3-030-49913-6_9]
41. Sabben G, Mudhune V, Ondeng'e K, Odero I, Ndivo R, Akelo V, et al. A smartphone game to prevent HIV among young Africans (tumaini): assessing intervention and study acceptability among adolescents and their parents in a randomized controlled trial. JMIR Mhealth Uhealth 2019 May 21;7(5):e13049 [FREE Full text] [doi: 10.2196/13049] [Medline: 31115348]

Abbreviations

CAB: community advisory board
PID: participant ID
PrEP: pre-exposure prophylaxis
STI: sexually transmitted infection

Edited by A Mavragani; submitted 04.03.22; peer-reviewed by C Biernesser, S Lindsay; comments to author 19.04.22; revised version received 10.05.22; accepted 25.05.22; published 28.07.22.

Please cite as:
Johnson AK, Haider S, Nikolajuk K, Kuhns LM, Ott E, Motley D, Hill B, Hirschhorn L
An mHealth Intervention to Improve Pre-Exposure Prophylaxis Knowledge Among Young Black Women in Family Planning Clinics: Development and Usability Study
JMIR Form Res 2022;6(7):e37738
URL: https://formative.jmir.org/2022/7/e37738
doi:10.2196/37738
PMID:35900830

Original Paper

Feature-Level Analysis of a Smoking Cessation Smartphone App Based on a Positive Psychology Approach: Prospective Observational Study

Bettina B Hoepper[1], MS, PhD; Kaitlyn R Siegel[1], BA; Hannah A Carlon[2], BS; Christopher W Kahler[3], PhD; Elyse R Park[4], MPH, PhD; Steven Trevor Taylor[1], MPH, MD; Hazel V Simpson[1], BA; Susanne S Hoeppner[5], MApStat, PhD

[1]Recovery Research Institute, Department of Psychiatry, Massachusetts General Hospital, Boston, MA, United States

[2]Department of Psychology, University of New Mexico, Albuquerque, NM, United States

[3]Center for Alcohol and Addiction Studies, Department of Behavioral and Social Sciences, Brown University School of Public Health, Providence, RI, United States

[4]Mongan Institute, Department of Psychiatry, Massachusetts General Hospital, Boston, MA, United States

[5]Obsessive-Compulsive Disorder and Related Disorders Program, Department of Psychiatry, Massachusetts General Hospital, Boston, MA, United States

Corresponding Author:
Bettina B Hoepper, MS, PhD
Recovery Research Institute
Department of Psychiatry
Massachusetts General Hospital
6th Floor
151 Merrimac St
Boston, MA, 02114
United States
Phone: 1 617 643 1988
Fax: 1 617 643 7667
Email: bhoeppner@mgh.harvard.edu

Abstract

Background: Smoking cessation smartphone apps have emerged as highly accessible tools to support smoking cessation efforts. It is unknown how specific app features contribute to user engagement over time and relate to smoking outcomes.

Objective: To provide a feature-level analysis of the Smiling Instead of Smoking app (version 2) and to link feature use to subsequent smoking cessation.

Methods: Nondaily smokers (N=100) used the app for a period of 49 days (1 week before quitting and 6 weeks after quitting). Participants self-reported 30-day point-prevalence abstinence at the end of this period and at a 6-month follow up (the survey response rate was 94% and 89% at these points, respectively). Self-reported 30-day point prevalence abstinence rates were 40% at the end of treatment and 56% at the 6-month follow up. The app engaged users in both positive psychology content and traditional behavioral smoking cessation content. The app sent push notifications to prompt participants to complete prescribed content (ie, a "happiness exercise" every day and a "behavioral challenge" to use the app's smoking cessation tools on 15 out of 49 days). Actions that participants took within the app were timestamped and recorded.

Results: Participants used the app on 24.7 (SD 13.8) days out of the 49 prescribed days, interacting with the happiness content on more days than the smoking content (23.8, SD 13.8 days vs 17.8, SD 10.3 days; t_{99}=9.28 [2-tailed]; *P*<.001). The prescribed content was frequently completed (45% of happiness exercises; 57% of behavioral challenges) and ad libitum tools were used on ≤7 days. Most participants used each ad libitum smoking cessation tool at least once, with higher use of personalized content (≥92% used "strategies," "cigarette log," "smoke alarms," and "personal reasons") than purely didactic content (79% viewed "benefits of quitting smoking"). The number of days participants used the app significantly predicted 30-day point-prevalence abstinence at the end of treatment (odds ratio [OR] 1.05, 95% CI 1.02-1.09; *P*=.002) and at the 6-month follow up (OR 1.04, 95% CI 1.008-1.07; *P*=.01). The number of days participants engaged with the happiness content significantly predicted smoking abstinence at the end of treatment (OR 1.05, 95% CI 1.02-1.08; *P*=.002) and at the 6-month follow up (OR 1.04, 95% CI 1.007-1.07;

P=.02). This effect was not significant for the number of days participants engaged with the smoking cessation content of the app, either at the end of treatment (OR 1.04, 95% CI 0.996-1.08, *P*=.08) or at the 6-month follow up (OR 1.02, 95% CI 0.98-1.06; *P*=.29).

Conclusions: Greater app usage predicted greater odds of self-reported 30-day point-prevalence abstinence at both the end of treatment and over the long term, suggesting that the app had a therapeutic benefit. Positive psychology content and prescriptive clarity may promote sustained app engagement over time.

Trial Registration: ClinicalTrials.gov NCT03951766; https://clinicaltrials.gov/ct2/show/NCT03951766

(JMIR Form Res 2022;6(7):e38234) doi:10.2196/38234

KEYWORDS

mHealth; smartphone; smartphone app; smoking; smoking cessation; nondaily smoking; positive psychology; happiness; positive affect; clinical trial; feasibility; acceptability; app usage; mobile health

Introduction

Mobile technologies have recently emerged as highly accessible support tools for health behavior change and mental health promotion. This has been particularly true for smoking cessation, with notable increases in the use of [1,2] and referral to [3] mobile technologies designed to support quitting. User engagement with smartphone apps, however, represents a critical challenge. It has been estimated that less than 5% of apps continue to be used 15 to 30 days after the initial app download [4]. Note that to our knowledge, a standardized operational definition of app engagement has not yet been established [5]; throughout this paper, we use the term "app engagement" to denote user behavior, that is, an app user interacting with the app's user interface.

App engagement is important for several reasons. First, it is through interaction with an app that app users engage in therapeutic activities. Such engagement can take different forms. An app may present information, provide a tool or behavior change strategy, assign homework, or prompt the app user to take specific actions (eg, call a clinician) [1]. It can be assumed that the more app users interact with the app, the more they will engage in thoughts, feelings, and actions that are believed to be beneficial to their smoking cessation goal. This logic is consistent with findings across the eHealth literature demonstrating that greater engagement with eHealth tools (eg, websites) is associated with more favorable outcomes related to smoking cessation [6-8].

A critically useful feature of smartphone apps is their potential utility in providing sustained support over time. Several studies suggest that extending behavioral support helps smokers remain abstinent in the long term, with longer treatments lasting 8 to 12 weeks [9-11]. If longer treatment is better, the next question is what app users should be doing within a smoking cessation app. The vast majority of publicly available smoking cessation apps focus on simple tools: calculators to track money saved and health benefits accrued or calendars to track the days until or since the chosen quit day [1]. The apps provided on Smokefree.gov, a recommended mHealth referral site for treating smokers in health care settings [3], are more sophisticated. They offer a variety of tools and trackers (eg, time-based and GPS-based reminders to abstain from smoking at high-risk times or in high-risk locations and daily tips and milestone achievement badges) while also providing tailored feedback for overcoming urges to smoke due to cravings and mood states. These apps are able to capture the initial engagement of a very large number of smokers (there are over 25,000 and 13,000 users to date for QuitGuide and quitSTART, respectively) [12]. It is unclear, however, to what extent and how these apps sustain engagement over time.

A new generation of smoking cessation apps is emerging. In such apps, there are additional ways that smoking cessation is supported. The nature of these approaches varies widely from app to app, including contingency management [13,14], motion-sensor detection of smoking [15], carbon monoxide level monitoring [16], using gaming to engage smokers in skills practice [17] or other activities promoting smoking cessation [18], promoting nicotine replacement therapy adherence [19], prescribing bouts of physical activity [20], mindfulness training [21], and acceptance and commitment therapy (ACT) [22,23]. Most of these apps are in the early stages of development, and study protocols have been published, but the studies are ongoing. A handful of pilot feasibility studies exist, only a few of which offer insight into app usage over time. App usage ranged from 7 days for a smartphone app using a gamification approach to smoking cessation [18], to 34 and 32 days, respectively, in a pilot randomized control trial that compared an ACT-based app to the National Cancer Institute (NCI)'s app QuitGuide [22]. The largest randomized trial to date reported 24 days of app use for the app iCanQuit, an ACT-based smoking cessation app, compared to 7 days of use for an earlier version of NCI's QuitGuide, a version that was largely text-based and did not include trackers or tools [23].

In this paper, we present a feature-level analysis of app usage over time of the Smiling Instead of Smoking (SiS) app. To our knowledge, this is only the second research project that links feature-level app use to subsequent smoking cessation. The first such project provided a feature-level analysis of the app SmartQuit [24], an early version of iCanQuit [23]. The feature-level analysis of SmartQuit [24] indicated that the use of 3 features was prospectively linked to smoking cessation at follow up; 2 of these features were ACT focused (ie, tracking ACT skills practice and tracking the practice of letting urges pass); the remaining feature was a traditional US Clinical Practice Guidelines (USCPG) feature (ie, viewing the quit plan) [24]. In that analysis, the team noted little overlap between the

popularity of the app's features and their subsequent link to smoking cessation success.

The app we are examining in this paper, the SiS app, was developed specifically for nondaily smokers [25-27]. Nondaily smoking is a widespread, increasingly prevalent pattern of smoking. Currently, 24.3% of all adult smokers smoke on a nondaily basis [28], which constitutes a 27% increase over the past decade [29]. Despite such prevalence, particularly in ethnic minority groups [30-34] and vulnerable populations, such as persons with mental health and substance use challenges [35], behavioral and pharmacological recommendations for nondaily smoking remain unaddressed in clinical practice guidelines [36]. Nicotine replacement therapies have been tried but so far have failed to show efficacy in achieving smoking abstinence in nondaily smokers [37,38], in line with a lack of interest among nondaily smokers in pharmacotherapy for smoking cessation [39,40]. Nondaily smokers are, however, highly motivated to quit smoking. Compared to daily smokers, they have greater current intentions to quit smoking [31,41,42] and more recent and planned cessation efforts [42-45]. These factors point to the utility of behavioral support, which can be delivered effectively via smartphone technology.

The therapeutic goal of the SiS app is to maintain positive affect while smokers undergo a quit attempt. Positive affect often decreases during a quit attempt [46], as smokers struggle with cravings and adjusting to a smoke-free life. Indeed, recently, a decrease in positive affect has been suggested as a new symptom of tobacco withdrawal, based on data from 24 trials involving 2054 participants showing a medium effect size (Cohen d=–0.40) for an overall decrease in positive affect [47]. Maintaining positive affect during this time, however, may be especially beneficial, because greater positive affect is associated with increased self-efficacy to quit smoking [48], decreased desire to smoke [49,50], and greater readiness to process self-relevant health information [51], all of which are constructs highlighted in dominant health behavior theories as causal agents in successful behavioral change [52-56].

Smoking cessation apps can deliver a wide array of content, but users must engage with the information to benefit. The SiS app engages users in short, daily exercises to boost their positive affect. This boost in positive affect is intended to increase their readiness to engage in the smoking cessation materials provided in the app [51]. Moreover, completing happiness-boosting exercises is intrinsically rewarding, in and of itself. Having rewarding experiences while using the app may, in turn, entice app users to return to the app on subsequent days. This dynamic is evident across a diversity of settings, and positive psychology interventions have been found to be highly appealing to patients [57], resulting in better treatment adherence [58,59] and engagement [60].

In the initial study of version 1 of the SiS app (N=30), the "happiness exercises" appeared to be a driving factor in app engagement [26]. In this paper, we are examining app usage of participants in a larger trial, using version 2 of this app. The overall app usage was quite high, with smokers using the SiS app for an average of 24 days within the prescribed period of 49 days [27]. Using these data, our goals were to (1) describe how nondaily smokers used the app and (2) test if app usage patterns (ie, any app use, use of happiness content, and use of smoking content) during the prescribed treatment period predicted smoking abstinence (ie, self-reported 30-day point-prevalence abstinence [PPA]) at the end of treatment and at a 6-month postquit follow up.

Methods

Participants

In this secondary data analysis, we examined app usage data from 100 adult nondaily smokers participating in a single-arm study. The eligibility criteria for the study were as follows: age at least 18 years, nondaily smoking habit (ie, smoking at least weekly but no more than 25 out of the past 30 days), smartphone ownership (Android or iPhone only), willingness to make a quit attempt as part of the study, willingness to name friends and family members who could help study staff with updating contact information for follow-up assessments, and fluency in the English language. Note that our operational definition of nondaily smokers was designed to include the majority of nondaily smokers while targeting nondaily smokers who would routinely engage in nondaily smoking and thus might benefit from an app providing continuous support, and nondaily smokers who were not too close to being daily smokers. We have kept this operational definition consistent across our studies on nondaily smokers [26,40]. Nationally representative data at the time suggested that it would include 72% of nondaily smokers [31]. Participants who completed a baseline survey and who were successfully onboarded to the app via phone were included in this secondary data analysis. The average age of participants was 35.9 (SD 11.4) years. More than half (61/100, 61%) were female, and the majority were white (75/100, 75%) and employed (63/100, 63%), either full-time (44/100, 44%) or part-time (19/100, 19%). Thirty-eight participants (38/100, 38%) had a college degree or higher. Most had previously smoked daily (70/100, 70%). Typically, approximately half of nondaily smokers are former daily smokers [61-63]. Most participants had previously tried to quit smoking (77/100, 77%), a rate that was slightly higher than that reported by nondaily smokers in the 2000 National Health Interview Survey, in which 65% reported having tried to quit [42]. Participants smoked on average on 14.7 (SD 4.6) days out of the past 30 days and smoked 4.6 (SD 3.3) cigarettes per smoking day. Half had tried e-cigarettes (57/100, 57%).

Procedures

Recruitment occurred between June and November 2019. Participants were recruited nationwide using online resources. After screening, participants completed an onboarding phone call with study staff, during which they were guided through downloading, installing, and using the app. This onboarding phone call marked the beginning of the treatment period: participants were instructed to use the app for the subsequent 49 days. The onboarding call was scheduled to occur 1 week prior to the participants' chosen quit day, so that the prescribed app use period covered 1 week before and 6 weeks after the initial quit day (participants could reschedule their quit day within the app). App usage for all participants was recorded by

the app. Participants also completed online surveys 2, 6, 12, and 24 weeks after the initial quit day, with response rates of 96/100 (96%), 96/100 (96%), 94/100 (94%), and 89/100 (89%), respectively.

Ethics Approval

The study procedures were approved by the Mass General Brigham Institutional Review Board (2018P002699) and are detailed elsewhere [27]. The trial has been registered on ClinicalTrials.gov (NCT03951766).

Outcomes

App Utilization

Actions that participants took within the app were timestamped by the app and recorded on a secure server. From these data, we coded the number of days participants used certain features of the app and the percentage of participants who used that feature at least once after onboarding.

Smoking Cessation

In online surveys, participants were asked to indicate their smoking status using the following options: "I smoke daily," "I smoke nondaily (and have smoked in the past 7 days)," "I smoke nondaily (but have NOT smoked in the past 7 days)," and "I do not smoke at all." Participants who reported not smoking at all were then asked if they had been completely abstinent since their originally chosen quit day, during the past 7 days, and during the past 30 days. From this, we coded 30-day PPA at the end of treatment and at the 6-month follow up.

Description of the SiS App

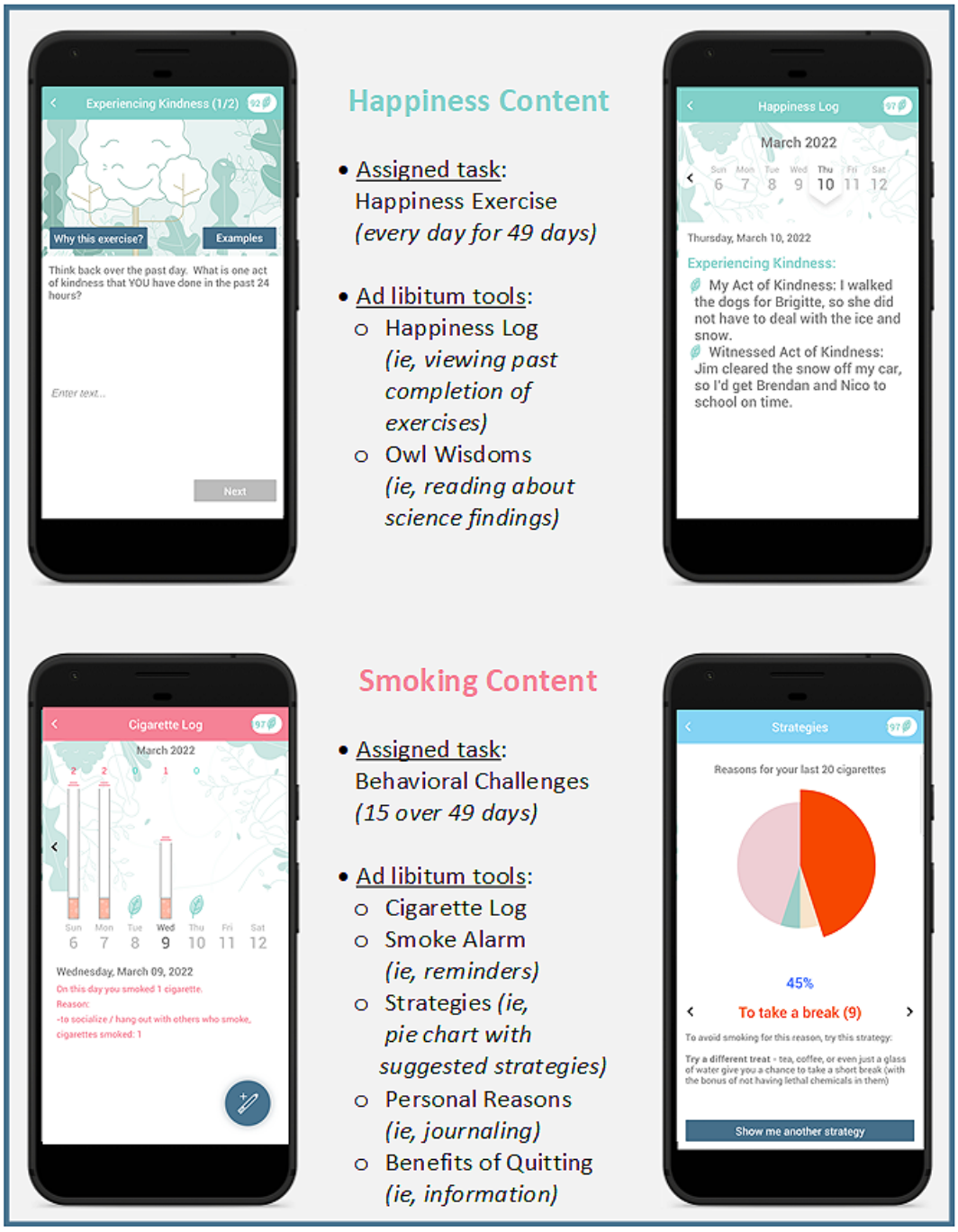
Version 2 of the SiS app (Figure 1) engaged app users in both positive psychology content designed to maintain their positive affect and traditional behavioral smoking cessation content to guide their quit attempt. During onboarding, study staff walked participants through the app and how to use it. They started with the happiness content, including showing the participants the specific buttons that explained the positive psychology framework used by the app (ie, the buttons labeled "why happiness" and "why this exercise"). These buttons were prominently displayed when engaging in the positive psychology content of the app and provided text explaining why app users were asked to complete happiness exercises in order to support their smoking cessation efforts. Study staff then moved on to "behavioral challenges," and used these as an organizing structure to guide participants through the smoking tools.

To elicit positive affect, participants were asked to complete a happiness exercise each day. Each day, the app chose 1 of 5 happiness exercises (Multimedia Appendix 1) to be completed that day. To complete the exercise, participants had to enter text into the app (eg, to describe good things that had happened to them or to describe something they had savored). These 5 exercises had been tested previously in an online survey that randomized survey takers into completing 1 of these exercises or 1 of 2 control exercises, which showed that these happiness exercises increased in-the-moment happiness [64]. Optionally, app users could review their past entries in the "happiness log," and could use the feature called "owl wisdoms" to read about scientific findings that showcase the utility of engaging in happiness-enhancing activities.

For smoking content, participants were asked to complete temporally appropriate "behavioral challenges" every 3 to 4 days (on 15 of 49 days). These behavioral challenges were anchored on the participant's quit day, which they specified in the app upon app installation. Participants could reset the quit day at any point, causing the app to adjust the schedule of the behavioral challenges accordingly. These behavioral challenges prompted users to use the smoking cessation tools provided within the app, in the order recommended by the NCI's "Clearing the Air" brochure [65]. The tools included a cigarette log to log smoked cigarettes, a strategy guide, which provided a pie chart of users' smoking triggers and suggested strategies for them, an alarm feature that let users set reminders to stay smoke free at upcoming times and events, a journal function to enter personal reasons for quitting smoking, and an informational section where the benefits of quitting smoking were presented. After the first month, behavioral challenges also directed participants to use the app's ad libitum happiness tools (ie, the happiness log and "owl wisdoms").

In total, app use entailed both prescribed (ie, happiness exercises and behavioral challenges) and ad libitum app activities that pertained to either happiness or smoking cessation (Figure 1). During the onboarding call, study staff set clear expectations that the participants should complete the happiness exercise every day for 49 days (ie, during the treatment period) and optionally thereafter and that they should complete every behavioral challenge. The app sent push notifications to prompt participants to complete the prescribed content. For the happiness exercise, the push notification was either sent at 10 AM to announce the exercise, at a random time between 12 PM and 2 PM to remind them to complete the exercise, or at 7 PM (if the exercise was still incomplete). For behavioral challenges, the push notification was always sent at 10 AM.

Figure 1. Features of the Smiling instead of Smoking 2 app.



Analyses

To describe app use during the prescribed 49-day app use period, we calculated the average number of days on which participants used the app overall and specific functions within the app. We also calculated the percentage of participants who used specific functions of the app at least once after onboarding and the percentage of times the prescribed activities (ie, happiness exercises and behavioral challenges) were completed. To test if participants seemed to prefer one happiness exercise over another, we calculated the number of times the participants completed each of the 5 exercises and then used a hierarchical linear model to test if the categorical variable denoting each exercise significantly predicted this number. Observations were modeled as nested within persons.

To test if app usage during the prescribed treatment period predicted smoking abstinence, we used a series of univariate logistic regressions, where self-reported 30-day PPA was the dependent variable (with 1 indicating “abstinent” and 0 indicating “not abstinent”) and app usage was the univariate predictor. We examined 3 different summaries of app usage: the overall number of days the app was used, the number of days the happiness content was used, and the number of days the smoking cessation content was used. Based on the app usage pattern observed in the first SiS study [26], we expected the correlation of the number of days the app was used and the number of days the happiness content was engaged with to be very high. We calculated them separately, however, to create conceptual clarity in our prediction of smoking cessation. We

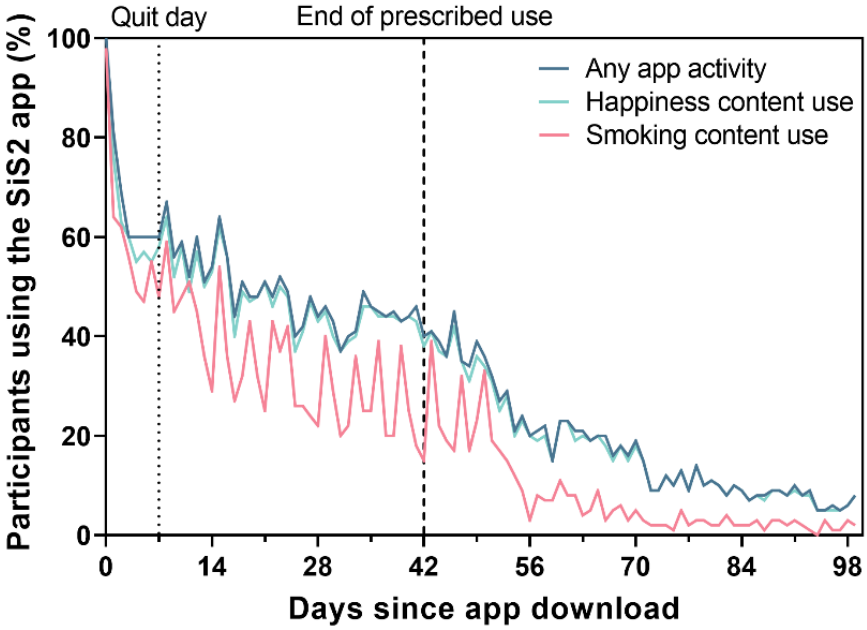
fit the same models for smoking abstinence at the end of treatment (ie, 6 weeks after the initially chosen quit day) and at the end of follow up (ie, 6 months after the initially chosen quit day). Participants were assumed to be smoking if they did not complete the surveys (there were 4% and 11% nonresponse rates at weeks 6 and 24, respectively). The logistic regression results are presented with odds ratios (ORs) and the Wald 95% CI, as well as the *C* statistic (an indicator of correct classification). All analyses were completed in SAS 9.4 for Windows (SAS Institute).

Results

App Usage

Participants used the SiS app an average of 24.1 (SD 14.1) days out of the 49 prescribed days (Figure 2). Overall, they interacted with the happiness content on more days than the smoking-related content (23.2 days, SD 14.1, vs 16.7 days, SD 10.3; t_{99}=9.47 [2-tailed]; *P*<.001). Participants completed the behavioral challenges more consistently than the happiness exercises, with participants completing 56.6% (SD 28.1%) of the behavioral challenges, on average, compared to 44.8% (SD 28.8%) of the happiness exercises (t_{99}=7.44; *P*<.001). The completion rate of the happiness exercises differed by exercise type ($F_{4,396}$=2.82; *P*=.03). Tukey adjusted posthoc pairwise comparisons showed that participants completed the "rose, thorn, and bud" exercise more often than the "savoring" exercise (4.6 days, SD 3.1, vs 4.1 days, SD 2.8; *P*=.02). The completion rates of the other 3 exercise types ("3 good things": 4.5 days, SD 3.0, "experiencing kindness": 4.4 days, SD 3.1, and "reliving happy moments": 4.4 days, SD 3.0) were intermediate to the "rose, thorn, and bud" and "savoring" exercises and did not differ from any other exercise type.

Figure 2. App usage over time. SiS2: Smiling Instead of Smoking, version 2.



Ad libitum tools were used relatively sparingly. The behavioral challenges appeared to have been successful in initially engaging participants with specific tools, as indicated by the high percentage of participants using each tool at least once after the onboarding day (Table 1). For example, 93 participants (93%) used the "personal reasons" tool, and 93 participants (93%) used "smoke alarms" at least once after onboarding. The number of days on which participants used these smoking cessation tools, however, was relatively low. Of these tools, the "cigarette log" was used the most (average 7.2 days), and the "benefits of quitting" the least (average 2.5 days). The happiness-focused ad libitum tools were similarly infrequently used. The exception was the happiness log, which was viewed on 22.1 of 49 days; however, it should be noted that upon completion of the assigned happiness exercises, participants automatically landed on the happiness log.

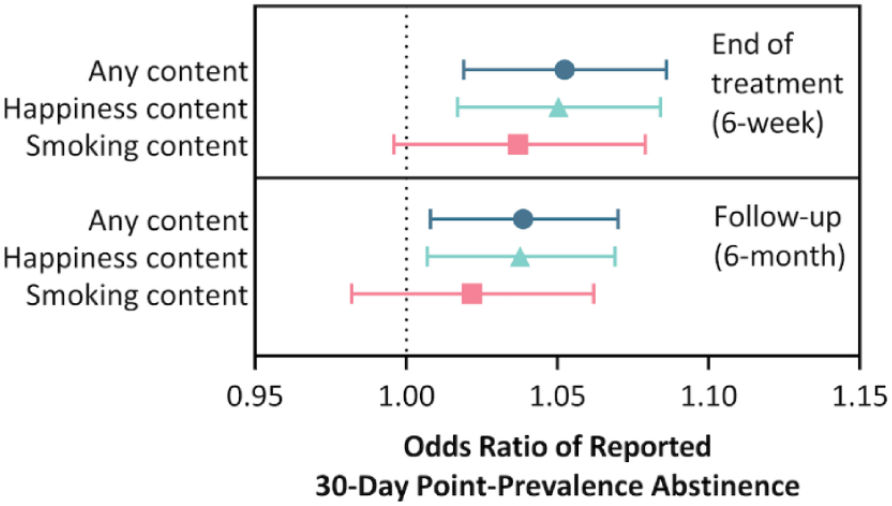
Table 1. Description of app use during the prescribed app use period (ie, 49 days).

App use	Days used of possible 49, mean (SD)	Participants (N=100) with at least one day of use after onboarding, n (%)
Overall		
Any use of the app	24.1 (14.1)	99 (99)
Any happiness content	23.2 (14.1)	98 (98)
Any smoking content	16.7 (10.3)	98 (98)
Assigned tasks		
“Happiness exercises” completed	22.0 (14.1)	96 (96)
“Behavioral challenges” completed	8.5 (4.2)	100 (100)
Ad libitum tasks		
“Happiness log” viewed	22.1 (14.1)	96 (96)
“Owl wisdoms” viewed	3.6 (3.3)	78 (78)
“Happiness information” viewed	2.0 (3.2)	60 (60)
“Cigarette log” viewed/edited	7.2 (6.5)	96 (96)
“Smoke alarms” viewed/edited	4.6 (3.0)	93 (93)
“Strategies” viewed/edited	4.1 (3.3)	94 (94)
“Personal reasons” viewed/edited	3.5 (3.2)	92 (92)
“Benefits of quitting” viewed	2.5 (2.3)	79 (79)

Relationship of App Usage to Smoking Abstinence

As illustrated in Figure 3, overall, the number of days with any use of the app significantly predicted smoking abstinence at 6 weeks (1 more day of use: OR 1.052, 95% CI 1.019-1.086; *P*=.002; *C*=.69) and 6 months postquitting (1 more day of use: OR 1.038, 95% CI 1.008-1.070; *P*=.014; *C*=.65). The number of days participants engaged with the SiS 2 app’s happiness content significantly predicted smoking abstinence at the end of treatment (1 more day of use: OR 1.050, 95% CI 1.017-1.084; *P*=.002; *C*=.69) and at 6-month follow up (1 more day of use: OR 1.037, 95% CI 1.007-1.069; *P*=.016; *C*=.65). This effect was not significant for the number of days participants engaged with the smoking cessation content of the SiS 2 app, at either the end of treatment (1 more day of use: OR 1.036, 95% CI 0.996-1.079; *P*=.08; *C*=.64) or at the 6-month follow up (1 more day of use: OR 1.021, 95% CI 0.982-1.062; *P*=.29; *C*=.59). The correlation between these 3 app usage indices was high, especially between any content and happiness content (*r*=0.995), but only somewhat lower for smoking content with any content (*r*=0.89) and happiness content (*r*=0.88).

Figure 3. Odds ratio of app usage predicting self-reported 30-day point prevalence abstinence. The odds ratio is based on a single-day increase in app usage of the indicated content (ie, "any content," "happiness content," "or smoking cessation content").



Discussion

Key Findings

This secondary data analysis of participants enrolled in a smartphone app–based smoking cessation study provided insight into how feature-level app usage behaviors relate to smoking cessation outcomes. The app under study, the SiS 2 app, is part of an emerging generation of smartphone apps that offer a treatment framework beyond standard USCPG content. In our feature-level analysis of the SiS 2 app, we found that overall greater app usage predicted higher chances of subsequent smoking cessation. This finding is in line with the feature-level analysis of the app SmartQuit [24], which showed that greater app use was positively related to subsequent smoking cessation. It suggests that there was a potential therapeutic effect of engaging with the SiS 2 app, though a causal relationship could not be established in this observational study.

Divergent from SmartQuit findings, for SiS 2 the popularity of the app's features aligned with smoking cessation success, where greater usage of the happiness components of the app predicted greater chances of 30-day PPA at both the end of treatment and at the 6-month follow up. This finding suggests that the positive psychology components of SiS 2 are an important factor in supporting smoking abstinence. This is in line with findings from in-person treatment studies that indicate the value of positive psychology in smoking cessation [66,67]. Our findings here suggest this value may extend to the smartphone app environment. Particularly noteworthy is the high level of engagement with the SiS 2 app (ie, sustained use over 49 days, with multiple uses per week), largely driven by the positive psychology content. In mHealth research, touch point frequency is an emerging area of investigation, with some apps focusing solely on reminding app users not to smoke at key timepoints [15], as simple reminders can be powerful tools in smoking cessation [68]. The high touch point frequency observed for the SiS 2 app speaks to its ability to remain present in smokers' minds as they navigate smoking cessation.

Prescriptive Clarity

Our results highlighted a rather stark difference in the completion rates of assigned versus ad libitum tasks. Happiness exercises were completed on 45% of days, and 57% of assigned behavioral challenges were completed. These completion rates are in line with the completion rates reported for a mindfulness smoking cessation app for adolescents; participants completed 13 of 22 (61%) of the assigned mindfulness modules [69]. The ad libitum tools, on the other hand, were sparsely used: they were used on only 7 days for the most popular ad libitum tool. This finding suggests that prescriptive clarity may be of critical importance in driving app usage, and therefore in achieving an app's therapeutic effect. In the SiS 2 app, there was prescriptive clarity: clear expectations were set about treatment length (ie, 49 days), and which specific actions to complete (ie, daily happiness exercises and 15 behavioral challenges). These expectations were reinforced with proactive push notifications. The iCanQuit app [70] also had prescriptive clarity: a set number of modules were required to be completed. Both apps had high app engagement over time. By contrast, the NCI app QuitGuide lacks prescriptive clarity. Many potentially useful tools are offered by QuitGuide, but it is not clear which tools to use, when, and for how long. Future research that experimentally tests whether prescribed content is more engaging would be useful to inform the development of health behavior apps.

To date, text messaging has shown greater smoking cessation benefits than smartphone apps [71], potentially in part due to prescriptive clarity. Text-messaging interventions have prescriptive clarity (ie, there is a set number of days in the program; information is provided on specific days, in proactive, succinct fashion; and actions to be taken are clearly spelled out), while many apps do not [1]. Our data show that participants are willing to complete assigned tasks much more than use ad libitum tools. Data from a randomized trial conducted in the United Kingdom show that assigning daily tasks within a smoking cessation app versus offering the same content without the specific daily tasks led to improved smoking cessation rates [72]. Combined, these findings lead us to believe that prescriptive clarity is a critically important feature in the development of smartphone apps targeting smoking cessation.

Long-term Engagement

Smartphone app technology has the potential to provide ongoing support for smoking cessation over long periods of time. To date, this potential has been largely unexplored, including in our own work. To our knowledge, few studies have examined app engagement; these studies have focused on factors contributing to initial app use [73] or have tested the value of push notifications in enhancing engagement [74]. These studies have not provided insight into the content features and app parameters that promote long-term engagement. In-person smoking cessation interventions typically provide 8 to 12 weeks of support. The SiS 2 app provided assigned tasks for 49 days and continued use of the app's ad libitum tools as needed, a treatment length in line with the support offered via the NCI's Smokefree text-messaging interventions, and similar to the UK Smoke Free app, which provides assigned daily tasks for 31 days [72]. This treatment length roughly covers the time from preparation to action according to the transtheoretical model of change and does not address maintenance [55]. In fact, originally, the SiS app spanned only 21 days [26], a length that was specifically chosen to provide support during the acute "cessation" phase [75] of the process of smoking cessation. Based on user feedback, we increased treatment length to 49 days in version 2. Our app usage data demonstrate that this increased treatment length was well tolerated, opening the door to potentially further increasing treatment length to provide support during the maintenance phase of smoking cessation. To date, however, very little research exists to guide the intervention content of smartphone apps to support sustained user engagement and long-term abstinence from smoking.

Limitations

This secondary data analysis was based on a single-arm trial, and therefore causal inferences about the observed effects cannot be drawn. Our analyses were exploratory. They hint at the value of positive psychology to engage app users in smoking cessation over time and the value of prescriptive clarity, but we did not design this study, nor the SiS 2 app, to address these questions.

In considering these effects, it should be kept in mind that the participants were asked to complete happiness content on a more frequent basis than smoking cessation content (ie, daily vs every 3 to 4 days) and that the smoking cessation tools had a shelf-life (eg, "smoke alarms" became less useful as cravings diminished; participants could log cigarettes, but not smoke-free days). In terms of generalizability, it should be noted that the SiS 2 study used an interactive onboarding procedure via phone. While this is in line with warm handoff models for smoking cessation [76,77], app usage patterns within the context of a clinical trial are typically higher than real-life app use [78].

Conclusions

In the SiS 2 app, greater app usage predicted greater chances of self-reporting 30-day PPA at both the end of treatment and at a 6-month follow up. This finding strengthens the rationale for testing this app in a randomized trial. Feature-level analysis of app usage patterns suggests that positive psychology content and prescriptive clarity may promote app engagement.

Acknowledgments

This research was supported by a grant to BBH from the American Cancer Society (RSG CPPB–130323).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Happiness Activities.
[DOCX File , 2436 KB - formative_v6i7e38234_app1.docx]

References

1. Hoeppner BB, Hoeppner SS, Seaboyer L, Schick MR, Wu GWY, Bergman BG, et al. How Smart are Smartphone Apps for Smoking Cessation? A Content Analysis. Nicotine Tob Res 2016 May 04;18(5):1025-1031. [doi: 10.1093/ntr/ntv117] [Medline: 26045249]
2. Abroms LC, Lee Westmaas J, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736. [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
3. Fiore MC, Baker TB. Clinical practice. Treating smokers in the health care setting. N Engl J Med 2011 Sep 29;365(13):1222-1231. [doi: 10.1056/NEJMcp1101512] [Medline: 21991895]
4. 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]
5. O'Brien HL, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. J Am Soc Inf Sci 2008 Apr;59(6):938-955. [doi: 10.1002/asi.20801]
6. Zeng EY, Heffner JL, Copeland WK, Mull KE, Bricker JB. Get with the program: Adherence to a smartphone app for smoking cessation. Addict Behav 2016 Dec;63:120-124. [doi: 10.1016/j.addbeh.2016.07.007] [Medline: 27454354]
7. Bricker JB, Mull KE, Kientz JA, Vilardaga R, Mercer LD, Akioka KJ, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. Drug Alcohol Depend 2014 Oct 01;143:87-94. [doi: 10.1016/j.drugalcdep.2014.07.006] [Medline: 25085225]
8. Taylor G, Dalili M, Semwal M, Civljak M, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2017 Sep 04;9:CD007078. [doi: 10.1002/14651858.CD007078.pub5] [Medline: 28869775]
9. Killen J, Fortmann S, Schatzberg A, Arredondo C, Murphy G, Hayward C, et al. Extended cognitive behavior therapy for cigarette smoking cessation. Addiction 2008 Aug;103(8):1381-1390. [doi: 10.1111/j.1360-0443.2008.02273.x] [Medline: 18855829]
10. Hall SM, Humfleet GL, Reus VI, Muñoz RF, Cullen J. Extended nortriptyline and psychological treatment for cigarette smoking. Am J Psychiatry 2004 Nov 01;161(11):2100-2107. [doi: 10.1176/appi.ajp.161.11.2100] [Medline: 15514412]
11. Keith A, Dong Y, Shuter J, Himelhoch S. Behavioral Interventions for Tobacco Use in HIV-Infected Smokers: A Meta-Analysis. J Acquir Immune Defic Syndr 2016 Aug 15;72(5):527-533. [doi: 10.1097/QAI.0000000000001007] [Medline: 27028502]
12. Prutzman YM, Wiseman KP, Grady MA, Budenz A, Grenen EG, Vercammen LK, et al. Using Digital Technologies to Reach Tobacco Users Who Want to Quit: Evidence From the National Cancer Institute's Smokefree.gov Initiative. Am J Prev Med 2021 Mar;60(3 Suppl 2):S172-S184 [FREE Full text] [doi: 10.1016/j.amepre.2020.08.008] [Medline: 33663705]
13. Hertzberg JS, Carpenter VL, Kirby AC, Calhoun PS, Moore SD, Dennis MF, et al. Mobile contingency management as an adjunctive smoking cessation treatment for smokers with posttraumatic stress disorder. Nicotine Tob Res 2013 Nov 03;15(11):1934-1938. [doi: 10.1093/ntr/ntt060] [Medline: 23645606]

14. Hicks TA, Thomas SP, Wilson SM, Calhoun PS, Kuhn ER, Beckham JC. A Preliminary Investigation of a Relapse Prevention Mobile Application to Maintain Smoking Abstinence Among Individuals With Posttraumatic Stress Disorder. J Dual Diagn 2017 Dec 05;13(1):15-20. [doi: 10.1080/15504263.2016.1267828] [Medline: 27918881]
15. 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]
16. Krishnan N, Elf J, Chon S, Golub J. COach2Quit: A Pilot Randomized Controlled Trial of a Personal Carbon Monoxide Monitor for Smoking Cessation. Nicotine Tob Res 2019 Oct 26;21(11):1573-1577. [doi: 10.1093/ntr/nty182] [Medline: 30169740]
17. Krebs P, Burkhalter J, Fiske J, Snow H, Schofield E, Iocolano M, et al. The QuitIT Coping Skills Game for Promoting Tobacco Cessation Among Smokers Diagnosed With Cancer: Pilot Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Jan 10;7(1):e10071 [FREE Full text] [doi: 10.2196/10071] [Medline: 30632971]
18. Marin-Gomez FX, Garcia-Moreno Marchán R, Mayos-Fernandez A, Flores-Mateo G, Granado-Font E, Barrera Uriarte ML, et al. Exploring Efficacy of a Serious Game (Tobbstop) for Smoking Cessation During Pregnancy: Randomized Controlled Trial. JMIR Serious Games 2019 Mar 27;7(1):e12835 [FREE Full text] [doi: 10.2196/12835] [Medline: 30916655]
19. Herbec A, Brown J, Shahab L, West R, Raupach T. Pragmatic randomised trial of a smartphone app (NRT2Quit) to improve effectiveness of nicotine replacement therapy in a quit attempt by improving medication adherence: results of a prematurely terminated study. Trials 2019 Sep 02;20(1):547 [FREE Full text] [doi: 10.1186/s13063-019-3645-4] [Medline: 31477166]
20. Hassandra M, Lintunen T, Hagger MS, Heikkinen R, Vanhala M, Kettunen T. An mHealth App for Supporting Quitters to Manage Cigarette Cravings With Short Bouts of Physical Activity: A Randomized Pilot Feasibility and Acceptability Study. JMIR Mhealth Uhealth 2017 May 26;5(5):e74 [FREE Full text] [doi: 10.2196/mhealth.6252] [Medline: 28550004]
21. Garrison K, Pal P, O'Malley SS, Pittman B, Gueorguieva R, Rojiani R, et al. Craving to Quit: A Randomized Controlled Trial of Smartphone App-Based Mindfulness Training for Smoking Cessation. Nicotine Tob Res 2020 Mar 16;22(3):324-331. [doi: 10.1093/ntr/nty126] [Medline: 29917096]
22. Vilardaga R, Rizo J, Palenski P, Mannelli P, Oliver J, Mcclernon F. Pilot Randomized Controlled Trial of a Novel Smoking Cessation App Designed for Individuals With Co-Occurring Tobacco Use Disorder and Serious Mental Illness. Nicotine Tob Res 2020 Aug 24;22(9):1533-1542. [doi: 10.1093/ntr/ntz202] [Medline: 31667501]
23. Bricker JB, Watson NL, Mull KE, Sullivan BM, Heffner JL. Efficacy of Smartphone Applications for Smoking Cessation: A Randomized Clinical Trial. JAMA Intern Med 2020 Nov 01;180(11):1472-1480. [doi: 10.1001/jamainternmed.2020.4055] [Medline: 32955554]
24. Heffner JL, Vilardaga R, Mercer LD, Kientz JA, Bricker JB. Feature-level analysis of a novel smartphone application for smoking cessation. Am J Drug Alcohol Abuse 2015 Jan 14;41(1):68-73. [doi: 10.3109/00952990.2014.977486] [Medline: 25397860]
25. Hoeppner BB, Hoeppner SS, Kelly L, Schick M, Kelly JF. Smiling Instead of Smoking: Development of a Positive Psychology Smoking Cessation Smartphone App for Non-daily Smokers. Int J Behav Med 2017 Oct 14;24(5):683-693. [doi: 10.1007/s12529-017-9640-9] [Medline: 28197846]
26. Hoeppner BB, Hoeppner SS, Carlon HA, Perez GK, Helmuth E, Kahler CW, et al. Leveraging Positive Psychology to Support Smoking Cessation in Nondaily Smokers Using a Smartphone App: Feasibility and Acceptability Study. JMIR Mhealth Uhealth 2019 Jul 03;7(7):e13436 [FREE Full text] [doi: 10.2196/13436] [Medline: 31271147]
27. Hoeppner BB, Siegel KR, Carlon HA, Kahler CW, Park ER, Hoeppner SS. A Smoking Cessation App for Nondaily Smokers (Version 2 of the Smiling Instead of Smoking App): Acceptability and Feasibility Study. JMIR Form Res 2021 Nov 17;5(11):e29760 [FREE Full text] [doi: 10.2196/29760] [Medline: 34787577]
28. Adhikari B, Kahende J, Malarcher A, Pechacek T, Tong V, National Center for Chronic Disease Prevention and Health Promotion. Smoking-Attributable Mortality, Years of Potential Life Lost, and Productivity Losses—United States, 2000-2004. JAMA 2009 Feb 11;301(6):593. [doi: 10.1001/jama.301.6.593]
29. Jamal A, King BA, Neff LJ, Whitmill J, Babb SD, Graffunder CM. Current Cigarette Smoking Among Adults - United States, 2005-2015. MMWR Morb Mortal Wkly Rep 2016 Nov 11;65(44):1205-1211 [FREE Full text] [doi: 10.15585/mmwr.mm6544a2] [Medline: 27832052]
30. Wortley P, Husten C, Trosclair A, Chrismon J, Pederson LL. Nondaily smokers: a descriptive analysis. Nicotine Tob Res 2003 Oct 1;5(5):755-759. [doi: 10.1080/1462220031000158753] [Medline: 14577992]
31. Hassmiller KM, Warner KE, Mendez D, Levy DT, Romano E. Nondaily smokers: who are they? Am J Public Health 2003 Aug;93(8):1321-1327. [doi: 10.2105/ajph.93.8.1321] [Medline: 12893622]
32. Husten CG, McCarty MC, Giovino GA, Chrismon JH, Zhu B. Intermittent smokers: a descriptive analysis of persons who have never smoked daily. Am J Public Health 1998 Jan;88(1):86-89. [doi: 10.2105/ajph.88.1.86] [Medline: 9584039]
33. Coggins CRE, Murrelle EL, Carchman RA, Heidbreder C. Light and intermittent cigarette smokers: a review (1989-2009). Psychopharmacology (Berl) 2009 Dec 3;207(3):343-363. [doi: 10.1007/s00213-009-1675-4] [Medline: 19830407]
34. Trinidad D, Pérez-Stable EJ, Emery S, White M, Grana R, Messer K. Intermittent and light daily smoking across racial/ethnic groups in the United States. Nicotine Tob Res 2009 Feb;11(2):203-210. [doi: 10.1093/ntr/ntn018] [Medline: 19246433]

35. Weinberger AH, Streck JM, Pacek LR, Goodwin RD. Nondaily Cigarette Smoking Is Increasing Among People With Common Mental Health and Substance Use Problems in the United States: Data From Representative Samples of US Adults, 2005-2014. J Clin Psychiatry 2018 Aug 14;79(5):1-14. [doi: 10.4088/JCP.17m11945] [Medline: 30153404]
36. Jaén C, Benowitz N, Curry S, Parsippany N, Kottke T, Mermelstein R. A clinical practice guideline for treating tobacco use and dependence: 2008 update. A U.S. Public Health Service report. Am J Prev Med 2008 Aug;35(2):158-176. [doi: 10.1016/j.amepre.2008.04.009] [Medline: 18617085]
37. Shiffman S, Scholl S, Mao J, Ferguson S, Hedeker D, Primack B, et al. Using Nicotine Gum to Assist Nondaily Smokers in Quitting: A Randomized Clinical Trial. Nicotine Tob Res 2020 Mar 16;22(3):390-397. [doi: 10.1093/ntr/ntz090] [Medline: 31125988]
38. Nollen NL, Cox LS, Mayo MS, Ellerbeck EF, Ahluwalia JS. Counseling alone or in combination with nicotine replacement therapy for treatment of black non-daily smokers: a randomized trial. Addiction 2020 Aug 06;115(8):1547-1560. [doi: 10.1111/add.14948] [Medline: 31899564]
39. Berg CJ, Sutfin EL, Mendel J, Ahluwalia JS. Use of and interest in smoking cessation strategies among daily and nondaily college student smokers. J Am Coll Health 2012 Apr;60(3):194-202. [doi: 10.1080/07448481.2011.586388] [Medline: 22420696]
40. Hoeppner B, Hoeppner S, Carlon H, Abry A, Darville A, Rohsenow D. Preparing for the Quit Day: Comparing Beliefs of Nondaily Versus Daily Young Adult Smokers as They Prepare for a Quit Attempt. Nicotine Tob Res 2021 May 24;23(6):1038-1046. [doi: 10.1093/ntr/ntaa166] [Medline: 32882037]
41. Rutten L, Augustson E, Doran K, Moser R, Hesse B. Health information seeking and media exposure among smokers: a comparison of light and intermittent tobacco users with heavy users. Nicotine Tob Res 2009 Feb;11(2):190-196. [doi: 10.1093/ntr/ntn019] [Medline: 19264865]
42. Tong EK, Ong MK, Vittinghoff E, Pérez-Stable EJ. Nondaily smokers should be asked and advised to quit. Am J Prev Med 2006 Jan;30(1):23-30. [doi: 10.1016/j.amepre.2005.08.048] [Medline: 16414420]
43. Levy D, Biener L, Rigotti N. The natural history of light smokers: a population-based cohort study. Nicotine Tob Res 2009 Feb;11(2):156-163. [doi: 10.1093/ntr/ntp011] [Medline: 19264862]
44. Haug S, Schaub MP, Venzin V, Meyer C, John U. Efficacy of a text message-based smoking cessation intervention for young people: a cluster randomized controlled trial. J Med Internet Res 2013 Aug 16;15(8):e171 [FREE Full text] [doi: 10.2196/jmir.2636] [Medline: 23956024]
45. Cooper TV, Taylor T, Murray A, DeBon MW, Vander Weg MW, Klesges RC, et al. Differences between intermittent and light daily smokers in a population of U.S. military recruits. Nicotine Tob Res 2010 May 04;12(5):465-473. [doi: 10.1093/ntr/ntq025] [Medline: 20203108]
46. Strong D, Kahler C, Leventhal A, Abrantes A, Lloyd-Richardson E, Niaura R, et al. Impact of bupropion and cognitive-behavioral treatment for depression on positive affect, negative affect, and urges to smoke during cessation treatment. Nicotine Tob Res 2009 Oct;11(10):1142-1153. [doi: 10.1093/ntr/ntp111] [Medline: 19574407]
47. Klemperer E, Hughes J, Peasley-Miklus CE, Callas PW, Cook JW, Streck JM, et al. Possible New Symptoms of Tobacco Withdrawal III: Reduced Positive Affect-A Review and Meta-analysis. Nicotine Tob Res 2021 Jan 22;23(2):259-266. [doi: 10.1093/ntr/ntaa044] [Medline: 32188995]
48. 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. [doi: 10.1037/hea0000075] [Medline: 25020151]
49. Shiffman S, Dunbar MS, Kirchner TR, Li X, Tindle HA, Anderson SJ, et al. Cue reactivity in non-daily smokers: effects on craving and on smoking behavior. Psychopharmacology (Berl) 2013 Mar 11;226(2):321-333. [doi: 10.1007/s00213-012-2909-4] [Medline: 23142992]
50. Rabois D, Haaga DA. The influence of cognitive coping and mood on smokers' self-efficacy and temptation. Addict Behav 2003 Apr;28(3):561-573. [doi: 10.1016/s0306-4603(01)00249-0] [Medline: 12628627]
51. Das E, Vonkeman C, Hartmann T. Mood as a resource in dealing with health recommendations: how mood affects information processing and acceptance of quit-smoking messages. Psychol Health 2012 Jan;27(1):116-127. [doi: 10.1080/08870446.2011.569888] [Medline: 21678163]
52. Becker MH, Maiman LA. Sociobehavioral determinants of compliance with health and medical care recommendations. Med Care 1975 Jan;13(1):10-24. [doi: 10.1097/00005650-197501000-00002] [Medline: 1089182]
53. Bandura A. Social cognitive theory of self-regulation. Organ Behav Hum Decis Process 1991 Dec;50(2):248-287. [doi: 10.1016/0749-5978(91)90022-l]
54. Ajzen I, Madden TJ. Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. J Exper Soc Psych 1986 Sep;22(5):453-474. [doi: 10.1016/0022-1031(86)90045-4]
55. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: Toward an integrative model of change. J Counsul Clin Psychol 1983;51(3):390-395. [doi: 10.1037/0022-006x.51.3.390]
56. Marlatt G, Donovan D. Relapse prevention: Maintenance strategies in the treatment of addictive behaviors. New York, NY: Guilford Press; 2005.

57. Moskowitz JT, Carrico AW, Duncan LG, Cohn MA, Cheung EO, Batchelder A, et al. Randomized controlled trial of a positive affect intervention for people newly diagnosed with HIV. J Consult Clin Psychol 2017 May;85(5):409-423. [doi: 10.1037/ccp0000188] [Medline: 28333512]
58. Huffman JC, Mastromauro CA, Boehm JK, Seabrook R, Fricchione GL, Denninger JW, et al. Development of a positive psychology intervention for patients with acute cardiovascular disease. Heart Int 2011 Sep 29;6(2):e14. [doi: 10.4081/hi.2011.e14] [Medline: 23825741]
59. Ogedegbe GO, Boutin-Foster C, Wells MT, Allegrante JP, Isen AM, Jobe JB, et al. A randomized controlled trial of positive-affect intervention and medication adherence in hypertensive African Americans. Arch Intern Med 2012 Feb 27;172(4):322-326. [doi: 10.1001/archinternmed.2011.1307] [Medline: 22269592]
60. Carrico AW, Gómez W, Siever MD, Discepola MV, Dilworth SE, Moskowitz JT. Pilot randomized controlled trial of an integrative intervention with methamphetamine-using men who have sex with men. Arch Sex Behav 2015 Oct 30;44(7):1861-1867. [doi: 10.1007/s10508-015-0505-5] [Medline: 26123068]
61. Tindle HA, Shiffman S. Smoking cessation behavior among intermittent smokers versus daily smokers. Am J Public Health 2011 Jul;101(7):e1-e3. [doi: 10.2105/AJPH.2011.300186] [Medline: 21566030]
62. Edwards S, Bondy S, Kowgier M, McDonald PW, Cohen JE. Are occasional smokers a heterogeneous group? An exploratory study. Nicotine Tob Res 2010 Dec;12(12):1195-1202. [doi: 10.1093/ntr/ntq168] [Medline: 20978108]
63. Nguyen Q, Zhu S. Intermittent smokers who used to smoke daily: a preliminary study on smoking situations. Nicotine Tob Res 2009 Feb;11(2):164-170. [doi: 10.1093/ntr/ntp012] [Medline: 19246632]
64. Hoeppner BB, Schick MR, Carlon H, Hoeppner SS. Do self-administered positive psychology exercises work in persons in recovery from problematic substance use? An online randomized survey. J Subst Abuse Treat 2019 Apr;99:16-23. [doi: 10.1016/j.jsat.2019.01.006] [Medline: 30797389]
65. Clearing the Air. National Cancer Institute. URL: https://www.cancer.gov/publications/patient-education/clearing-the-air-pdf [accessed 2022-06-08]
66. Kahler CW, Spillane NS, Day A, Clerkin EM, Parks A, Leventhal AM, et al. Positive Psychotherapy for Smoking Cessation: Treatment Development, Feasibility and Preliminary Results. J Posit Psychol 2014 Jan 01;9(1):19-29. [doi: 10.1080/17439760.2013.826716] [Medline: 24683417]
67. Kahler CW, Spillane NS, Day AM, Cioe PA, Parks A, Leventhal AM, et al. Positive Psychotherapy for Smoking Cessation: A Pilot Randomized Controlled Trial. Nicotine Tob Res 2015 Nov 02;17(11):1385-1392. [doi: 10.1093/ntr/ntv011] [Medline: 25646352]
68. Tobias R. Changing behavior by memory aids: a social psychological model of prospective memory and habit development tested with dynamic field data. Psychol Rev 2009 Apr;116(2):408-438. [doi: 10.1037/a0015512] [Medline: 19348548]
69. Pbert L, Druker S, Crawford S, Frisard C, Trivedi M, Osganian SK, et al. Feasibility of a Smartphone App with Mindfulness Training for Adolescent Smoking Cessation: Craving to Quit (C2Q)-Teen. Mindfulness (N Y) 2020 Mar 20;11(3):720-733. [doi: 10.1007/s12671-019-01273-w] [Medline: 33343761]
70. Bricker JB, Watson NL, Heffner JL, Sullivan B, Mull K, Kwon D, et al. A Smartphone App Designed to Help Cancer Patients Stop Smoking: Results From a Pilot Randomized Trial on Feasibility, Acceptability, and Effectiveness. JMIR Form Res 2020 Jan 17;4(1):e16652 [FREE Full text] [doi: 10.2196/16652] [Medline: 31951215]
71. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y, Dobson R. Mobile phone text messaging and app-based interventions for smoking cessation. Cochrane Database Syst Rev 2019 Oct 22;10:CD006611. [doi: 10.1002/14651858.CD006611.pub5] [Medline: 31638271]
72. Crane D, Ubhi HK, Brown J, West R. Relative effectiveness of a full versus reduced version of the 'Smoke Free' mobile application for smoking cessation: an exploratory randomised controlled trial. F1000Res 2018 Jan 9;7:1524 [FREE Full text] [doi: 10.12688/f1000research.16148.2] [Medline: 30728950]
73. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Feb 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]
74. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To Prompt or Not to Prompt? A Microrandomized Trial of Time-Varying Push Notifications to Increase Proximal Engagement With a Mobile Health App. JMIR Mhealth Uhealth 2018 Nov 29;6(11):e10123 [FREE Full text] [doi: 10.2196/10123] [Medline: 30497999]
75. Baker TB, Mermelstein R, Collins LM, Piper ME, Jorenby DE, Smith SS, et al. New methods for tobacco dependence treatment research. Ann Behav Med 2011 Apr 3;41(2):192-207. [doi: 10.1007/s12160-010-9252-y] [Medline: 21128037]
76. Richter KP, Faseru B, Shireman TI, Mussulman LM, Nazir N, Bush T, et al. Warm Handoff Versus Fax Referral for Linking Hospitalized Smokers to Quitlines. Am J Prev Med 2016 Oct;51(4):587-596. [doi: 10.1016/j.amepre.2016.04.006] [Medline: 27647059]
77. Mussulman LM, Faseru B, Fitzgerald S, Nazir N, Patel V, Richter KP. A randomized, controlled pilot study of warm handoff versus fax referral for hospital-initiated smoking cessation among people living with HIV/AIDS. Addict Behav 2018 Mar;78:205-208. [doi: 10.1016/j.addbeh.2017.11.035] [Medline: 29216569]

78. Baumel A, Edan S, Kane J. Is there a trial bias impacting user engagement with unguided e-mental health interventions? A systematic comparison of published reports and real-world usage of the same programs. Transl Behav Med 2019 Nov 25;9(6):1020-1033. [doi: 10.1093/tbm/ibz147] [Medline: 31689344]

Abbreviations

ACT: acceptance and commitment therapy
NCI: National Cancer Institute
OR: odds ratio
PPA: point-prevalence abstinence
SiS: Smiling instead of Smoking
USCPG: US Clinical Practice Guidelines

Edited by A Mavragani; submitted 25.03.22; peer-reviewed by J Heffner, J Bricker; comments to author 14.04.22; revised version received 27.04.22; accepted 28.04.22; published 28.07.22.

Please cite as:
Hoepper BB, Siegel KR, Carlon HA, Kahler CW, Park ER, Taylor ST, Simpson HV, Hoeppner SS
Feature-Level Analysis of a Smoking Cessation Smartphone App Based on a Positive Psychology Approach: Prospective Observational Study
JMIR Form Res 2022;6(7):e38234
URL: https://formative.jmir.org/2022/7/e38234
doi:10.2196/38234
PMID:35900835

Original Paper

Development of a Theory-Based, Culturally Appropriate Message Library for Use in Interventions to Promote COVID-19 Vaccination Among African Americans: Formative Research

Jennifer Cunningham-Erves[1], MAEd, MS, MPH, PhD; Heather M Brandt[2], PhD; Maureen Sanderson[3], PhD; Kristin Clarkson[4], MHA; Omaran Lee[4], MDiv; David Schlundt[5], PhD; Kemberlee Bonnet[5], MA; Jamaine Davis[6], PhD

[1]Department of Internal Medicine, Meharry Medical College, Nashville, TN, United States

[2]Department of Epidemiology and Cancer Control, St Jude Children's Research Hospital, Memphis, TN, United States

[3]Department of Family Medicine, Meharry Medical College, Nashville, TN, United States

[4]Nashville General Hospital, Nashville, TN, United States

[5]Department of Psychology, Vanderbilt University, Nashville, TN, United States

[6]Department of Biochemistry and Cancer Biology, Meharry Medical College, Nashville, TN, United States

Corresponding Author:
Jennifer Cunningham-Erves, MAEd, MS, MPH, PhD
Department of Internal Medicine
Meharry Medical College
1005 Dr DB Todd Jr Blvd
Nashville, TN, 37208
United States
Phone: 1 251 714 4233
Email: jerves@mmc.edu

Abstract

Background: Disparities in COVID-19 incidence, hospitalization, and mortality rates among African Americans suggest the need for targeted interventions. Use of targeted, theory-driven messages in behavioral and communication interventions could empower African Americans to engage in behaviors that prevent COVID-19.

Objective: To address this need, we performed a formative study that aimed to develop and design a culturally appropriate, theory-based library of messages targeting concerns around COVID-19 vaccines that could be used in behavioral and communication interventions for African Americans.

Methods: Message development occurred between January 2021 and February 2022. Initial messages were designed by a multidisciplinary team of researchers, community leaders, and community members. Kreuter's 5 strategies (ie, linguistic, peripheral, evidential, sociocultural, and constituent-involving strategies) were used to achieve cultural appropriateness. After forming a community-academic partnership, message development occurred in 4 phases: (1) adaptation of a message library using the literature, (2) review by 6 clinical and research experts for content validation, (3) input and review by a 6-member community advisory panel (CAP), and (4) message pretesting with African Americans via semistructured interviews in a qualitative study.

Results: Themes from the semistructured interviews among 30 African Americans were as follows: (1) community reactions to the messages, (2) community questions and information needs, (3) suggestions for additional content, and (4) suggestions to improve comprehension, relevance, and trustworthiness. Feedback from the CAP, community members, and scientific experts was used by members of the community-academic partnership to iteratively update message content to maximize cultural appropriateness. The final message library had 18 message subsets for adults and 17 message subsets for parents and caregivers of children. These subsets were placed into 3 categories: (1) vaccine development, (2) vaccine safety, and (3) vaccine effectiveness.

Conclusions: We used a 4-phase, systematic process using multiple community engagement approaches to create messages for African Americans to support interventions to improve COVID-19 vaccination rates among adults and children. The newly developed messages were deemed to be culturally appropriate according to experts and members of the African American community. Future research should evaluate the impact of these messages on COVID-19 vaccination rates among African Americans.

(JMIR Form Res 2022;6(7):e38781) doi:10.2196/38781

KEYWORDS

African American; Black American; Black; minority; ethnic; culturally sensitive; cultural sensitivity; inclusive; vulnerable; COVID-19; vaccination; vaccine; health promotion; campaign; messaging; culturally appropriate; theory; adults; children; disparity; health belief model; community engagement; public engagement; public awareness; community-based; health information; health communication; health intervention; vulnerable population; community health; patient education

Introduction

Background

Since December 2020, 3 vaccines have been approved in the United States to prevent severe disease and death caused by COVID-19. The 2-dose vaccinations developed by *Pfizer* and *Moderna* are approved for use in individuals aged 18 years and older and under emergency use authorization (EAU) for children aged 6 months to 17 years [1]. The 1-dose *Johnson & Johnson* vaccination is under EAU for adults aged 18 years and older [1]. Although highly effective [2], vaccination rates remain suboptimal, especially among populations who could benefit most. For example, African Americans comprise 12.4% of the US population but only 10.1% of those who have initiated the series and 10.3% of those who have completed the series as of March 10, 2022 [3]. Despite African Americans being almost 2 times more likely than White Americans to die from COVID-19 [4], vaccine hesitancy remains a major hindrance to reduced vaccine uptake among African Americans [5-9].

Emerging studies demonstrate that vaccine hesitancy is deeply rooted in several overlapping areas: (1) mistrust in health care, government, and research [10-13]; (2) structural racism [14]; and (3) lack of understanding of science related to vaccine-specific issues (eg, efficacy, safety, speed of development) [13,15]. Lack of information, misinformation, and disinformation further drive vaccine hesitancy [10], with social or mass media as the primary source [15]. Because effective communication is necessary to help African Americans make informed decisions about COVID-19 vaccines [16], studies have begun to explore the communication strategies necessary to increase COVID-19 vaccination [17-19]. Trusted messengers are key to COVID-19 information being well received and used [20].

The specific messages related to COVID-19 vaccination are as important as the messenger. Information sources have been developed and disseminated widely to educate communities on messages to use to educate communities on COVID-19 vaccination to increase uptake [21,22]. A few emerging studies have tested messages, including persuasive messaging [18], video-based messages [23], and behavioral nudges [24], on vaccination intention or uptake. African Americans suggest the need for messages that are accurate, targeted, culturally appropriate, and community based [25,26]. However, perceptions of the messages remain unknown, and none (to the best of our knowledge) have actively engaged the African American community to develop or refine the messages on COVID-19 vaccination to ensure cultural appropriateness. Such engagement is key because African Americans' values and decision-making about the COVID-19 vaccine are strongly shaped by culture [10,12,13,27], and targeting will maximize "fit" of information to an individual's unique characteristics [28,29].

Conceptual Framework

Message development was guided by 2 psychosocial decision-making models: (1) theory of reasoned action (TRA) [30] and (2) the Health Belief Model (HBM) [31]. These health behavior theoretical models are commonly used to understand vaccination decision-making. The TRA predicts that behavioral intentions to vaccinate against COVID-19 are based on attitudes and subjective norms. The HBM predicts that the likelihood of vaccinating against COVID-19 is based on perceived susceptibility of the individual to SARS-CoV-2, perceived severity of COVID-19, and whether perceived benefits of vaccination outweigh perceived barriers.

Kreuter et al [32] proposed 5 strategies to achieve cultural appropriateness that were used to guide message development. *Peripheral strategies* increase communication appeal through the title, fonts, colors, and images. *Evidential strategies* provide data on impact of a health issue in a certain group. *Sociocultural strategies* address health issues from the social and cultural values of a group. *Linguistic strategies* fit the program to the native language of a certain group. Lastly, *constituent-involving strategies* ensure community members' inclusion in program planning. Developing culturally appropriate, theory-based messages that can be used in communication and behavioral interventions may address concerns about COVID-19 vaccination among African Americans. Because social marketing campaigns have been effective in changing knowledge, attitudes, intentions, and behavior at the community level [33-35], they can disseminate theory-based, culturally appropriate messages and potentially increase COVID-19 vaccination acceptability and uptake. Social marketing is "the application of proven concepts and techniques drawn from the commercial sector to promote changes in diverse socially important behaviors such as drug use, sexual behavior…This marketing approach has an immense potential to affect major social problems if we can only learn how to harness its power" [34].

Study Objectives

We describe the development of a theory-based, culturally appropriate library of motivational messages for a social marketing campaign to promote COVID-19 vaccination among African Americans who are vaccine hesitant. Message development occurred in 4 phases: (1) adaptation of a message library based on the literature, (2) review by clinical and research experts for content validation, (3) input and review by a community advisory panel (CAP), and (4) message pretesting via a qualitative study with African American community members to evaluate the accuracy, relevance, and persuasiveness

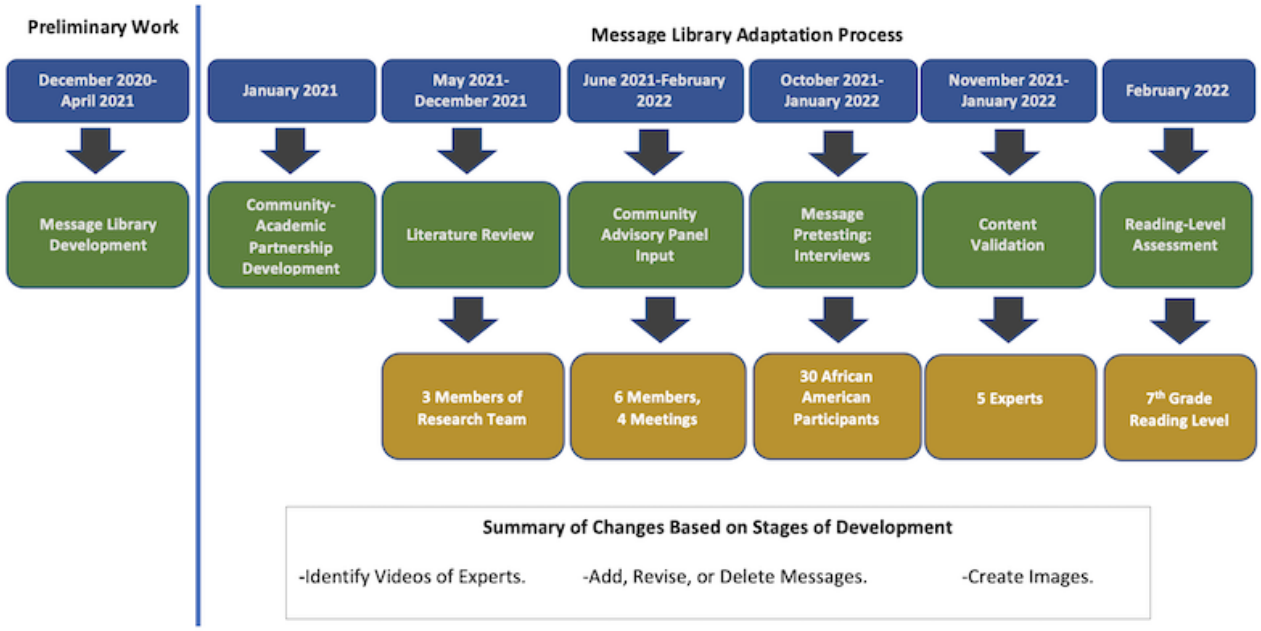
of the messages. The long-term goal is for these messages to be used within interventions aimed at increasing COVID-19 vaccination among African Americans.

Methods

Study Design

We conducted a formative study to design and develop a theory-based, culturally appropriate message library that could be used in behavioral and communication interventions to increase COVID-19 vaccine uptake among African Americans. In the existing COVID-19 library, 1 message subset for adults and 1 for parents and caregivers of adolescents was developed and iteratively adapted by coauthors JCE (a behavioral scientist with a background in biology and vaccine hesitancy) and JD (a basic scientist with a background in infectious disease) on the basis of emerging literature and feedback from over 30 educational sessions provided to communities on COVID-19 and the vaccine. A community-academic partnership was formed to iteratively adapt the message library over a 1-year period to ensure that it was theory based, culturally appropriate, and up to date. Our adaptation process occurred through 6 phases, with community input at each phase. See Figure 1 for a depiction of the 1-year process to yield the final message library of adult and parent message sets.

Figure 1. Message library development process.



Development of a Community-Academic Partnership, January 2021

A community-academic partnership was developed between 2 academic partners, Meharry Medical College and St Jude Children's Research Hospital, and 1 community partner, the Congregational Health and Education Network (CHEN) in 2021. This partnership yielded an interdisciplinary team of experts in basic science, epidemiology, behavioral science, communication, and community engagement. The purpose of this partnership was (1) to develop messages and products for a social marketing campaign using community engagement principles and (2) to implement and evaluate the social marketing campaign. CHEN, a 501c(3), is a collaboration between Nashville General Hospital at Meharry Medical College and other academic institutions, along with faith-based organizations serving African American and Latino congregants [36]. A focus for CHEN during the COVID-19 pandemic was to improve vaccine health literacy and uptake to reduce health inequities in vaccination and related outcomes. The community partner (CHEN) ensured that the messaging represented the top vaccination concerns and was culturally appropriate. The academic partners offered guidance, when needed, to ensure that the partnership was equitable, and bidirectional engagement occurred in all research phases. They also ensured scientific accuracy of messages.

Adaptation of the Existing Message Library, May-January 2022

Members of the research team conducted a literature search to identify reasons for COVID-19 vaccine hesitancy and acceptance, along with potential strategies to improve coverage among African Americans. See Multimedia Appendix 1 for the methods and results of the search. This search was conducted from May 2021 until January 2022 to ensure that the library had the latest updates before the launch of the social marketing campaign. Using the literature, the TRA [30] and the HBM [31], and the experiences and skills of the research team, we adapted the existing message library to ensure accuracy and relevance.

Input from the Community Advisory Panel, June 2021-February 2022

An 8-member CAP was formed to provide feedback on concerns about the vaccine and potential strategies to increase vaccination. This panel meets quarterly and comprises African Americans in the Nashville, Tennessee, metropolitan statistical area. Organizations represented included the Matthew Walker

Comprehensive Health Center, Health Leads, the National Association for the Advancement of Colored People, the Community Partners' Network, and the Nashville Health Disparities Coalition. Additional members included a young adult, a physician, and a parent. A subgroup of the CAP (n=4, 50%) reviewed the content and provided feedback on the messages. The feedback from the meetings and content review from the subgroup were used to iteratively develop the messages and images for the library.

Message Pretesting With African Americans, October 2021-January 2022

Study Design

We conducted a phenomenological, qualitative study [37] to create messages that could be used to assist in decision-making on COVID-19 vaccination among African Americans. Specifically, we conducted semistructured interviews (1) to identify reasons participants decided to receive or decline the vaccine for their self or child and (2) to gain feedback on messaging relevance, acceptability, and comprehensibility. This protocol was guided by the HBM [31], the TRA [30], and the community partner (CHEN) and CAP input. Social marketing campaign content (ie, draft messages, images/graphics) was iteratively revised using the data.

Sampling and Recruitment

We recruited a purposive sample of African Americans in the southeastern United States who met the following eligibility criteria: (1) adults aged 18 years or older and vaccinated or unvaccinated and (2) parents or caregivers of children aged 5-18 years and vaccinated or unvaccinated. Our community partner (CHEN) and the CAP members recruited participants via email, telephone, or word-of-mouth by using their social network and existing databases. ResearchMatch (RM), an online research recruitment tool, was also used for recruitment [38].

Data Collection

Interested participants completed a screener. If they qualified, then they completed informed consent procedures and a brief 12-item survey on barriers to COVID-19 vaccination, with response options on a 5-point Likert-type agreement scale. The screener, consent, and demographic survey were completed via Research Electronic Data Capture (REDCap; Vanderbilt University) [39], a secure web-based data collection application. Then, participants were emailed a copy of the adult and parent message sets a minimum of 3 days before the interview.

Participant interviews were conducted by a trained interviewer and lasted 45-60 minutes. Open-ended scripted questions were asked using an interview guide. Questions included (1) attitudes and beliefs about COVID-19 and vaccination, (2) facilitators and barriers to COVID-19 vaccination, (3) message library content and images, and (4) dissemination strategies. For the message development section of interviews, specific questions included the following:

- What is your overall view and purpose of the message?
- Which parts of this message did you not understand or were not clear?
- What should be added to the message?
- What message should be removed or changed?

Follow-up questions were asked for clarification and to facilitate in-depth discussion. Participants were paid a US $50 gift card. Interviews were recorded, transcribed verbatim, and de-identified for data analysis.

Data Analyses

Qualitative data coding and analysis was managed by the Vanderbilt University Qualitative Research Core, led by a PhD-level psychologist. Data coding and analysis were conducted following Consolidated Criteria for Reporting Qualitative Studies (COREQ) guidelines, an evidence-based qualitative methodology [40]. A hierarchical coding system was developed and refined using the interview guide and a preliminary review of the transcripts. Experienced qualitative coders first established reliability in using the coding system on 2 transcripts, reconciling any discrepancies, and then independently coded the remaining transcripts. We used an iterative inductive/deductive approach to qualitative data analysis [41,42]. Inductively, we sorted the quotations by coding category to identify higher-order themes and relationships between themes. Deductively, we were guided by the HBM and the TRA. The transcripts, quotations, and codes were managed using Microsoft Excel 2016 and IBM SPSS Statistics version 28.0. Survey data were also analyzed using SPSS Statistics version 28.0. Descriptive analysis (eg, means, frequencies) and bivariate analysis (eg, chi-square tests, Fisher exact tests) were performed to describe patterns in the data.

Message Library Revision Process

Using the original message library, the research team iteratively modified the message sets and added visuals and videos to "match" the messages to increase comprehensibility and appropriateness. A subgroup of the research team met regularly to discuss interview findings and the message library, and subsequent changes were made. Then, all members of the partnership met to discuss the changes and identify additional needed modifications. Each new iteration of the messages and visuals was developed using peripheral, evidential, linguistic, and sociocultural strategies [32]. A final meeting was held to ensure that all feedback was incorporated into the message library and was culturally appropriate.

Content Validation of Messages by Experts, November 2021-January 2022

Five experts were identified to review the content for accuracy and relevance. These reviews were conducted strategically due to the ever-changing nature of COVID-19 pandemic updates. Specifically, 2 experts reviewed the content prior to, 1 during, and 3 after community review. Selection criteria for these content reviewers were experience in SARS-CoV-2/COVID-19 research, including vaccination, expertise in the use of psychosocial theory, and willingness and ability to review the items.

Adapted from Lawshe [43], we used qualitative and quantitative methods in a 2-phase content review process. Specifically, each message subset was quantitatively evaluated for relevance by using a 3-point Likert scale: "essential," "useful but not

essential," and "not necessary." Then, experts reviewed the messages for perceived accuracy and clarity by providing written comments and edits for improvement. Using these methods, we iteratively refined the messages.

Ethical Considerations

This research was approved by Meharry Medical College's Institutional Review Board (Protocol 21-03-1076). All participants provided oral informed consent.

Results

Community Advisory Panel Feedback

The CAP members' feedback evolved throughout the message library development process. At the beginning of each meeting, the researchers gained insight into the community's thinking about the COVID-19 vaccine to determine trends in hesitancy. Top concerns included vaccine safety, the speed of vaccine development, mistrust in research and health care, politicization of the vaccine, and conspiracy theories (eg, tracking chip in the vaccine). We then asked about the presentation of messaging. Members suggested that messages should be concise yet comprehensible across reading levels. Members further indicated the need to discuss immediate and long-term benefits and the risks of vaccination so that the community can make an informed decision. A few members further suggested the need to use numbers and images to explain these concerns more clearly.

Select members of the CAP were asked to conduct a detailed review of the messages and images to identify ways to make the content more relatable and comprehensible. Some even provided preferred sources of content (eg, NPR (National Public Radio), which is media organization that seeks to create a more informed public via air, online, or in-person) to help develop the material. Collectively, we used the feedback from the quarterly advisory panel meetings and subgroup review of messages to update the library.

Semistructured Interviews

Sociodemographics

Most participants were female and had a college degree or higher. More than half had a household income less than US $80,000. The mean age was 38.6 years (SD 9.49 years). See Table 1 for sociodemographics by subgroup: vaccinated adult, unvaccinated adult, adult with unvaccinated child, and adult with vaccinated child.

Table 1. Sociodemographics of African American interview participants (N=30).

Characteristic	Parent[a] with vaccinated child (N=7)	Parent with unvaccinated child (N=7)	Adult, vaccinated (N=9)	Adult, unvaccinated (N=7)
Age (years), mean (SD)	42.4 (6.1)	37.4 (6.2)	36.1 (12.5)	39.1 (11.1)
Gender, n (%)				
Male	2 (29)	1 (14)	2 (22)	1 (14)
Female	5 (71)	6 (86)	7 (78)	6 (86)
Education, n (%)				
Some college or lower	2 (29)	2 (29)	2 (22)	1 (14)
College degree or higher	5 (71)	5 (71)	7 (78)	6 (86)
Household income (US $), n (%)				
≤40,000	2 (29)	2 (29)	2 (22)	3 (42)
40,001-80,000	2 (29)	0	2 (22)	2 (29)
>80,000	3 (42)	4 (57)	3 (34)	0
Not available	0	1 (14)	2 (22)	2 (29)

[a]Parents had children aged 5-18 years.

Summary of Findings

Using the inductive-deductive approach, we identified 4 primary themes specific to the development of messages related to COVID-19 and the vaccines in the qualitative study. Messages were referenced in the text by the theme and message number (eg, 1.01 is the first quotation related to theme 1). See Multimedia Appendix 2 for quotations related to each theme.

Theme 1: Community Reactions to the Messages

Overall, the community members found that the messages were "very helpful" and had a "community feel." Specifically, the messages were inclusive and comprehensive and had a good balance between science and simple language. The messages were viewed as "good," "persuasive" (quotation 1.02), and "educational" (quotations 1.01-1.02). Most participants stated that the reading level was good, and suggested a few edits to specific messages (eg, whether the messenger RNA [mRNA] vaccine changes your DNA). Although there were mixed reviews, most perceived the length and number of messages in each set to be appropriate.

For message presentation, many perceived that there were good analogies and comparisons to increase comprehension (quotation 1.03). If applicable to the message set, participants liked the balance of benefits and risks of vaccines. Lastly, a few

participants cited the messages as relatable, trustworthy, and credible. Collectively, participants perceived that the information helped guide decision-making and did not simply "tell you" to take the vaccine.

Theme 2: Questions and Information Needs

Some participants had questions after reviewing the messages. A participant wanted to know where specific evidence for the general numbers on some websites (eg, the Centers for Disease Control and Prevention [44]) could be found. Other questions were related to long-term effects of the vaccine, "why we even need the vaccine," mixing vaccines and boosters, or the number of boosters after the first dose (quotation 2.01). Others asked about the relationship of COVID-19 vaccines and fertility. Lastly, participants asked how existing conditions (eg, diabetes, asthma) were related to the severity of COVID-19.

Theme 3: Suggestions for Additional Content

Most participants had suggestions for additional content or context for specific message sets. A suggestion for overall messages was to add a statement that science evolves as more data are collected to communicate new findings, which are constantly being added to increase our knowledge of COVID-19. For vaccine safety, additional information was requested on mRNA, along with the 30-year history of studies of mRNA and its use in vaccines (quotation 3.01). Participants asked for more information on the clinical trial process and better justification for boosters and their side effects. As it relates to vaccines and infertility, a suggestion was made to provide recommendations from gynecology professionals and experts and information about long-term effects of COVID-19 vaccines on infertility. Lastly, a few participants wanted to know whether the vaccines can lead to sexual dysfunction (quotation 3.02).

Distrusting in the government and pharmaceutical companies, participants wanted more details about the COVID vaccines (ie, development, testing, and ingredients; quotation 3.03). For vaccine effectiveness, participants asked about healthy people getting the virus and how outcomes of the vaccinated compare to those of the unvaccinated (quotation 3.04). They also asked why natural immunity is not better than vaccine-induced immunity. Participants wanted more information on variants (eg, Omicron), their severity, and how variants affect vaccine effectiveness.

For children specifically, participants asked for data on successes and challenges during development (quotation 3.05), along with updates on vaccine safety. Participants also suggested reinforcing other preventive behaviors, such as sanitizing, healthy eating, and physical distancing (quotation 3.06). Some further wanted to know alternative ways to boost their immune system without taking a vaccine or booster (quotations 3.07 and 3.08). A few suggested the need to encourage conversations with doctors, especially for those with underlying medical conditions, prior to getting the vaccine.

Theme 4: Suggestions to Increase Comprehension, Relevance, and Trustworthiness

A few suggestions were made to enhance comprehension. One suggestion was to provide definitions for specific terms (eg, high risk; quotation 4.01). Providing an easier-to-understand presentation of statistics was commonly mentioned. For example, participants further suggested the "need for statistics or more data" to help understand the vaccine development process. Participants emphasized the importance of clarifying the magnitude of potential side effects. Visuals were suggested to improve understandability of messages. Participants also suggested the use of specific terms such as "vaccination" and not "shot." Lastly, 1 (3%) participant suggested having information available in other languages (eg, Arabic).

A few participants indicated the need for messages to be tied to things people already understand, such as flu or smallpox vaccines, to increase relevance. Other suggestions were to use videos and images to reflect content. Some participants also wanted videos of personal testimonies of individuals who were undecided about the vaccine and their decision-making process to get vaccinated (quotation 4.02). Notably, testimonials were also perceived by others to be too contrived (quotation 4.03). To increase trustworthiness, participants suggested providing proof that doctors or health professionals (quotation 4.04) supported this work, along with the addition of informational sources, especially links to studies that provide supportive evidence. Lastly, participants encouraged honesty and transparency in information related to the COVID-19 vaccines.

Content Validation

In total, 6 reviewers validated the content. Of these 6, 2 (33%) were White Americans, and 4 (67%) were African Americans; 4 (67%) were female, and 2 (33%) were male. In addition, 1 (17%) reviewer provided feedback for only adult concerns. Reviewers' areas of expertise included vaccine development, immunology, vaccine-preventable disease and immunizations, and nurse safety. Messages were iteratively updated based on feedback of experts. Most message subsets were classified as essential or useful but not essential. Content edits and additions to the message library reflected the updates on the coronavirus and the vaccine and strategies to ensure comprehensibility and accuracy. Reordering, rephrasing, and adding (eg, analogies) of content were conducted to increase clarity. Because updates are ongoing for COVID-19, content was deemed evidential once primary concepts (eg, process of vaccine development or purpose of boosters) were validated, with the intent to continue to update the library with expert review.

Reading Grade Level Assessment

To finalize the library messages, a readability assessment was conducted. We used 3 primary reading grade level assessment tools and an online consensus tool to assess the reading grade level of the adult and parent message sets. First, the Flesch Reading Ease Score [45] was calculated in Microsoft Word. Higher scores indicate easier readability by the user. Second, the Flesch-Kincaid Grade Level [46] was calculated in Microsoft Word to determine a US school grade level. The Flesch Reading Ease and Flesch-Kincaid Grade Levels are calculated by considering the average sentence length (total number of words divided by total number of sentences) and the average syllables per word (total number of syllables divided by the total number of words) using different underlying formulas. Third, the Simple Measure of Gobbledygook (SMOG) index was hand-scored, in

addition to being automatically calculated to determine a US [47] school grade level. The SMOG index is calculated using the number of words with multiple syllables in three 10-sentence samples at the beginning, middle, and end of the text. These 3 readability assessment tools all use word difficulty and sentence length as the main factors in determining how easy or how hard the material is to read. Finally, an online readability calculator [48] was used to determine readability consensus. The online calculator applied 7 commonly used readability formulas to provide a consensus rating. See Table 2 for readability assessment results.

The readability assessment revealed acceptable reading grade level scores for all readability formulas applied. The usual recommended reading grade level is fifth-sixth grade to optimize comprehension, according to the American Medical Association and the United States Department of Health and Human Services. However, in combination with the iterative development process and ongoing review by the community partner (CHEN), the CAP, and experts, the messages reflected use of plain language when polysyllabic or complex words were unavoidable.

Table 2. Reading grade level results.

Readability assessment tool	Adult library	Parent library
Flesch Reading Ease Score	62.0 (standard/average)	62.5 (standard/average)
Flesch-Kincaid Grade Level	8.4 (8th grade)	8.3 (8th grade)
SMOG[a] index	7.7 (8th grade)	7.6 (8th grade)
Readability consensus[b]	• Grade level: 8 • Reading level: standard/average • Age of reader: 12-14 years (7th-8th grade)	• Grade level: 8 • Reading level: standard/average • Age of reader: 12-14 years (7th-8th grade)

[a]SMOG: Simple Measure of Gobbledygook.

[b]Readability consensus was based on the application of 7 readability formulas using an online calculator available [44].

Message Set Finalization

After completion of the readability assessment, members of the research team conducted a final review to ensure comprehensibility and accuracy. Minor edits were made. See Table 3 for an example of a concern, along with an example of a message for the concern after the adaptation process.

Table 3. Example of a message for tailoring a variable postiterative development process.

Vaccine concern	Example message
Category 1: vaccine development	
"Human protections in research" [adult and parent]	Many people wonder about taking part in research and if they will be protected. We know there have been past research studies that were not done right [provides examples of historical research abuses]. To begin to address this issue, we give a few examples to show every person is protected when they take part in research and how the community could benefit after the research study is done [provides examples of human protections in research].
"Who is at the table?" [Adult and parent]	Many people wonder if people of all racial backgrounds were involved when the vaccines were developed. Individuals from all races were at the table to help guide the process [provides examples of scientists of all racial/ethnic backgrounds and their role in the development process].
Category 2: vaccine safety	
"mRNA[a] and DNA" [adult]	We all have mRNA in every cell in our bodies. mRNA is known as messenger RNA. It is the "recipe" that tells the cells in our body to make certain proteins. The mRNA protein in the Pfizer and Moderna COVID-19 vaccines shows up, teaches the immune system how to develop antibodies against SARS-CoV-2 (the virus that causes COVID-19), and then quickly dissolves. mRNA never enters the nucleus of the cell where your DNA is kept. Your body learns how to protect itself against future SARS-CoV-2 infection without ever having to risk getting the virus or the serious outcomes of getting sick with COVID-19.
"Infertility" [adult]	The antibody to the spike protein does not make a woman infertile or unable to get pregnant. There was a false claim that there were similarities between the SARS-CoV-2 spike protein and the surface of a protein on placental cells. Placental cells are needed for a successful pregnancy. SARS-CoV-2 spike protein and the placental cells are not the same. This means the vaccine will not cause the immune system to make antibodies against the placental protein.
"Your child's heart" [parent]	Many parents ask about the COVID-19 vaccine and how it can affect the hearts of children. In the United States, there has been an increase in myocarditis and pericarditis cases after getting the mRNA COVID-19 vaccine. Myocarditis is mild inflammation of the heart. Pericarditis is mild inflammation of the sack around the heart. For children under 16 years of age, myocarditis risk is 37 times higher for children with COVID-19 than the children without COVID-19. So, myocarditis does not happen often. The American Academy of Pediatrics says children and teens should get the COVID-19 vaccines.
"My child has underlying medical conditions." [Parent]	The Pfizer vaccine can be given to children 5 [years] and older with underlying medical conditions like diabetes or autoimmune diseases. It cannot cause COVID-19, even in those with weak immune systems. Children with underlying medical conditions took part in the clinical trials and serious reactions to the vaccine [were] rare. However, children with underlying medical conditions are more apt to have problems from COVID-19.
Category 3: vaccine effectiveness	
"Boosters. Why?" [Adult]	All routine [vaccines] require booster doses to have full protection [adds examples]. So the COVID-19 vaccine is not any different. Booster shots are given to jumpstart the body's immune system to produce more antibodies against the original SARS-CoV-2 (the virus that causes COVID-19) and help protect against new variants. Because antibody levels decrease over time, boosters are needed to keep us protected.
"Variants and the vaccine" [adult]	As SARS-CoV-2 (the virus that causes COVID-19) continues to infect people, it is more likely to mutate. This means that the virus makes a new version of itself also known as a mutation. It is common for this to happen. Mutations affect how viruses work, like to help the virus better attach to our cells or lower the virus's ability to attach to our cells. So it is important for people to complete COVID-19 vaccination. More mutations and new variants may lower or stop the protection provided by the vaccines.
"Natural immunity or vaccine immunity"	Natural immunity happens when your child's body gets infected with the SARS-CoV-2 virus, the virus that causes COVID-19. While your child's body will make antibodies against the virus, the danger is in your child getting very sick and maybe even dying. Immunity from getting a vaccine is very similar to immunity developed through natural infection but does not carry the increased risk of your child getting very sick or even death. Natural immunity provides less protection over time than the immunity gained by COVID-19. While people can gain immunity after getting the virus, studies show that more than one third of COVID-19 infections results in low levels of protective antibodies.
"Too many vaccines" [parent]	We all got different vaccines when we were babies, adolescents, and event adults [gives an example of multiple vaccines being given]. These vaccines are routinely given at the same time without serious side effects. So, getting more than one vaccine is something we have been doing since birth.

[a]mRNA: messenger RNA.

Final Message Library

The final message library had 2 message sets, 1 for adults and 1 for parents. There were 18 message subsets for adults and 17 message subsets for parents. These subsets were placed into 3 categories: vaccine development, vaccine safety, and vaccine effectiveness. Each message subset begins with expressing empathy toward the individuals' concern. Then the facts are provided around each concern, positive or negative. Lastly, the message subset ends with a positive statement related to COVID-19 vaccination that addresses concerns. All message subsets were reviewed by community leaders and members (constituent-involving strategy). We briefly describe each subset for each group next. Each message subset was presented using 3 modes: content, image, and video. See Table 4 to identify each concern, along with key message attributes and the associated strategy to achieve cultural appropriateness.

Table 4. Vaccine concern, key message attributes, types of visuals, and type of cultural targeting strategy.

Vaccine concern[a]	Key message attributes	Types of visuals	Cultural targeting strategy
Vaccine development			
Human protections in research	• Acknowledge historical research abuses (eg, U.S. Public Health Service Syphilis Study, Henrietta Lacks). • Provide examples of protections provided in research. • Discuss how the community can be involved in the research process (eg, co–primary investigator, community advisory board, consultant).	• Not applicable	Sociocultural, evidential
Who is at the table?	• Discuss researchers (ie, current job, expertise) across all backgrounds and their role in the development process. • List their current jobs.	• Researchers of diverse backgrounds (visual) • Researcher describing role in vaccine development (video)	Sociocultural, evidential, peripheral
Too new and too quick	• Define mRNA[b] and its role in the body. • Discuss the mRNA vaccine history and how it works in the body. • Define what mRNA does not do (change DNA). • Compare mRNA vaccine development to existing vaccine development processes. • Define the EUA[c]. • Discuss the number of vaccines given to date, adverse events, and how to identify those events. • Confirm that being unvaccinated places one at higher risk of death compared to those who received the vaccine.	• Timeline of COVID-19 vaccine development (visual and video)	Evidential, peripheral
How research works	• Define research. • Define clinical trials and their phases. • Define types of researchers. • Define sites of research and who can participate. • Discuss what happens after research.	• Demonstration of phases in the clinical process and steps within each phase (visual)	Evidential, peripheral
Vaccine safety			
mRNA and DNA	• State years of mRNA existence and mRNA's role in the body. • Explain the process of mRNA technology. • Identify vaccines that use mRNA technology. • Emphasize benefits of vaccination over natural immunity.	• Demonstration of the mRNA technology process (visual and video)	Evidential, peripheral, linguistic
Infertility	• Demonstrate how the proteins needed for pregnancy and needed to make the spike protein are not the same. • Emphasize that women are able to conceive, have a healthy pregnancy and baby, and breastfeed after vaccination. • State that babies receive antibodies from vaccinated mothers. • Highlight that COVID-19 may impact fertility in men. • Highlight cons of nonvaccination in pregnant women (eg, increased risk of stillbirth, newborn deaths, hospitalization).	• Explanation of pregnant women getting the vaccine or women conceiving getting vaccinated (video)	Evidential, linguistic

Vaccine concern[a]	Key message attributes	Types of visuals	Cultural targeting strategy
Underlying medical conditions	• State the reason to vaccinate with an underlying condition. • State that the vaccine will not give an individual COVID-19. • Discuss the vaccine schedule for those immunocompromised. • Discuss the severity in COVID-19 if not vaccinated. • State to consult with a doctor in getting the vaccine.	• Explanation of why those with underlying medical conditions need the vaccine (video)	Evidential, linguistic
Your heart	• Define myocarditis and pericarditis. • Compare the rate of heart problems in those who get vaccinated compared to those who get COVID-19. • Demonstrate the symptoms and treatment of heart disorders. • Emphasize the recommendation by infectious disease experts and the American Academy of Pediatrics for children.	• Explanation of myocarditis and vaccination (video)	Evidential, linguistic
GBS[d] (adult only)	• Define GBS. • Discuss the signs of GBS. • Discuss the number of cases to date after vaccination. • Emphasize that it is rare.	• Not applicable	Evidential, linguistic
Blood clots (adult only)	• Identify the number of cases with the Johnson & Johnson vaccine. • Discuss why the Food and Drug Administration (FDA) halted the clinical trial to determine whether risks of blood clots outweigh the benefits of the vaccine. • State recommendation of Moderna and Pfizer vaccines over the Johnson & Johnson vaccine. • Discuss the symptoms of blood clots. • State the blood clot risk for those vaccinated and unvaccinated.	• Symptoms of blood clots vs symptoms of COVID-19 (visual)	Peripheral, evidential, linguistic
Side effects	• Emphasize the number of years for COVID-19 research. • Discuss the number of lives and hospitalizations prevented with vaccines. • Identify the risk of allergic reactions and short-term side effects. • Discuss that side effects are short-lived and everyone reacts differently. • State that routine vaccinations show no long-term side effects.	• Side effects of vaccination compared to natural infection through SARS-CoV-2 (visual and video)	Peripheral, evidential, linguistic
Too young (parent only)	• Emphasize the impact of COVID-19 on children. • Provide recommendations for COVID-19 vaccination by age. • Emphasize that vaccination protects them and others. • State that the long-term effects of COVID-19 in children are unknown, but long COVID is seen in many.	• Statistics of current COVID-19 cases in children and increases in COVID-19 cases, hospitalizations, and deaths in children overtime (visual)	Evidential, peripheral, linguistic
Vaccine effectiveness			

Vaccine concern[a]	Key message attributes	Types of visuals	Cultural targeting strategy
Unsure if it works	• Define effectiveness and how to obtain it (ie, fully vaccinated). • Demonstrate risks if not vaccinated.	• Comparison of risk of hospitalization and death of those vaccinated vs not vaccinated (visual)	Peripheral, evidential, linguistic
Variants	• Define "breakthrough case." • Discuss mutations and how new variants are created. • Discuss the impact of emerging variants on vaccines and health. • Emphasize the impact of virus on short- and long-term health.	• Not applicable	Evidential, linguistic
Natural immunity	• Define natural immunity versus vaccine-induced immunity. • Emphasize vaccine-induced immunity being much safer than natural immunity. • Discuss the "gamble" in natural immunity over vaccine-induced immunity. • Discuss the benefits of vaccination despite having COVID-19.	• Comparison of the health risks of those with natural immunity and those vaccinated (visual and video)	Peripheral, evidential, linguistic
Too many vaccines	• Discuss the lack of danger of multiple vaccines at a time. • Remind people of receiving many vaccines at once as a baby and preteen. • Compare the number of proteins in the vaccine to the number of proteins if exposed to SARS-CoV-2.	• Not applicable	Evidential, linguistic
Boosters. Why?	• Define boosters and why they are needed. • Discuss booster recommendations. • Emphasize discussing getting a booster with a provider.	• Stating the vaccine dose and booster schedule of each vaccine (image) • Defining a booster and why we need it (video)	Evidential, peripheral, linguistic
Is it even needed?	• Discuss transmission routes and rates by variant. • Compare COVID-19 hospitalization, long COVID, and death rates among those vaccinated and unvaccinated. • Discuss the susceptibility and severity of COVID-19 and the importance of vaccination.	• Tracker of COVID-19 rates and deaths (United States and Tennessee)	Evidential, peripheral, linguistic

[a]All vaccine concerns were vetted by community leaders and members (ie, constituent-involving strategy) and edited to be comprehendible (ie, linguistics).

[b]mRNA: messenger RNA.

[c]EAU: emergency use authorization.

[d]GBS: Guillain-Barré syndrome.

Vaccine Development

These message subsets target individuals who have concerns about the COVID-19 vaccines and the development process. The goal is to positively influence attitudes toward researchers and the process. There are 4 message sets in this category:

- Human protections in research/child protections in research
- Who is at the table?
- Too new and too quick
- How research works

Vaccine Safety

These message subsets target individuals who have concerns about the safety of the COVID-19 vaccines. The goal is to demonstrate that the benefits of COVID-19 vaccination outweigh the harms of COVID-19 vaccination. There are 8 message sets in this category:

- mRNA and DNA
- Infertility/youth infertility
- Underlying medical conditions
- Your heart/your child's heart
- Guillain-Barré syndrome (GBS; adults only)
- Blood clots (adults only)
- Side effects
- Child is too young (parents only)

Vaccine Effectiveness

These message subsets seek to demonstrate that the risk of SARS-CoV-2 and the severity of COVID-19 (ie, long-haul COVID-19, hospitalization, and death) are far greater when not vaccinated against COVID-19. These sets further demonstrate that the vaccine is effective and how variants may affect effectiveness. We also discussed the dosing schedule and role of boosters. There are 7 message sets in this category:

- Unsure if it works
- Variants and the vaccine
- Natural immunity or vaccine immunity
- Too many vaccines
- Boosters. Why?
- Is it even needed?

Discussion

Principal Findings

Our study aimed to develop and validate a message library for a social marketing campaign to increase COVID-19 vaccination among African Americans. The goal was to provide African American adults and parents with theory-based, culturally appropriate messaging on COVID-19 vaccines to motivate vaccine uptake. We described a multiphase process using community engagement approaches for the message library development with the HBM [31], the TRA [30], and Kreuter's [32] cultural targeting strategies serving as the conceptual frameworks. Our existing library allowed us to expeditiously adapt the messaging to meet the needs of African Americans. This process can be used by researchers and health care professionals to inform the development of culturally appropriate messages.

Use of formative research to build theory-based, culturally appropriate messaging while applying community engagement principles is critical for communities to play an active role in disease prevention and control measures, such as COVID-19 vaccination [26,49]. This method holds great promise in addressing health disparities, yet is in its infancy [50,51]. Applying Boyer et al's [52] multilevel approach to stakeholder engagement, we had community member involvement at all phases and varying levels to develop a message library to promote COVID-19 vaccination among African Americans. Engagement approaches included formation of a community-academic partnership, a CAP, and inclusion of community interviewees. Having the community-academic partnership and CAP allowed the community voice to be at the root of the messaging. Using each engagement approach, there was a balance of power to ensure that there was bidirectional communication and a deliberative process to foster respect, and even trust in some instances [53]. Furthermore, this process increased the likelihood of achieving cultural appropriateness of the messages.

Content validation has been recognized as a necessary component of message development and is highly valued [54]. The feedback provided by experts in the content review process was used to evolve the library with accurate and relevant messages. Furthermore, the suggestions for modification enriched the messages. Messages were further tested with a purposeful sample of African Americans for cultural appropriateness (ie, evidential, linguistics, peripheral, and sociocultural strategies). Our results indicated that African American adults and parents viewed the messages positively and indicated that the messages were *persuasive*, *useful*, and *trustworthy*. Feedback yielded distinct strategies to increase relevance, comprehension, and appeal. It is important to understand the target audiences' response early to determine the likelihood of message effectiveness for the intervention [55].

Using this feedback from a multiphase process, our final message library yielded 18 message subsets for adults, and there were 17 message subsets for parents that were grounded in theory and cultural-targeting strategies. There were 3 preferred modes (ie, messages, images, and videos) for African American adults and parents. Studies demonstrate that multiple modes of communication are effective in increasing health literacy among populations, and plain-language messages, pictures, and videos are commonly cited, particularly in the context of community-level interventions [56]. We believe that this approach will be effective in reaching different characteristics of individuals.

Strengths

A major strength of this study is the use of theory and culturally appropriate strategies inclusive of community engagement to develop the COVID-19 message library for African American adults and parents. We used different levels of community engagement (ie, community-academic partnership, CAP, and interviewees) to ensure that the messages met the needs of our target population. In addition, we equipped the community with information about COVID-19 vaccines to ease concerns postvaccination or to make an informed decision about getting the vaccine. Furthermore, these individuals can now serve as education resources to their communities.

The next step in our partnership will be to test these messages in a 5-month social marketing campaign in a pilot study. Specifically, these messages are used on a website to provide information on COVID-19 vaccines. Shortened versions of these messages are used to market the website. We will evaluate the impact on attitudes, willingness, and self-reported vaccination status to be reported in a future manuscript. If the intervention demonstrates effectiveness, it could prove that theory-based, culturally appropriate messages in a social marketing campaign can be used as a motivational tool among African Americans.

Limitations

This study has limitations. Messages may not be generalizable to African Americans outside the southeastern United States. We had a small, purposeful sample, yet findings explained diverse perspectives to ensure messages encompassed multiple viewpoints toward the vaccine. There is potential for selection bias among content experts as they are medical professionals and clinicians from different disciplines and with clinical or research expertise. Furthermore, lack of access (ie, geographical barriers) to the vaccine could prevent uptake regardless of other concerns being addressed.

Conclusion

Vaccine hesitancy continues to negatively impact COVID-19 vaccination among African Americans. Effective interventions are needed to increase vaccine uptake. We believe we have developed validated and pretested theory-based, culturally appropriate messages that can be motivational in different interventions aimed at increasing the COVID-19 vaccination rate among African Americans.

Acknowledgments

The authors thank members of the African American community for shedding light on the reasons for COVID-19 vaccination hesitancy. We also thank the COVID-19 experts for their invaluable insights into the intervention, particularly its content, images, and channels of delivery. This work was supported by the National Institute on Minority Health and Health Disparities (Award 3U54MD007586-34S7), the National Cancer Institute of the National Institutes of Health (Award P30CA021765 and 1K01CA237748-01A1), and the American Lebanese and Syrian Associated Charities (ALSAC) of St Jude Children's Research Hospital. This work is also supported by a subaward from the Association of American Medical Colleges through its cooperative agreement with the Centers for Disease Control and Prevention (CDC): "Improving Clinical and Public Health Outcomes through National Partnerships to Prevent and Control Emerging and Re-Emerging Infectious Disease Threats" (Award 6 NU50CK000586-02-02). The CDC is an agency within the Department of Health and Human Services (HHS). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the CDC, or the HHS and should not be considered an endorsement by the federal government.

Authors' Contributions

JCE was responsible for conceptualization, methodology, validation, formal analysis, resources, data curation, writing—original draft, writing—review and editing, visualization, supervision, planning administration, and funding acquisition; HMB and JD for conceptualization, writing—review and editing, visualization, supervision, planning administration, and funding acquisition; MS for conceptualization, writing—review and editing, visualization, and funding acquisition; KC and OL for conceptualization and writing—review and editing; and DS and KB for formal analysis and writing—review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Methods and results of the literature search.
[DOCX File , 37 KB - formative_v6i7e38781_app1.docx]

Multimedia Appendix 2
Qualitative quotations for themes.
[DOCX File , 17 KB - formative_v6i7e38781_app2.docx]

References

1. Centers for Disease Control and Prevention. Different COVID-19 Vaccines. 2022. URL: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines.html [accessed 2022-01-21]
2. Centers for Disease Control and Prevention. COVID-19 Vaccines Are Effective. 2022. URL: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/effectiveness/index.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fvaccines%2Feffectiveness.html [accessed 2022-03-11]
3. Centers for Disease Control and Prevention. Demographic Characteristics of People Receiving COVID-19 Vaccinations in the United States. 2022. URL: https://covid.cdc.gov/covid-data-tracker/#vaccinations_vacc-total-admin-rate-total [accessed 2022-03-10]
4. Centers for Disease Control and Prevention. Hospitalization and Death by Race/Ethnicity. 2022. URL: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html [accessed 2022-03-25]
5. Webb Hooper M, Nápoles AM, Pérez-Stable EJ. No populations left behind: vaccine hesitancy and equitable diffusion of effective COVID-19 vaccines. J Gen Intern Med 2021 Jul 22;36(7):2130-2133 [FREE Full text] [doi: 10.1007/s11606-021-06698-5] [Medline: 33754319]
6. Willis DE, Andersen JA, Bryant-Moore K, Selig JP, Long CR, Felix HC, et al. COVID-19 vaccine hesitancy: race/ethnicity, trust, and fear. Clin Transl Sci 2021 Nov 02;14(6):2200-2207 [FREE Full text] [doi: 10.1111/cts.13077] [Medline: 34213073]
7. Cunningham-Erves J, Mayer CS, Han X, Fike L, Yu C, Tousey PM, et al. Factors influencing intent to receive COVID-19 vaccination among Black and White adults in the southeastern United States, October - December 2020. Hum Vaccin Immunother 2021 Dec 02;17(12):4761-4798. [doi: 10.1080/21645515.2021.1984134] [Medline: 34847822]

8. McElfish PA, Willis DE, Shah SK, Bryant-Moore K, Rojo MO, Selig JP. Sociodemographic determinants of COVID-19 vaccine hesitancy, fear of infection, and protection self-efficacy. J Prim Care Community Health 2021 Aug 24;12:21501327211040746 [FREE Full text] [doi: 10.1177/21501327211040746] [Medline: 34427126]
9. Majee W, Anakwe A, Onyeaka K, Harvey IS. The past is so present: understanding COVID-19 vaccine hesitancy among African American adults using qualitative data. J Racial Ethn Health Disparities 2022 Feb 19:1-13 [FREE Full text] [doi: 10.1007/s40615-022-01236-3] [Medline: 35182372]
10. Kerrigan D, Mantsios A, Karver TS, Davis W, Taggart T, Calabrese SK, et al. Context and considerations for the development of community-informed health communication messaging to support equitable uptake of COVID-19 vaccines among communities of color in Washington, DC. J Racial Ethn Health Disparities 2022 Feb 03:1-15 [FREE Full text] [doi: 10.1007/s40615-022-01231-8] [Medline: 35118609]
11. Sharma M, Batra K, Batra R. A theory-based analysis of COVID-19 vaccine hesitancy among African Americans in the United States: a recent evidence. Healthcare (Basel) 2021 Sep 27;9(10):1273 [FREE Full text] [doi: 10.3390/healthcare9101273] [Medline: 34682953]
12. Bogart LM, Dong L, Gandhi P, Klein DJ, Smith TL, Ryan S, et al. COVID-19 vaccine intentions and mistrust in a national sample of Black Americans. J Natl Med Assoc 2022 Jan;113(6):599-611 [FREE Full text] [doi: 10.1016/j.jnma.2021.05.011] [Medline: 34158171]
13. Momplaisir F, Haynes N, Nkwihoreze H, Nelson M, Werner RM, Jemmott J. Understanding drivers of coronavirus disease 2019 vaccine hesitancy among Blacks. Clin Infect Dis 2021 Nov 16;73(10):1784-1789 [FREE Full text] [doi: 10.1093/cid/ciab102] [Medline: 33560346]
14. Bajaj SS, Stanford FC. Beyond Tuskegee — vaccine distrust and everyday racism. N Engl J Med 2021 Feb 04;384(5):e12. [doi: 10.1056/nejmpv2035827]
15. Khubchandani J, Macias Y. COVID-19 vaccination hesitancy in Hispanics and African-Americans: a review and recommendations for practice. Brain Behav Immun Health 2021 Aug;15:100277 [FREE Full text] [doi: 10.1016/j.bbih.2021.100277] [Medline: 34036287]
16. Reddy B, Gupta A. Importance of effective communication during COVID-19 infodemic. J Family Med Prim Care 2020 Aug;9(8):3793-3796 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_719_20] [Medline: 33110769]
17. Cooper LA, Stoney CM. Messages to increase COVID-19 knowledge in communities of color: what matters most? Ann Intern Med 2021 Apr;174(4):554-555. [doi: 10.7326/m20-8057]
18. James EK, Bokemper SE, Gerber AS, Omer SB, Huber GA. Persuasive messaging to increase COVID-19 vaccine uptake intentions. Vaccine 2021 Dec 03;39(49):7158-7165 [FREE Full text] [doi: 10.1016/j.vaccine.2021.10.039] [Medline: 34774363]
19. Borah P, Hwang J, Hsu YC. COVID-19 vaccination attitudes and intention: message framing and the moderating role of perceived vaccine benefits. J Health Commun 2021 Aug 03;26(8):523-533. [doi: 10.1080/10810730.2021.1966687] [Medline: 34424140]
20. Privor-Dumm L, King T. Community-based strategies to engage pastors can help address vaccine hesitancy and health disparities in Black communities. J Health Commun 2020 Oct 02;25(10):827-830. [doi: 10.1080/10810730.2021.1873463] [Medline: 33719889]
21. World Health Organization. COVID-19 Message Library. URL: https://www.who.int/publications/i/item/covid-19-message-library [accessed 2022-07-07]
22. The Rockefellar Foundation. Vaccine Confidence Message Brief. 2021. URL: https://www.covidcollaborative.us/assets/uploads/pdf/STAT-Vaccine-Confidence-Message-Brief.pdf [accessed 2022-07-07]
23. Jensen UT, Ayers S, Koskan AM. Video-based messages to reduce COVID-19 vaccine hesitancy and nudge vaccination intentions. PLoS One 2022 Apr 6;17(4):e0265736 [FREE Full text] [doi: 10.1371/journal.pone.0265736] [Medline: 35385505]
24. Dai H, Saccardo S, Han MA, Roh L, Raja N, Vangala S, et al. Behavioural nudges increase COVID-19 vaccinations. Nature 2021 Sep 02;597(7876):404-409 [FREE Full text] [doi: 10.1038/s41586-021-03843-2] [Medline: 34340242]
25. Dada D, Djiometio JN, McFadden SM, Demeke J, Vlahov D, Wilton L, et al. Strategies that promote equity in COVID-19 vaccine uptake for Black communities: a review. J Urban Health 2022 Feb 11;99(1):15-27 [FREE Full text] [doi: 10.1007/s11524-021-00594-3] [Medline: 35018612]
26. Shafiq M, Elharake JA, Malik AA, McFadden SM, Aguolu OG, Omer SB. COVID-19 sources of information, knowledge, and preventive behaviors among the US adult population. J Public Health Manag Pract 2021;27(3):278-284. [doi: 10.1097/PHH.0000000000001348] [Medline: 33762543]
27. Yates JF, de Oliveira S. Culture and decision making. Organ Behav Hum Decis Process 2016 Sep;136:106-118 [FREE Full text] [doi: 10.1016/j.obhdp.2016.05.003] [Medline: 32288179]
28. Kreuter M, Farrell D, Olevitch L, Brennan L. Tailored Health Messages: Customizing Communication with Computer Technology. Mahwah, NJ: Lawrence Erlbaum; 1999.
29. 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. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]

30. Hill RJ, Fishbein M, Ajzen I. Belief, attitude, intention and behavior: an introduction to theory and research. Contemp Sociol 1977 Mar;6(2):244. [doi: 10.2307/2065853]
31. Rosenstock IM. The Health Belief Model and preventive health behavior. Health Educ Monogr 1974 Dec 01;2(4):354-386. [doi: 10.1177/109019817400200405]
32. Kreuter MW, Lukwago SN, Bucholtz RDDC, 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]
33. Jedele J, Ismail A. Evaluation of a multifaceted social marketing campaign to increase awareness of and screening for oral cancer in African Americans. Community Dent Oral Epidemiol 2010 Aug;38(4):371-382. [doi: 10.1111/j.1600-0528.2010.00545.x] [Medline: 20646014]
34. Ngui E, Hamilton C, Nugent M, Simpson P, Willis E. Evaluation of a social marketing campaign to increase awareness of immunizations for urban low-income children. WMJ 2015 Feb;114(1):10-15 [FREE Full text] [Medline: 25845130]
35. Graham JL, Andreasen AR. Marketing social change: changing behavior to promote health, social development, and the environment. J Mark Res 1997 May;34(2):294. [doi: 10.2307/3151867]
36. Nashville General Hospital. Congregational Health and Education Network. URL: https://nashvillegeneral.org/resources/chen [accessed 2022-03-11]
37. Creswell J. Qualitative Inquiry and Research Design: Choosing among Five Traditions. Newbury Park, CA: SAGE Publications; 1998.
38. 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;87(1):66-73. [doi: 10.1097/acm.0b013e31823ab7d2]
39. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
40. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
41. Azungah T. Qualitative research: deductive and inductive approaches to data analysis. Qual Res J 2018 Oct 31;18(4):383-400. [doi: 10.1108/qrj-d-18-00035]
42. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2016 Nov 29;5(1):80-92. [doi: 10.1177/160940690600500107]
43. LAWSHE CH. A quantitative approach to content validity. Pers Psychol 1975 Dec;28(4):563-575. [doi: 10.1111/j.1744-6570.1975.tb01393.x]
44. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/ [accessed 2022-07-07]
45. Flesch R. A new readability yardstick. J Appl Psychol 1948 Jun;32(3):221-233. [doi: 10.1037/h0057532] [Medline: 18867058]
46. Kincaid JP, Fishburned RP, Rogers RL, Chissom BS. Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel. URL: http://stars.library.ucf.edu/istlibrary/56. [accessed 2022-07-07]
47. McLaughlin G. SMOG grading: a new readability formula. J Read 1969;12(8):639-646.
48. Readability Formulas. Free Readability Tools to Check for Reading Levels, Reading Assessment, and Reading Grade Levels. URL: https://readabilityformulas.com/ [accessed 2022-07-07]
49. Gilmore B, Ndejjo R, Tchetchia A, de Claro V, Mago E, Diallo AA, et al. Community engagement for COVID-19 prevention and control: a rapid evidence synthesis. BMJ Glob Health 2020 Oct;5(10):e003188 [FREE Full text] [doi: 10.1136/bmjgh-2020-003188] [Medline: 33051285]
50. Vastine A, Gittelsohn J, Ethelbah B, Anliker J, Caballero B. Formative research and stakeholder participation in intervention development. Am J Health Behav 2005 Jan 01;29(1):57-69. [doi: 10.5993/ajhb.29.1.5] [Medline: 15604050]
51. Cunningham-Erves J, Barajas C, Mayo-Gamble TL, McAfee CR, Hull PC, Sanderson M, et al. Formative research to design a culturally-appropriate cancer clinical trial education program to increase participation of African American and Latino communities. BMC Public Health 2020 Jun 03;20(1):840 [FREE Full text] [doi: 10.1186/s12889-020-08939-4] [Medline: 32493245]
52. Boyer A, Fair A, Joosten Y, Dolor R, Williams N, Sherden L. A multilevel approach to stakeholder engagement in the formulation of a clinical data research network. Med Care 2018;56(10 Suppl 1):S22-S26. [doi: 10.1097/mlr.0000000000000778]
53. Wallerstein N, Muhammad M, Sanchez-Youngman S, Rodriguez Espinosa P, Avila M, Baker EA, et al. Power dynamics in community-based participatory research: a multiple-case study analysis of partnering contexts, histories, and practices. Health Educ Behav 2019 Oct 24;46(1_suppl):19S-32S. [doi: 10.1177/1090198119852998] [Medline: 31549557]
54. Helitzer D, Hollis C, Cotner J, Oestreicher N. Health literacy demands of written health information materials: an assessment of cervical cancer prevention materials. Cancer Control 2009 Jan 01;16(1):70-78 [FREE Full text] [doi: 10.1177/107327480901600111] [Medline: 19078933]

55. Brown KM, Lindenberger JH, Bryant CA. Using pretesting to ensure your messages and materials are on strategy. Health Promot Pract 2008 Apr 01;9(2):116-122. [doi: 10.1177/1524839908315134] [Medline: 18340086]
56. Sudore R, Schillinger D. Interventions to improve care for patients with limited health literacy. J Clin Outcomes Manag 2009 Jan 01;16(1):20-29 [FREE Full text] [Medline: 20046798]

Abbreviations

CAP: community advisory panel
CHEN: Congregational Health and Education Network
EUA: Emergency Use Authorization
GBS: Guillain-Barré syndrome
HBM: Health Belief Model
mRNA: messenger RNA
SMOG: Simple Measure of Gobbledygook
TRA: Theory of Reasoned Action

Edited by A Mavragani; submitted 15.04.22; peer-reviewed by M Gwadz, N Martin; comments to author 26.05.22; revised version received 23.06.22; accepted 27.06.22; published 28.07.22.

Please cite as:
Cunningham-Erves J, Brandt HM, Sanderson M, Clarkson K, Lee O, Schlundt D, Bonnet K, Davis J
Development of a Theory-Based, Culturally Appropriate Message Library for Use in Interventions to Promote COVID-19 Vaccination Among African Americans: Formative Research
JMIR Form Res 2022;6(7):e38781
URL: https://formative.jmir.org/2022/7/e38781
doi:10.2196/38781
PMID:35781223

Original Paper

Health-Related Quality of Life Among Members Using an On Demand Behavioral Health Platform: Pilot Observational Study

Emily Shih[1], PhD; Brandon S Aylward[1], PhD; Sarah Kunkle[1], MPH; Grant Graziani[1], PhD

Ginger, San Francisco, CA, United States

Corresponding Author:
Emily Shih, PhD
Ginger
116 New Montgomery St. Suite 500
San Francisco, CA, 94105
United States
Phone: 1 6282079974
Email: eshih@ginger.io

Abstract

Background: Despite the well-known adverse health conditions and negative economic outcomes associated with mental health problems, accessing treatment is difficult due to reasons such as availability and cost. As a solution, digital mental health services have flooded the industry, and new studies are quickly emerging that support their potential as an accessible and cost-effective way to improve mental health outcomes. However, many mental health platforms typically use clinical tools such as the Patient Health Questionnaire-9 (PHQ-9) or General Anxiety Disorder-7 (GAD-7). Yet, many individuals that seek out care do not have clinical symptomatology and thus, traditional clinical measures may not adequately capture symptom improvement in general well-being. As an alternative, this study used the health-related quality of life (HRQoL) tool from the Centers for Disease Control and Prevention "Healthy Days" measure. This subjective measure of well-being is an effective way to capture HRQoL and might be better suited as an outcome measure for treatments that include both clinical and subclinical individuals.

Objective: The purpose of this study was to describe changes in HRQoL in clinical and subclinical members assessing virtual care and to examine the association between text-based behavioral coaching and virtual clinical sessions with changes in HRQoL.

Methods: A total of 288 members completed the 4-item HRQoL measure at baseline and at 1 month following use of the Ginger on demand behavioral health platform. Baseline anxiety and depression levels were collected using the GAD-7 and PHQ-9, respectively.

Results: Members completed on average 1.92 (SD 2.16) coaching sessions and 0.91 (SD 1.37) clinical sessions during the assessment month. Paired samples *t* tests revealed significant reductions in the average number of unhealthy mental health days between baseline (mean 16, SD 8.77 days) and follow-up (mean 13.2, SD 9.02 days; t_{287}=5.73; *P*<.001), and in the average number of days adversely impacted (mean$_{baseline}$ 10.9, mean$_{follow\text{-}up}$ 8.19; t_{287}=6.26; *P*<.001). Both subclinical members (t_{103}=3.04; *P*=.003) and clinical members (t_{183}=5.5; *P*<.001) demonstrated significant improvements through reductions in adversely impacted days over a month. Clinical members also demonstrated significant improvements through reductions in unhealthy mental health days (t_{183}=5.82; *P*<.001). Finally, member engagement with virtual clinical sessions significantly predicted changes in unhealthy mental health days (*B*=–0.96; *P*=.04).

Conclusions: To our knowledge, this study is one of the first to use the HRQoL measure as an outcome in an evaluation of a digital behavioral health platform. Using real-world longitudinal data, our preliminary yet promising results show that short-term engagement with virtual care can be an effective means to improve HRQoL for members with subclinical and clinical symptoms. Further follow-up of reported HRQoL over several months is needed.

(JMIR Form Res 2022;6(7):e35352) doi:10.2196/35352

KEYWORDS

behavioral coaching; mental health; telehealth; Healthy Days; clinical care; behavior; coach; quality of life; platform; tool; pilot study; observational; health-related quality of life; virtual care; association; text-based; outcome; evaluation

Introduction

Nearly 1 in 5 adults in the United States (51.5 million people) experience mental health issues [1]. The World Health Organization estimates that anxiety and depression alone cost the global economy US $1 trillion dollars each year in lost productivity, absenteeism, and medical costs [2]. Mental health issues have been exacerbated with the recent COVID-19 pandemic and underscore a critical moment of global need [3,4]. A recent meta-analysis found the global prevalence of diagnosable anxiety and depression during the pandemic was 27% and 28%, respectively [5]. Even among those with subclinical symptoms, nearly half of adults in the United States have reported symptoms of anxiety or depression during this time [6]. Timely intervention for those with subclinical symptoms is just as important to prevent development of more serious symptoms requiring more costly treatment.

Despite the well-known adverse health conditions and negative economic outcomes, accessing treatment for common mental health problems is difficult [7]. The demand for mental health services has outpaced the availability of qualified mental health professionals. A recent survey found that 1 in 4 individuals with depression or anxiety lack access to care or have unmet mental health needs [8]. In addition, long wait lists, high out-of-pocket expenses, and transportation burdens all continue to serve as barriers to receipt of effective services [9,10]. There is a growing need for scalable mental health solutions that increase both the availability of professionals and access to care for common mental health conditions. This is particularly important with the recent increase of mental health issues during the pandemic. Digital mental health services have flooded the industry, and new studies are emerging that support their potential to serve as cost-effective ways to manage anxiety and depression [11,12]. This type of support can even be beneficial for individuals who may be at risk for but do not yet experience clinically significant symptoms [13].

Many mental health platforms typically use clinical tools such as the Patient Health Questionnaire-9 (PHQ-9) or General Anxiety Disorder-7 (GAD-7) for assessing initial and treatment outcomes of depressive and anxiety symptoms, respectively. As behavioral coaching focuses on goal-oriented behavior and typically targets those with subclinical symptomatology, traditional clinical measures may not adequately capture symptom improvement in general mental health and well-being. State and federal health agencies have supported the population surveillance of health-related quality of life (HRQoL), which is a multidimensional concept that examines overall health related to perceived physical and mental health as well as daily functioning [14,15]. One common HRQoL tool is the Centers for Disease Control and Prevention (CDC) “Healthy Days” measure that asks about self-rated general health, physical health, mental health, and activity limitations over the past 30 days. This subjective measure of well-being is an effective way to capture HRQoL and might be better suited as an outcome measure for treatments that include both clinical individuals and individuals with symptoms not meeting clinical thresholds [16]. Yet, few studies have used this measure when evaluating digital behavioral health platforms. Financially, Humana found that the cost of each reported unhealthy day is equivalent to 10 hospital admissions per thousand patients, with a potential increase of US $15.64 per member per month in medical costs for each unhealthy day [17]. This highlights the potential long-term savings that could result from interventions targeting individual HRQoL. A previous health coaching study has already demonstrated significant reductions in reported unhealthy days among participants [18].

The purpose of this study was to examine self-reported HRQoL among members using an on demand digital health platform and the association of short-term text-based behavioral health coaching and virtual clinical sessions with healthy days over time. To that end, the study will describe baseline characteristics of members in terms of reported unhealthy days and changes over 1 month, describe changes in unhealthy days as a function of baseline anxiety and depressive symptoms, and examine the association between member engagement and changes in unhealthy days.

Methods

Participants

Participants were members who had access to the Ginger on demand behavioral health platform as part of their employer or health plan benefits. Internal clinical protocols include exclusionary criteria where self-directed telehealth is likely not appropriate and where more specialized and urgent psychiatric services are required (eg, active suicide ideation or active high-risk self-harm behavior; see Kunkle et al [19] for exhaustive list). This study included Ginger members 18 years or older who completed the baseline Healthy Days measure between November 2020 to November 2021 and who first accessed care within 1 month of completing their Healthy Days baseline.

Procedures

The Ginger platform provides members with access to virtual behavioral health coaching, teletherapy, telepsychiatry, and self-guided content and assessments, primarily via a mobile app platform. After downloading the mobile app, members can start texting with a behavioral health coach within minutes of requesting to connect. Ginger coaches are full-time employees who have an advanced degree in a field related to mental health or have accredited coach certification. While many members are solely engaged with text-based coaching services, some will request or require escalation to clinical services (teletherapy or telepsychiatry) depending on preference or clinical severity. When members are escalated to therapy or psychiatry, they may continue working with a coach provided they also seek additional specialized care concurrently. Additional detail regarding Ginger can be found in prior publications [19,20].

The Healthy Days measure was administered to members 4 times across the span of 4 months (once per month). Data were collected externally using the Survey Monkey platform. Only responses from survey items pertaining to the number of unhealthy mental health days and impacted days were of focus for this study. The PHQ-9 and GAD-7 were typically completed

at intake within 1 month of the Healthy Days baseline assessment.

Measures

The CDC Healthy Days measure contains four items: (1) "Would you say that in general your health is excellent, very good, good, fair, or poor?" (2) "Now thinking about your physical health, which includes physical illness and injury, for how many days during the past 30 days was your physical health not good?" (3) "Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?" and (4) "During the past 30 days, for about how many days did poor physical or mental health keep you from doing your usual activities, such as self-care, work, or recreation?" (referred to here as impacted days). For this study a change variable was calculated by subtracting reported unhealthy scores from time 1 from scores from time 2, where positive values indicate an increase in unhealthy days, whereas negative values indicate a reduction in unhealthy days.

The PHQ-9 is a 9-item self-report questionnaire that assesses the frequency and severity of depression symptomatology within the previous 2 weeks. Each of the 9 items is based on the *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition; *DSM-IV*) criteria for major depressive disorder and are scored on a 0 (not at all) to 3 (nearly every day) scale. Items include "Little interest or pleasure in doing things" and "Feeling down, depressed, or hopeless." Total scores can range from 0 to 27 with higher scores indicating more depressive symptoms. A score of 10 was used as the clinical threshold [21].

The GAD-7 is a valid brief self-report tool to assess the frequency and severity of anxious thoughts and behaviors over the past 2 weeks. Each of the 7 items are based on the *DSM-IV* diagnostic criteria for generalized anxiety disorder and are scored on a 0 (not at all) to 3 (nearly every day) scale, with total scores ranging from 0 to 21. Items include "Feeling nervous, anxious, or on edge" and "Not being able to stop or control worrying." Consistent with existing literature, a score of 10 was used as the clinical threshold [22].

Member engagement with Ginger services was quantified as the number of coaching and clinical sessions. Coaching sessions were operationalized as the number of unique days where both members and coaches each sent at least 5 text messages. Ginger coaching is an on demand text-based service, and the operationalization of a "text-based coach session" has not been predetermined in the literature. As such, our threshold was decided based upon internal work that highlighted approximately 5 texts each way as the number of text messages needed to capture a productive conversation between members and their coaches. Clinical sessions were operationalized as the number of completed video sessions with a clinician.

Statistical Analysis

Analyses were conducted using RStudio (version 1.4.1717; RStudio, PBC). Data were first screened for outliers and normality. Descriptive statistics were used to describe baseline member characteristics. For changes in reported unhealthy days, paired sample *t* tests were used. Next, members were divided into groups as a function of clinical thresholds using the PHQ-9 and GAD-7 scores at intake (ie, clinical vs subclinical). Additional paired sample *t* tests were performed to evaluate member differences in responses between time 1 and time 2 for clinical and subclinical groups separately. A Benjamini-Hochberg correction was used to adjust for multiple comparisons [23]. Finally, scatterplots suggested a linear trend between member engagement and changes in unhealthy days. As such, multiple linear regressions were performed to examine the association of member engagement (ie, coaching and clinical sessions) with changes in the number of unhealthy days. Baseline Healthy Days scores and the number of prior engagement levels were entered as covariates. All continuous variables were standardized for interpretability.

Ethics Approval

This is a secondary analysis of pre-existing deidentified data. The authors do not have access to participant identifying information and do not intend to recontact participants. Ginger's research protocols and supporting policies have been reviewed and approved by Advarra's institutional review board (Pro00046797) in accordance with the US Department of Health and Human Services regulations at 45 CFR 46.

Results

Descriptive Statistics

A total of 1496 members completed the Healthy Days measure at time 1 (intake), 351 (23.5%) members at time 2 (~30 days following intake; mean 31.9, SD 1.48 days), 114 members at time 3 (~60 days following intake), and 37 members at time 4 (~90 days following intake). The current analyses examined only members who had completed surveys at both time 1 (intake) and at time 2 (N=288). Data were missing at random for all primary outcome variables ($t>-0.70$ and $t<1.54$; $P>.12$). Potential reasons for earlier drop-offs that should be taken into consideration when interpreting our results include members having achieved their coaching goals, members no longer interested in care, and members engaged at a monthly cadence and returned after the study evaluation month was finished. Demographic information about members was provided by employers but contained missing data. Of members in the analytical sample, 82 (28.5%) members were between the ages of 18-34 years, 96 (33.3%) members were 35 years of age or older, and 110 (38.2%) members did not have age reported. Regarding gender identity, 125 (43.4%) members identified as female, 33 (11.5%) as male, 14 (4.9%) as other, and 116 (40.3%) did not have gender reported.

Descriptive statistics for the primary variables are presented in Table 1. Members, on average, completed 1.92 (SD 2.16, range 0-12) coaching sessions and 0.91 (SD 1.37, range 0-5) video sessions with a clinician within a single month. A total of 179 (62.2%) members engaged exclusively with text-based coaching (no clinical sessions). Subclinical depression and anxiety levels were reported in 104 (36.1%) members, whereas 184 (63.9%) members reported clinical levels of depression or anxiety. Of members in the analytical sample, 71% (n=205) at time 1 and 77% (n=223) of members at time 2 reported feeling "good" or better in response to the question "Would you say that in general

your health is excellent, very good, good, fair, or poor?" (includes members who reported feeling "very good" and "excellent"). Bivariate correlations among the primary variables are presented in Figure 1. Of note, the number of unhealthy mental health days was positively correlated with the number of impacted health days at each respective time point (r=0.62 at time 1, r=0.65 at time 2; P<.001).

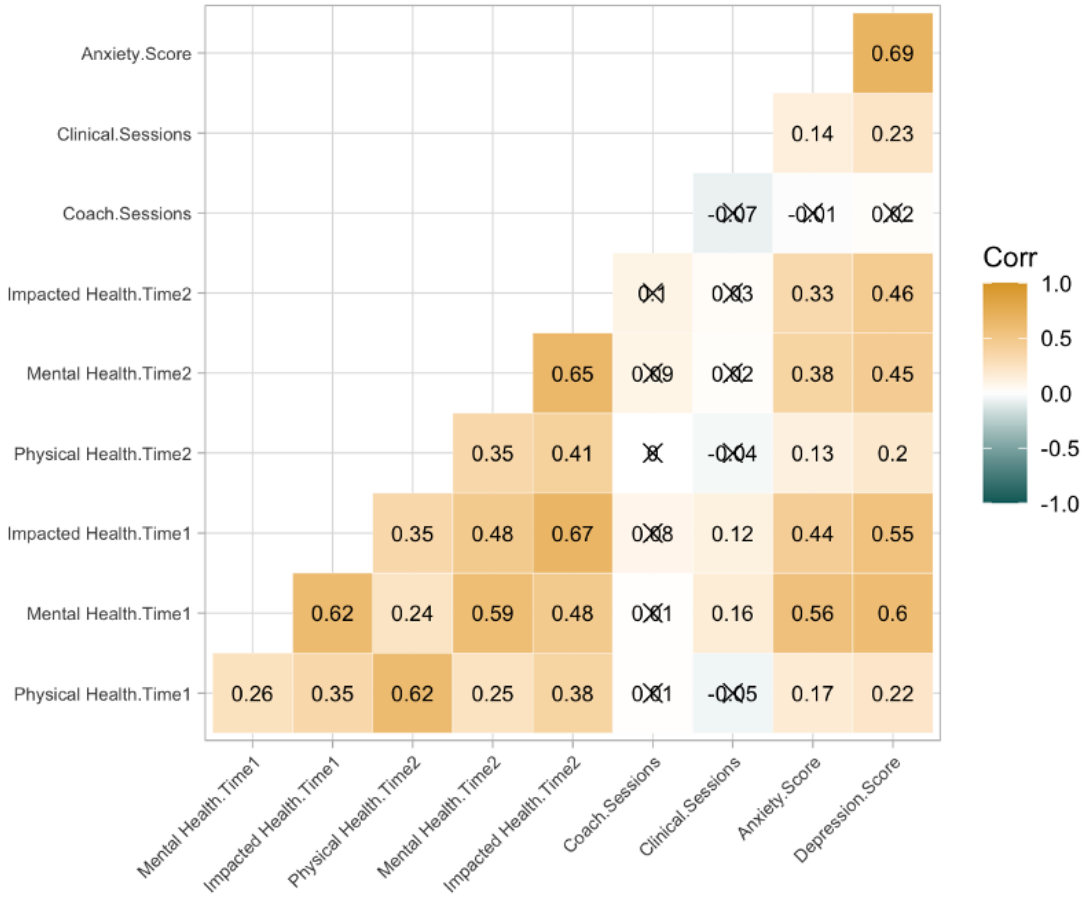
Table 1. Descriptive statistics among primary variables.

	Values, mean (SD)	Min	Max
Physical health (time 1)	5.1 (7.8)	0	30
Mental health (time 1)	16.0 (8.8)	0	30
Impacted health (time 1)	10.9 (9.6)	0	30
Physical health (time 2)	5.6 (8.3)	0	30
Mental health (time 2)	13.2 (9.0)	0	30
Impacted health (time 2)	8.2 (8.5)	0	30
Coaching sessions	1.9 (2.2)	0	12
Clinical sessions	0.9 (1.4)	0	5
Depression score (PHQ-9[a])	11.3 (6.1)	1	27
Anxiety Score (GAD-7[b])	9.8 (5.7)	0	21

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

[b]GAD-7: General Anxiety Disorder-7.

Figure 1. Correlations among primary variables. Note: Insignificant correlations where P>.05 are marked. Corr: correlation.

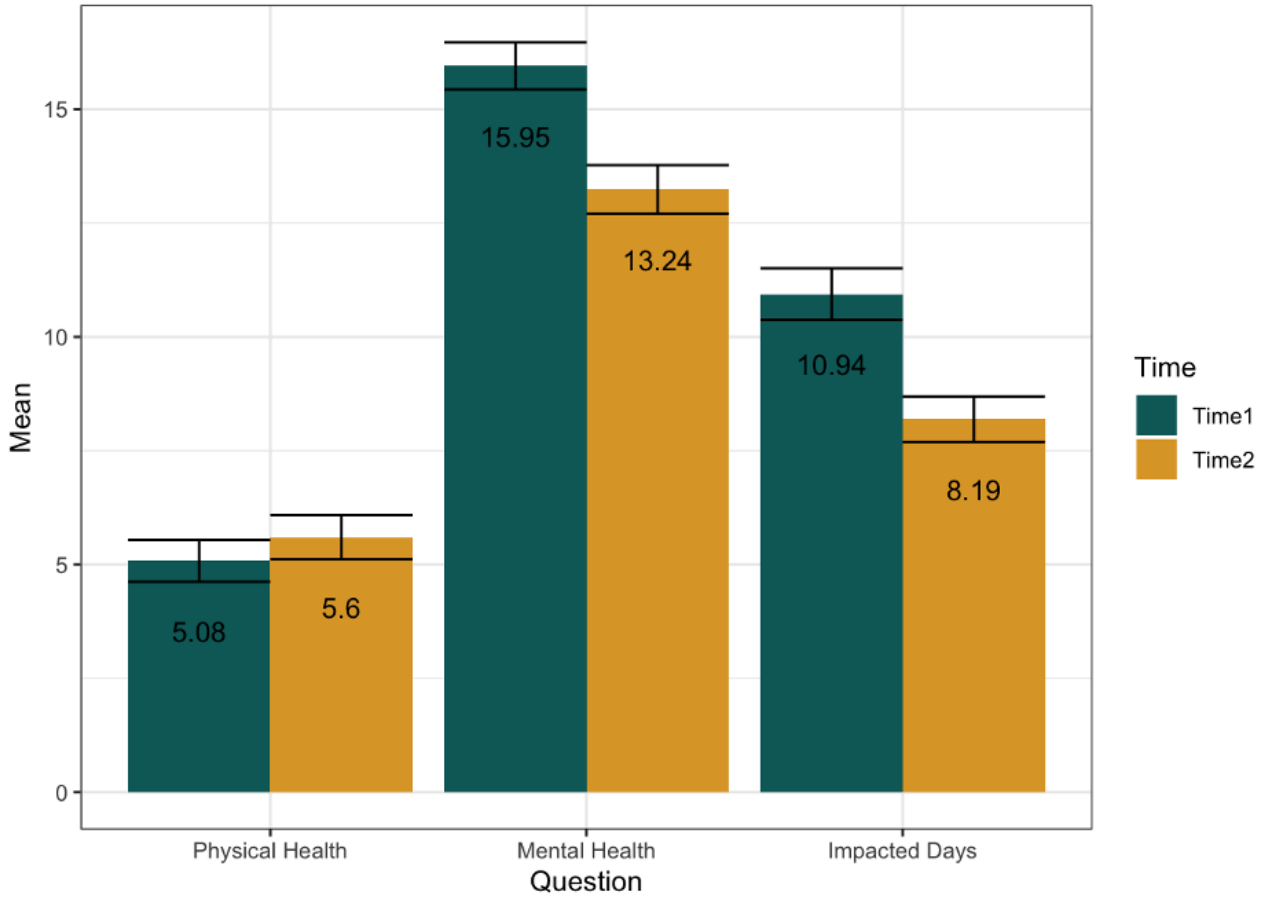


Pre-Post Changes in Reported Unhealthy Days

Members reported on average nearly 3 fewer unhealthy mental health days (mean –2.71, SD 8.03) between baseline and 1 month later. Of the analytical sample, 61% (n=175) of members reported an improvement in unhealthy mental health days, whereas 39% (n=113) reported no improvement or an increase in unhealthy mental health days. Paired sample t tests were

performed to evaluate differences in member Healthy Days responses at time 1 compared to time 2 (Figure 2) across all continuous items. Results showed no significant improvements in unhealthy physical health days between time 1 (mean 5.08, SD 7.78 days) and time 2 (mean 5.60, SD 8.25 days; t_{287}=–1.25; *P*=.21). However, results showed significant improvements in unhealthy mental health days between time 1 (mean 16, SD 8.77 days) and time 2 (mean 13.2, SD 9.02 days; t_{287}=5.73; *P*<.001), as well as significant improvements in adversely impacted days between time 1 (mean 10.9, SD 9.60 days) and time 2 (mean 8.19, SD 8.51 days; t_{287}=6.26; *P*<.001). Given Ginger is a mental health platform and significant changes were only observed for unhealthy mental health days and adversely impacted days, these two outcomes were explored in subsequent analyses.

Figure 2. Display of means across the items from the Healthy Days measure at time 1 and time 2 (N=288).



Comparison of Change in Healthy Days Between Clinical and Subclinical Members

Subclinical members showed trending reductions in reported unhealthy mental health days between time 1 (mean 9.92, SD 6.78 days) and time 2 (mean 8.44, SD 7.83 days; t_{103}=1.87; *P*=.06, adjusted *P*=.06). Clinical members also showed reductions in reported unhealthy mental health days between time 1 (mean 19.4, SD 7.91 days) and time 2 (mean 16.0, SD 8.52 days; t_{183}=5.82; *P*<.001; adjusted *P*<.001).

Similarly, subclinical members showed significant reductions in reported impacted days at time 1 (mean 5.15, SD 6.64 days) compared to time 2 (mean 3.47, SD 5.3 days; t_{103}=3.04; *P*=.003, adjusted *P*=.003). Clinical members also showed significant reductions in reported impacted days at time 1 (mean 14.2, SD 9.48 days) compared to time 2 (mean 10.9, SD 8.83 days; t_{183}=5.50; *P*<.001, adjusted *P*=.001).

Member Engagement on Changes in Reported Unhealthy Mental Health Days

The linear regression model predicting changes in reported unhealthy mental health days was significant ($F_{5,282}$=14.6; *P*<.001) and accounted for 21% of the variance. No significant main effects of coaching sessions (*B*=0.61; *P*=.19) were observed. However, there was a significant main effect of clinical sessions (*B*=–0.96; *P*=.04), where more clinical sessions was associated with a decrease in unhealthy mental health days. The model predicting changes in adversely impacted days was also significant ($F_{5,282}$=22.2; *P*<.001) and accounted for 28% of the variance. No significant main effects of coaching sessions (*B*=0.43; *P*=.30) or clinical sessions (*B*=–0.40; *P*=.33) were observed. Coefficients for both models are presented in Table 2.

Table 2. Summary of regression coefficients (N=288).

	Beta (SE)	*P* value
Model 1: Changes in the number of unhealthy mental health days		
(Intercept)	–2.71 (0.43)	<.001
Unhealthy mental health days (baseline)	–3.32 (0.43)	<.001
Prior coaching sessions	0.39 (0.47)	.42
Prior clinical sessions	0.59 (0.47)	.21
Clinical sessions	–0.96 (0.47)	.04
Coaching sessions	0.61 (0.47)	.19
Model 2: Changes in the number of adversely impacted days		
(Intercept)	–2.75 (0.38)	<.001
Unhealthy impacted days (baseline)	–3.87 (0.38)	<.001
Prior coaching sessions	–0.30 (0.42)	.48
Prior clinical sessions	0.10 (0.41)	.80
Clinical sessions	–0.40 (0.41)	.33
Coaching sessions	0.43 (0.41)	.30

Discussion

Principal Findings

This study evaluated the real-world association between digital care utilization in members with both subclinical and clinical symptoms of anxiety or depression. HRQoL at baseline suggested that members were, on average, demonstrating "frequent distress" and reporting more *unhealthy* mental health days than *healthy* mental health days (mean 16, SD 8.77 days; 53% of the month). The CDC defines having ≥14 unhealthy mental health days as "frequent distress [24]." Of note, our results also observed a relatively high number of unhealthy mental health days (mean 9.92, SD 6.78 days; 33% of the month) for subclinical members at baseline, highlighting the need for care for those that might not traditionally be recommended for clinical services (eg, individuals who might not have exceeded clinical thresholds using traditional PHQ-9 and GAD-7 assessment surveys). Bivariate correlations revealed a positive association between unhealthy mental health days and adversely impacted days, underscoring the relationship between mental health and daily functioning [25,26]. Overall, members evidenced significant improvements in reported unhealthy mental health days and adversely impacted days over the month. Furthermore, improvements in reported adversely impacted days were significant for both subclinical members and clinical members, and improvements in reported unhealthy mental health days were significant for clinical members. Our results also found that clinical sessions, but not coaching sessions, predicted changes in reported unhealthy mental health days over the month. Taken together, this study offers preliminary descriptives on a valuable but less commonly used outcome measure, specifically in a traditionally understudied but increasing population of individuals seeking out virtual care. The study further supports how virtual care is a promising strategy to meet the growing demand of mental health services.

Not all individuals seeking out care exceeded industry clinical thresholds. Thus, additional outcome measures, such as the Healthy Days measure, are needed to evaluate the effects of digital mental health care beyond clinically focused measures (eg, PHQ-9 and GAD-7). To our knowledge, we are one of the first to use the Healthy Days measure within this population (ie, individuals seeking out virtual mental health care). Overall, members reported a reduction of 2.71 unhealthy mental health days. Extrapolating from the Humana data [16], this would be equivalent to a decrease of 27.1 hospital admissions per thousand patients and a potential cost savings of US $42.38 per member per month. Thus, virtual mental health care can be seen as a low-intensity approach to achieve better health outcomes at lower cost [12,13].

Our results found a significant association between the reduction in the number of reported unhealthy mental health days and member engagement with clinical sessions, but not with coaching sessions. Coaching, and even more so text-based coaching, differs fundamentally in their objectives and practices compared to clinical care [27,28]. Little is understood regarding the effects of text-based coaching on mental health outcomes. Our findings suggest that the amount of care needed to drive member improvement might vary between text-based coaching and clinical practices [29]. It is possible that additional time/sessions might be needed for coaching goals to be formed, implemented, and subsequently have an impact on behavioral change via a text-based medium [28-30]. Future studies should extend the follow-up window when evaluating coaching sessions and assess alternate trajectories of improvement in mental health (eg, nonlinear).

Limitations

There are several limitations to consider. One limitation is the potential for bias in our estimates and the increased likelihood that our results may not generalize to all individuals who engage with teletherapy. Furthermore, our cohort design did not have

a comparison group or random assignment to the treatment intervention. Thus, our ability to draw causal inferences is limited and improvements in reported unhealthy days could simply be due to a passage of time; however, we were able to demonstrate significant changes in members with both subclinical and clinical symptoms using real-world longitudinal data. Even though data were missing at random and may not bias results, future studies should implement procedures (eg, incentives) to encourage and capture more complete follow-up data. Future studies can also examine obstacles and facilitators for engagement in teletherapy. The study was also limited to available self-reported outcome data, and there was a large amount of attrition in members reporting unhealthy days over time. This could be due to most members not experiencing clinically meaningful baseline symptomatology and potentially quick improvements in functioning. It is also possible that because the survey was administered outside of the Ginger platform (ie, Survey Monkey), the additional step of completing the measure might have been an added time burden. However, this approach allowed us to pilot and demonstrate the real-world attrition rate when using external data collection platforms.

Conclusions

To our knowledge, this study is one of the first to use the HRQoL measure as a primary outcome in an evaluation of a digital behavioral health platform. Using real-world longitudinal data, our preliminary yet promising results show that short-term engagement with virtual care can be an effective means to improve HRQoL for members with subclinical and clinical symptoms. Virtual care represents a scalable and well-suited approach to meet the growing need for mental health services that has outpaced the in-person availability of clinical mental health professionals. Future studies should examine the long-term impact of text-based coaching and clinical support on HRQoL.

Acknowledgments

We are grateful for the participation of the members who completed the survey, as well as the hard work and dedication of Ginger coaches and clinicians.

Conflicts of Interest

All authors are paid employees of Ginger.

References

1. Depression. World Health Organization. 2020. URL: https://www.who.int/news-room/fact-sheets/detail/depression [accessed 2021-09-16]
2. Mental health in the workplace. World Health Organization. URL: https://www.who.int/teams/mental-health-and-substance-use/promotion-prevention/mental-health-in-the-workplace [accessed 2021-09-06]
3. Holman EA, Thompson RR, Garfin DR, Silver RC. The unfolding COVID-19 pandemic: a probability-based, nationally representative study of mental health in the United States. Sci Adv 2020 Oct;6(42):eabd5390 [FREE Full text] [doi: 10.1126/sciadv.abd5390] [Medline: 32948511]
4. Yao H, Chen J, Xu Y. Patients with mental health disorders in the COVID-19 epidemic. Lancet Psychiatry 2020 Apr;7(4):e21 [FREE Full text] [doi: 10.1016/S2215-0366(20)30090-0] [Medline: 32199510]
5. Nochaiwong S, Ruengorn C, Thavorn K, Hutton B, Awiphan R, Phosuya C, et al. Global prevalence of mental health issues among the general population during the coronavirus disease-2019 pandemic: a systematic review and meta-analysis. Sci Rep 2021 May 13;11(1):10173. [doi: 10.1038/s41598-021-89700-8] [Medline: 33986414]
6. Panchal N, Kamal R, Cox C, Garfield R. The implications of COVID-19 for mental health and substance use. KFF. URL: https://www.kff.org/coronavirus-covid-19/issue-brief/the-implications-of-covid-19-for-mental-health-and-substance-use/ [accessed 2021-09-30]
7. 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]
8. Nagata JM, Ganson KT, Bonin SL, Twadell KL, Garcia ME, Langrock O, et al. Prevalence and sociodemographic correlates of unmet need for mental health counseling among adults during the COVID-19 pandemic. Psychiatr Serv 2022 Feb 01;73(2):206-209. [doi: 10.1176/appi.ps.202100111] [Medline: 34189929]
9. Bishop TF, Seirup JK, Pincus HA, Ross JS. Population Of US practicing psychiatrists declined, 2003-13, which may help explain poor access to mental health care. Health Aff (Millwood) 2016 Jul 01;35(7):1271-1277. [doi: 10.1377/hlthaff.2015.1643] [Medline: 27385244]
10. Weil TP. Insufficient dollars and qualified personnel to meet United States mental health needs. J Nerv Ment Dis 2015 Apr;203(4):233-240. [doi: 10.1097/NMD.0000000000000271] [Medline: 25816044]
11. 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]

12. Firth J, Torous J, Carney R, Newby J, Cosco TD, Christensen H, et al. Digital technologies in the treatment of anxiety: recent innovations and future directions. Curr Psychiatry Rep 2018 May 19;20(6):44 [FREE Full text] [doi: 10.1007/s11920-018-0910-2] [Medline: 29779065]
13. Bendtsen M, Müssener U, Linderoth C, Thomas K. A mobile health intervention for mental health promotion among university students: randomized controlled trial. JMIR Mhealth Uhealth 2020 Mar 20;8(3):e17208 [FREE Full text] [doi: 10.2196/17208] [Medline: 32196462]
14. Moriarty DG, Zack MM, Kobau R. The Centers for Disease Control and Prevention's Healthy Days Measures - population tracking of perceived physical and mental health over time. Health Qual Life Outcomes 2003 Sep 02;1:37 [FREE Full text] [doi: 10.1186/1477-7525-1-37] [Medline: 14498988]
15. Zahran HS, Kobau R, Moriarty DG, Zack MM, Holt J, Donehoo R, Centers for Disease Control and Prevention (CDC). Health-related quality of life surveillance--United States, 1993-2002. MMWR Surveill Summ 2005 Oct 28;54(4):1-35 [FREE Full text] [Medline: 16251867]
16. Slabaugh SL, Shah M, Zack M, Happe L, Cordier T, Havens E, et al. Leveraging health-related quality of life in population health management: the case for Healthy Days. Popul Health Manag 2017 Feb;20(1):13-22 [FREE Full text] [doi: 10.1089/pop.2015.0162] [Medline: 27031869]
17. Healthy Days. Humana: Population Health. URL: https://populationhealth.humana.com/healthy-days/ [accessed 2021-10-01]
18. Cole S, Zbikowski SM, Renda A, Wallace A, Dobbins JM, Bogard M. Examining changes in Healthy Days after health coaching. Am J Health Promot 2019 Jun;33(5):774-777. [doi: 10.1177/0890117118816286] [Medline: 30497272]
19. Kunkle S, Yip M, Ξ W, Hunt J. Evaluation of an on-demand mental health system for depression symptoms: retrospective observational study. J Med Internet Res 2020 Jun 18;22(6):e17902 [FREE Full text] [doi: 10.2196/17902] [Medline: 32554387]
20. Kunkle S, Yip M, Hunt J, Ξ W, Udall D, Arean P, et al. Association between care utilization and anxiety outcomes in an on-demand mental health system: retrospective observational study. JMIR Form Res 2021 Jan 26;5(1):e24662 [FREE Full text] [doi: 10.2196/24662] [Medline: 33496679]
21. 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]
22. 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]
23. Thissen D, Steinberg L, Kuang D. Quick and easy implementation of the Benjamini-Hochberg procedure for controlling the false positive rate in multiple comparisons. J Educ Behav Statistics 2016 Nov 23;27(1):77-83. [doi: 10.3102/10769986027001077]
24. Frequent mental distress. America's Health Rankings. URL: https://www.americashealthrankings.org/explore/annual/measure/mental_distress/state/ALL [accessed 2021-09-15]
25. Erickson SR, Guthrie S, Vanetten-Lee M, Himle J, Hoffman J, Santos SF, et al. Severity of anxiety and work-related outcomes of patients with anxiety disorders. Depress Anxiety 2009;26(12):1165-1171. [doi: 10.1002/da.20624] [Medline: 19842165]
26. Jain G, Roy A, Harikrishnan V, Yu S, Dabbous O, Lawrence C. Patient-reported depression severity measured by the PHQ-9 and impact on work productivity: results from a survey of full-time employees in the United States. J Occup Environ Med 2013 Mar;55(3):252-258. [doi: 10.1097/JOM.0b013e31828349c9] [Medline: 23439268]
27. Lattie EG, Graham AK, Hadjistavropoulos HD, Dear BF, Titov N, Mohr DC. Guidance on defining the scope and development of text-based coaching protocols for digital mental health interventions. Digit Health 2019;5:2055207619896145 [FREE Full text] [doi: 10.1177/2055207619896145] [Medline: 31897306]
28. Jordan M, Livingstone JB. Coaching vs psychotherapy in health and wellness: overlap, dissimilarities, and the potential for collaboration. Glob Adv Health Med 2013 Jul;2(4):20-27 [FREE Full text] [doi: 10.7453/gahmj.2013.036] [Medline: 24416682]
29. Uher R. Genes, environment, and individual differences in responding to treatment for depression. Harv Rev Psychiatry 2011;19(3):109-124. [doi: 10.3109/10673229.2011.586551] [Medline: 21631158]
30. ICF Code of Ethics. International Couching Federation. URL: https://coachingfederation.org/ethics/code-of-ethics [accessed 2021-10-08]

Abbreviations

CDC: Centers for Disease Control and Prevention
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition)
GAD-7: General Anxiety Disorder-7
HRQoL: health-related quality of life
PHQ-9: Patient Health Questionnaire-9

Edited by A Mavragani; submitted 01.12.21; peer-reviewed by A Solomonides; comments to author 03.01.22; revised version received 21.01.22; accepted 14.03.22; published 08.07.22.

Please cite as:
Shih E, Aylward BS, Kunkle S, Graziani G
Health-Related Quality of Life Among Members Using an On Demand Behavioral Health Platform: Pilot Observational Study
JMIR Form Res 2022;6(7):e35352
URL: https://formative.jmir.org/2022/7/e35352
doi:10.2196/35352
PMID:35802408

Original Paper

The Burden of Cold Agglutinin Disease on Patients' Daily Life: Web-Based Cross-sectional Survey of 50 American Patients

Florence Joly[1], MSc, PharmD; Lisa Anne Schmitt[2], MBA; Patricia Ann McGee Watson[3], MBA; Emilie Pain[4], MSc; Damien Testa[5], PhD

[1]Health Economics And Outcomes Research, Sanofi, Chilly-Mazarin, France

[2]Public Affairs and Patient Advocacy, Sanofi Genzyme, Cambridge, MA, United States

[3]Cold Agglutinin Disease Foundation, Cherry Hill, NJ, United States

[4]Carenity, Paris, France

[5]Moona, Paris, France

Corresponding Author:
Florence Joly, MSc, PharmD
Health Economics And Outcomes Research
Sanofi
1 Avenue Pierre Brossolette
Chilly-Mazarin, 91380
France
Phone: 33 01 60 49 77 77
Email: florence.joly@sanofi.com

Abstract

Background: Cold agglutinin disease (CAD) is a rare disorder, affecting 15% of patients with autoimmune hemolytic anemia. Few studies have assessed CAD symptoms and their impact on daily life, but these studies did not address the patients' perspectives.

Objective: The aims of this study were to increase the knowledge about CAD through a patient-centric survey and to gain a better understanding of the burden of this disease.

Methods: We conducted an internet-based survey in September 2020 among American patients registered on the CAD Unraveled website and members of the Cold Agglutinin Disease Foundation.

Results: A total of 50 respondents were included in this study. Totally, 90% (45/50) of the patients reported having experienced fatigue. Fatigue was mainly reported on a daily basis, and approximately one-third of these patients (13/45, 29%) said that their fatigue was constant throughout the day. It has also been shown that CAD has a great impact on patients' physical well-being, emotional well-being, social life, and household finances. The disease varies over time, with or without symptoms. A total of 88% (44/50) of the patients reported previous episodes of the increased intensity or sensitivity of their CAD symptoms, with a mean of 4.5 (SD 5.4) episodes reported during the past year. More than half of the patients (27/50, 54%) considered their disease to be moderate or severe, and 42% (21/50) of the study group reported that their symptoms had worsened since the time of diagnosis.

Conclusions: Our study has provided new data on CAD symptoms, particularly data on the importance and type of fatigue and the fluctuation of CAD symptoms.

(JMIR Form Res 2022;6(7):e34248) doi:10.2196/34248

KEYWORDS

autoimmune diseases; fatigue; perception; rare diseases; surveys and questionnaires; cold agglutinin disease; cold autoimmune hemolytic anemia

Introduction

Background

Autoimmune hemolytic anemias (AIHAs) are rare and heterogeneous disorders characterized by the destruction of red blood cells by warm or cold antibodies, thereby causing anemia and other related health issues [1-4]. AIHA is classified into three categories: warm, cold, and mixed [1,5,6]. Warm AIHA is characterized by the binding of polyclonal immunoglobulin (often immunoglobulin G) to red blood cell antigens (Rh proteins or glycophorins A-D). This binding is referred to as *warm* in that it occurs at most temperatures but is maximal at 37 °C [4,5]. Cold agglutinin disease (CAD) is the most common form of cold AIHA, accounting for 15 to 25% of AIHA cases [5-7]. CAD is recognized by the presence of immunoglobulin M autoantibodies, also known as cold agglutinins, which are active and cause hemolysis at cold temperatures, usually 3 °C to 4 °C [5,7]. This cold AIHA is composed of CAD, formerly known as primary CAD, and cold agglutinin syndrome, formerly known as secondary CAD [7]. There is no known cause for CAD. Cold agglutinin syndrome is associated with underlying conditions such as infection, malignancy, or immune disease [4,7].

CAD is a rare disease. Berentsen et al [8] suggested variations between cold and warm climates, reporting an incidence of 9 cases per million people per year in North Italy and 2 cases per million people per year in Norway, with a prevalence of 50 per million people in Italy and 200 per million people in Norway. An incidence of 1.8 cases per million person-years was reported based on the Danish national patient register [4]. CAD primarily affects middle-aged to older individuals; however, the disease has been observed in people aged as young as 30 years [7,9-11]. Some studies have indicated that women are slightly more affected by the disease than men [9].

Diagnosis of CAD is established with hemolytic anemia, reticulocytosis, hyperbilirubinemia, elevated lactate dehydrogenase, and positive Coombs test for anti-C3d and classically negative anti– immunoglobulin G [9]. After the test findings suggest CAD, the antibody titer and thermal activity should be determined to prevent overdiagnosis, because most agglutinins are clinically insignificant [9].

Most people with CAD have symptoms of hemolytic anemia such as paleness, shortness of breath, rapid heart rate, fatigue, weakness, dark urine, or pain [6,9,12]. Many people with CAD also experience pain and bluish coloring of the hands and feet (acrocyanosis) or Raynaud disease owing to slow or poor blood circulation [6,9,12]. CAD symptoms can vary throughout the course of a patient's illness, involving fluctuation in CAD severity [7]. Anemia in CAD is often mild (hemoglobin level >10 g/dL) to moderate (hemoglobin level between 8 and 10 g/dL) and, in some cases, fully compensated, but it may be severe (hemoglobin level <8 g/dL) [13].

Importantly, febrile illness, trauma, and surgery can exacerbate hemolytic anemia [7,14], which can increase the risk of thromboembolic events and death [15]. Patients with mild anemia and minimal circulatory symptoms typically do not require specific pharmacological treatment [4,6,13]. These patients are advised to avoid exposure to cold [6,13], to limit the intensity of acrocyanosis and risk of developing trophic disorders. Moreover, any bacterial or viral infection should be treated [6].

Seasonal variation of these symptoms has been reported [16]. Symptoms were reported to be more severe when the patient is exposed to cold temperatures. This can lead to delay in diagnosis during warm periods.

At the time of this study, there was no approved treatment for CAD. However, patients with moderate anemia and hemoglobin levels below approximately 10 g/dL or disabling cold-induced circulatory symptoms may require blood transfusions. Rituximab monotherapy or rituximab along with bendamustine are recommended as first-line therapy for patients with severe symptoms, depending on individual patient characteristics [6,17]. Corticosteroid treatment is not recommended for patients with CAD, especially for long-term treatment [6]. Novel pharmacological treatment options are also currently under development to improve clinical management [13,17] and consequently, the quality of life of patients [4].

In addition to these treatments, patients are recommended to avoid cold. Mild CAD symptoms may be managed by avoiding exposure to cold temperatures, avoiding cold food and water, using room heaters, and wearing warm clothing (warm shoes, scarves, gloves, earmuffs, warm inners, and stockings) [17].

Objectives

To enhance diagnosis and clinical management, it is necessary to better understand the symptom severity and impact of CAD from the patient's perspective. Qualitative studies on CAD are scarce and lacking. Su et al [12] found from qualitative interviews of 16 patients with CAD that the most frequently reported symptoms were fatigue, tiredness, or lack of energy and reaction to cold environments [2,12]. The rarity of CAD limits the ability to perform large-scale studies. The aims of this study were to increase the knowledge about CAD through a patient-centric survey and gain a better understanding of the burden of this disease from patients living with CAD in the United States.

Methods

Study Design and Participant Recruitment

This study included qualitative and quantitative research using a web-based questionnaire. The recruitment period started on September 1, 2020, and lasted for less than a month. Invitations to complete a self-administered questionnaire and follow-up emails were sent to members of the CAD Unraveled website, a website created by Sanofi and dedicated to providing support, information, and tools to patients with CAD to help them manage their condition. The Cold Agglutinin Disease Foundation (CADF), a nonprofit foundation dedicated to educating and supporting patients living with CAD, also shared the survey with its members. The study was based on voluntary and free participation of patients who agreed to participate in the survey. The study's inclusion criteria were as follows: patients self-reporting a diagnosis of CAD, aged ≥18 years,

residing in the United States, and being a registered member of the CAD Unraveled website or CADF.

Data Collection

The questionnaire (39 closed-ended questions and 5 open-ended questions) was designed by Carenity, in collaboration with Sanofi and CADF. The questionnaire was written in English, and the average estimated time needed to answer was 30 minutes.

The questionnaire started with a set of questions about the respondent's profile (eg, age, gender, place of residence, and profession), CAD characteristics, and disease status (eg, disease duration, age at diagnosis, type of CAD, symptoms experienced, and episodes experienced). CAD episodes were defined as a few hours during which a patient is affected by severe CAD symptoms.

The second part of the questionnaire was related to patient's perception of CAD severity and progression.

The third part assessed the patient's expectations of CAD-related fatigue management. Distinction between fatigue or tiredness, weakness, and lack of stamina was provided. Fatigue was defined as the feeling of tiredness or exhaustion or the need to rest because of lack of energy or strength and the need for extra sleep. Weakness was considered as a lack of physical or muscle strength and the feeling that extra effort is required to move arms, legs, or other muscles. Finally, lack of stamina was defined as an inability to maintain or sustain prolonged physical effort or activity.

The fourth part of the questionnaire evaluated the management methods used at the time of the survey, previous management methods, and patient satisfaction regarding management methods used at the time of the survey.

Finally, the last part explored the patient's perception of the CAD burden and the impact that CAD symptoms had on the patient's daily life, professional activities, and household finances (eg, impact owing to health care costs and work absenteeism). CAD burden was assessed at three time points (eg, when usual or regular symptoms were experienced, when symptom intensity or sensitivity increased or new symptoms occurred, and when no symptoms were experienced).

Statistical Analysis

Categorical variables are expressed as absolute frequencies and percentages. Continuous variable data are presented as mean (SD) for normal distribution and as median for nonnormal distribution. Narratives from open-ended questions were manually analyzed by identifying themes and subcategories.

Data processing and analysis were performed using RStudio (version 3.5.0; RStudio, Inc). Excel (Microsoft Corporation) 2013 was used to analyze the open-ended questions. The participants' personal information is not publicly available owing to privacy laws.

Ethics Approval

Written informed consent was obtained from all participants through information notice. The study protocol and patient materials were reviewed by the Western and Copernicus Group Institutional Review Board (formerly, New England Institutional Review Board; approval number 20202393).

Results

Description of the Study Population

A total of 116 participants started to fill in the questionnaire (registered as living in the United States or who had not indicated their country of residence), with 50 (43.1%) participants fulfilling the inclusion criteria and completing the questionnaire by September 24, 2020 (Figure 1). Patients were excluded from the study mainly because they indicated not living in the United States or not having CAD diagnosed by a health care professional.

Participants were located throughout the United States, with few participants from the Midwest. The mean age at enrollment was 66.7 (SD 8.5) years, and the men-to-women sex ratio was 0.22. Regarding employment status, 76% (38/50) of the patients declared that they were not working at the time of the survey and 20% (10/50) of the patients were working (Table 1).

At diagnosis, patients were on average aged 59.2 (SD 8.5) years. The main signs that led to diagnosis were blood tests (44/50, 88%) and presence of symptoms (26/50, 52%). Most patients (44/50, 88%) experienced symptoms before the diagnosis of CAD. The most frequently reported symptoms included fatigue (fatigue, tiredness, lack of stamina, or weakness; 37/50, 74%), shortness of breath (19/50, 38%), and acrocyanosis (18/50, 36%)

Among the 88% (44/50) of the patients who had experienced symptoms before being diagnosed, 34% (15/44) patients were diagnosed ≥2 years after their initial symptoms (Table 2). The delay between the development of initial symptoms and diagnosis was primarily owing to unawareness that the symptoms were associated with a disease (14/44, 32%) or because the patient had consulted multiple physicians before being properly diagnosed (13/44, 30%). The sociodemographic and clinical characteristics of the study participants are summarized in Tables 1 and 2.

Figure 1. Disposition of survey respondents. CAD: cold agglutinin disease; CADF: Cold Agglutinin Disease Foundation. *Indicates people who had agreed to receive invitations to participate in the study.

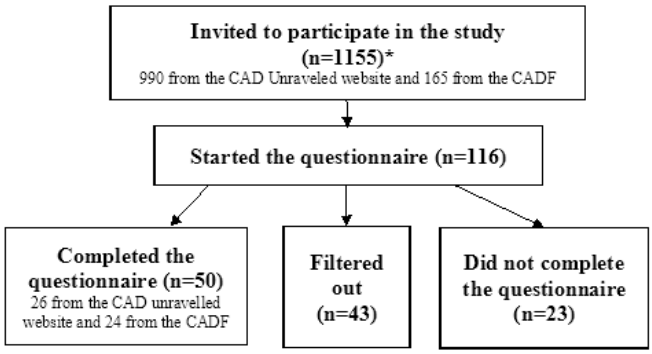


Table 1. Sociodemographic characteristics of patients with cold agglutinin disease in the United States recruited for the web-based survey (N=50).

Sociodemographic characteristics	Values
Sex, n (%)	
Women	41 (82)
Men	9 (18)
Age at the time of the survey (years), mean (SD; range)	66.7 (8.5; 45-86)
≤60, n (%)	12 (24)
61-65, n (%)	5 (10)
66-70, n (%)	14 (28)
71-75, n (%)	13 (26)
>75, n (%)	6 (12)
Employment status[a] at the time of the survey, n (%)	
Work—full-time	6 (12)
Work—part-time	4 (8)
Do not work	38 (76)
Disabled	2 (4)

[a]Multiple-choice question.

Table 2. Clinical characteristics of patients with CAD[a] in the United States recruited for the web-based survey (N=50).

Medical profile	Values
Type of CAD (according to patients' knowledge), n (%)	
Primary CAD	25 (50)
Secondary CAD	6 (12)
Do not remember	3 (6)
Do not know	16 (32)
Time since CAD diagnosis (years), mean (SD; range)	7.5 (7.4; 1-30)
<5, n (%)	21 (42)
5-10, n (%)	8 (16)
11-15, n (%)	5 (10)
>15, n (%)	7 (14)
Age at CAD diagnosis (years), mean (SD; range)	59.2 (8.5; 41-76)
<55, n (%)	17 (34)
55-60, n (%)	9 (18)
61-65, n (%)	13 (26)
>65, n (%)	11 (22)
Symptoms experienced before diagnosis (top 5)[b], n (%)	
Fatigue or tiredness	37 (74)
Decreased stamina	23 (46)
Increased weakness	22 (44)
Shortness of breath	19 (38)
Acrocyanosis[c]	18 (36)
Number of symptoms experienced before diagnosis, mean (SD)	4.6 (3.5)
Interval between initial symptoms and diagnosis[d], n (%)	
<6 months	15 (34)
6 months-1 year	7 (16)
1-2 years	6 (14)
2-3 years	7 (16)
>3 years	8 (18)
Do not remember	1 (2)

[a]CAD: cold agglutinin disease.

[b]Multiple-choice question; complete list is given in Figure S1 in Multimedia Appendix 1.

[c]Cold and blue limbs.

[d]Sample size: n=44.

CAD Symptoms Experienced and Perception of Disease Severity and Progression

More than half of the patients (27/50, 54%) considered their disease to be moderate or severe (Table 3). Patients with a severe form were principally affected by severe symptoms and were not satisfied with the efficiency of their management methods. Of the 50 patients, 31 (62%) patients reported that their CAD has progressed since their diagnosis: 21 (42%) patients thought that it has worsened and 10 (20%) thought that it has improved. Respondents who had experienced worsening of their CAD symptoms indicated that it occurred when the intensity or sensitivity of their symptoms increased (8/21, 38%), when new CAD symptoms appeared (4/21, 19%), and when both CAD symptom intensity or sensitivity increased and new CAD symptoms appeared (2/21, 10%). CAD varied over time: 88% (44/50) of the patients experienced episodes with increased intensity and sensitivity of CAD symptoms, 72% (36/50) experienced episodes with new CAD symptoms, and 54% (7/50) experienced episodes without symptoms. A total of 40% (20/50) of the patients had experienced all 3 types of episodes. Among the 94% (47/50) patients who reported CAD episodes with

increased intensity or sensitivity or new CAD symptoms, the most cited symptoms experienced during a CAD episode were fatigue (fatigue, tiredness, lack of stamina, or weakness; 42/47, 89%) and shortness of breath (31/47, 66%; Table 3). In total, 94% (44/47) of the patients identified at least one factor that caused an increase in symptom intensity and sensitivity. Among the 13 factors cited, the three major triggers were cold temperature (39/47, 83%), winter (32/47, 68%), and air conditioning (26/47, 55%; Table 3).

Totally, 90% (45/50) of the patients reported having experienced fatigue (fatigue, tiredness, lack of stamina, or weakness). Among these 45 patients, a total of 40 (89%) patients reported the symptoms at the time of the survey. In total, 47% (21/45) of the participants reported that it was moderate, 29% (13/45) reported that it was mild, and 13% (6/45) reported that it was severe. Fatigue was reported on a daily basis by 44% (20/45) of the patients, several times a week by 27% (12/45), and only after physical exertion by 16% (7/45) of the patients. When fatigue was experienced, 31% (14/45) of the patients mentioned that it usually fluctuated during the course of the day, 29% (13/45) reported that fatigue was constant throughout the day, and 22% (10/45) reported that fatigue symptoms were more intense in the afternoon.

Table 3. Episodes of CAD[a] reported by patients in the United States recruited for the web-based survey (N=50).

Episodes of CAD	Values
Perception of CAD severity, n (%)	
Mild	16 (32)
Moderate	21 (42)
Severe	6 (12)
Do not know	7 (14)
Progression of the disease, n (%)	
Worsened	21 (42)
Same	15 (30)
Improved	10 (20)
Do not know	4 (8)
Number of CAD episodes experienced in the past 12 months, mean (SD)	4.5 (5.4)
Do not know, n (%)	23 (46)
0-1, n (%)	9 (18)
2-3, n (%)	6 (12)
≥4, n (%)	9 (18)
Symptoms experienced during CAD episodes (top 5)[b,c], n (%)	
Fatigue or tiredness	39 (83)
Increased weakness	32 (68)
Decreased stamina	31 (66)
Shortness of breath	31 (66)
Dark urine	22 (47)
Number of symptoms experienced during CAD episodes[c], mean (SD)	6 (2.9)
1-2, n (%)	8 (17)
3-4, n (%)	8 (17)
5-6, n (%)	8 (17)
>6, n (%)	23 (49)
Triggering factors[b,c], n (%)	
Cold temperatures	39 (83)
Winter	32 (68)
Air conditioning	26 (55)
Sudden change in temperature	20 (43)
Infection	15 (32)
Psychological stress	10 (21)
Surgery	6 (13)
High humidity	3 (6)
Other[d]	5 (11)
None	3 (6)
Number of factors that triggered CAD symptoms[c], mean (SD)	3.3 (1.8)
<3, n (%)	16 (34)
3-4, n (%)	19 (40)
>4, n (%)	12 (26)

[a]CAD: cold agglutinin disease.

[b]Multiple-choice question.

[c]Sample size: n=47.

[d]Handling cold items (1/47, 2%), stress and physical activity (1/47, 2%), constant hemolysis creating chest pain in my sternum (1/47, 2%), overexertion (1/47, 2%), and unspecified (1/47, 2%).

Impact of CAD on Patients' Daily Life

When patients were asked in an open-ended question ("Which aspects of your daily life (emotionally, physically, socially, etc) are the most impacted by CAD?"), they named physical well-being (40 quotations), emotional well-being (31 quotations), and social life (28 quotations). Patients who reported an impact on their physical well-being primarily indicated the need to take naps (7 quotations), the inability to perform physical activities (7 quotations), and the need to limit their daily tasks (6 quotations). Emotional well-being was primarily affected owing to depression (7 quotations), frustration (6 quotations), or anxiety (4 quotations).

Among the 20% (10/50) of the patients who were employed at the time of the study, 90% (9/10) declared that CAD affected their professional life (Table 4). A total of 60% (30/50) of the patients reported that their household finances were affected by the disease. Totally, 83% (10/12) of the patients aged <60 years said that CAD affected their household finances, compared with 42% (8/19) of those aged >70 years. In total, 74% (37/50) of the patients had to cover out-of-pocket cost associated with the disease. Alternative medicines (21/50, 42%), transportation costs to medical appointments (16/50, 32%), and office visits or hospital care (16/50, 32%) were the most cited out-of-pocket costs to be covered, with an average of 2 out-of-pocket costs per patient (Table 4).

When patients were asked, with suggested items, which CAD-related symptoms had the greatest impact on their daily life, 90% (45/50) of the patients reported fatigue, 58% (29/50) of the patients reported shortness of breath, and 44% (22/50) of the patients reported joint pain, headaches, or acrocyanosis.

Patients were most affected during episodes with increased symptom intensity, increased sensitivity, or new CAD symptoms (median ≥8 out of 10, with 10=very strong impact), followed by episodes with usual symptoms (median ≥6 out of 10, with 10=very strong impact). Patients were less affected when no symptoms were experienced (median ≤3 out of 10), but they still indicated a negative impact on their quality of life (Figure 2).

Table 4. Impact of CAD[a] on patients with CAD in the United States recruited for the web-based survey (N=50).

Impact of CAD	Values, n (%)
Professional life[b,c]	
Had to take time off work	6 (60)
Could not work as much as they would like to	5 (50)
Not as efficient at work	2 (20)
Concerned that they would not be able to do the job when applying for a new position	2 (20)
Not able to work outside from late fall to early spring	1 (10)
Had to make modifications to avoid handling cold items	1 (10)
Unspecified	1 (10)
Not affected	1 (10)
Household finances	
No impact	20 (40)
Mild impact	15 (30)
Moderate impact	11 (22)
Severe impact	3 (6)
Not sure	1 (2)
Out-of-pocket costs related to CAD	
Alternative medicines (eg, vitamins and herbal medicines)	21 (42)
Transportation for medical appointments	16 (32)
Office visits or hospital care	16 (32)
Treatment for CAD and any related side effects	14 (28)
Transportation to relocate to a warmer region	10 (20)
Household services (eg, house cleaning and yard work)	9 (18)
Supportive care (eg, psychological support and nutrition)	7 (14)
Other[d]	5 (10)
None	13 (26)

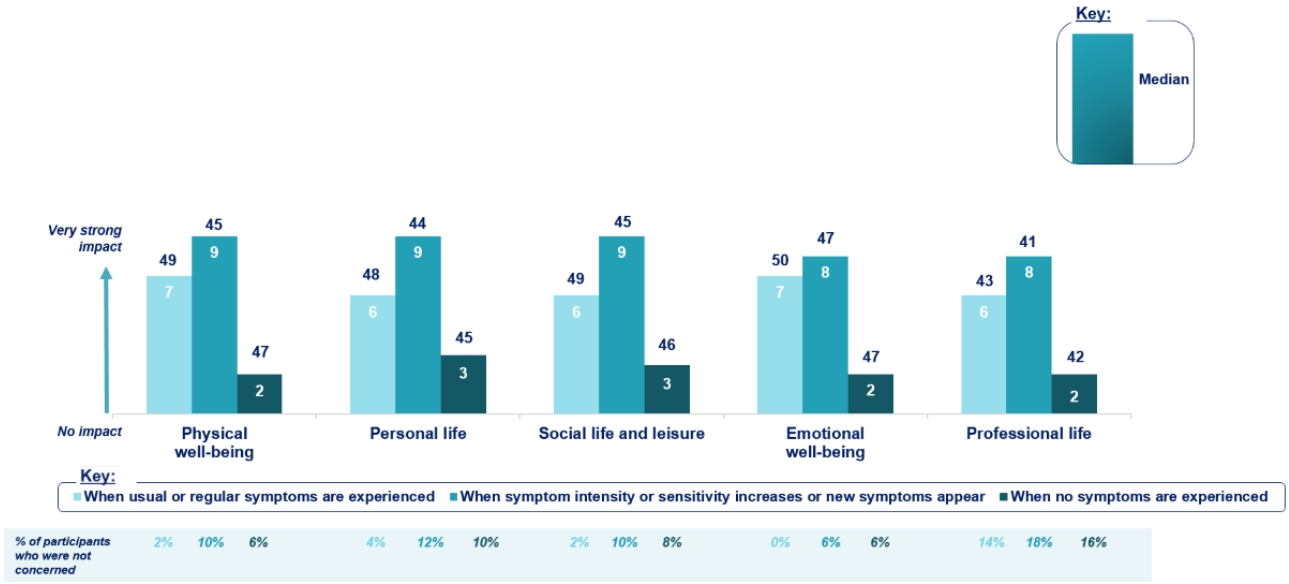
[a]CAD: cold agglutinin disease.

[b]Multiple-choice question.

[c]Sample size: n=10.

[d]Rental in warm area for 4 months each year (1/50, 2%); purchase more expensive health insurance (1/50, 2%); install a generator, which can be used if power is lost (1/50, 2%); move to a place for people aged >55 years, where everything is already available (1/50, 2%); and unspecified (1/50, 2%).

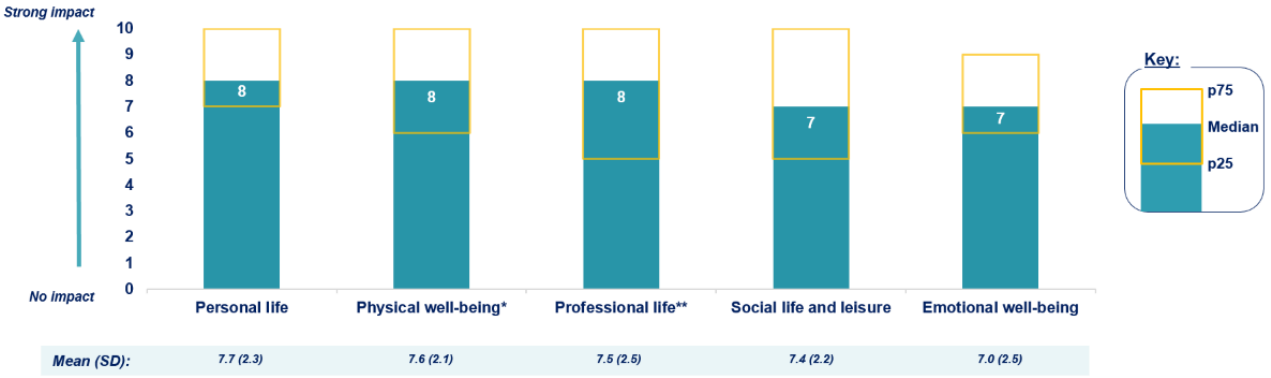
Figure 2. Impact of cold agglutinin disease on daily life reported by patients with that disease in the United States who were recruited for the web-based survey (N=50).



Impact of CAD-Related Fatigue

When patients who experienced fatigue, tiredness, lack of stamina, or weakness were asked in an open-ended question to describe how fatigue affected their daily life, they indicated that fatigue affected daily activities (38 quotations), associated fatigue with symptoms (37 quotations), and reported the need to implement solutions to cope with fatigue (31 quotations). Patients who reported fatigue-related symptoms primarily indicated weakness (22 quotations), difficulty in concentrating (6 quotations), and headaches (3 quotations). Those who indicated that fatigue affected their daily activities primarily cited an impact in terms of difficulty in completing daily tasks or starting new projects (12 quotations), difficulty in doing household chores (8 quotations), and the need to avoid physical activities (7 quotations). When asked on a scale from 0 to 10 (0=fatigue has no impact; 10=fatigue has a strong impact), it was found that fatigue had a significant impact on all aspects of daily life that were cited: personal life, physical well-being, professional life, social life and leisure, and emotional well-being (median ≥7 out of 10; Figure 3). Patients who experienced fatigue on a daily basis reported a more severe impact on each aspect of their daily life.

Figure 3. Impact of fatigue on daily life, reported by patients with cold agglutinin disease in the United States who were recruited for the web-based survey (N=50). p25: 25th percentile; p75: 75th percentile. *n=44 respondents affected; **n=37 respondents affected.



Symptom Management Methods

In total, 92% (46/50) of the patients reported using a management method at the time of the survey. The management method most often reported was cold avoidance (41/50, 82%), followed by rituximab treatment (10/50, 20%) and blood transfusion (4/50, 8%). On average, patients reported using 1.4 (SD 0.79) management methods at the time of the survey.

The main management methods used by patients with CAD in the past were cold avoidance (41/50, 82%), rituximab treatment (23/50, 46%), and corticosteroid treatment (15/50, 30%). Before the survey, patients reported using an average of 2.1 (SD 1.2) management methods.

A total of 32% (16/50) of the patients had to relocate away from where they used to live because of their CAD. Of these 16 patients, 9 (56%) patients relocated for part of the year and 7 (44%) patients relocated permanently. Totally, 59% (20/34) of the patients who had not relocated owing to CAD said that they want to do so. Incidentally, patients who relocated owing to CAD were more satisfied with their current management methods than those who did not relocate (7/15, 47% and 9/31, 29%, respectively).

Only 24% (12/50) of the patients declared having received blood transfusions owing to CAD in the previous 12 months. Among these 12 patients, 6 (50%) patients received at least three blood transfusions in the previous 12 months. When asked about their level of satisfaction regarding blood transfusion, half of the patients were satisfied, whereas the other half were unsatisfied (median 5 out of 10, with 10=agree with the item "I am satisfied with the way blood transfusions are performed"). Patients also reported that they often required assistance with transportation for blood transfusion appointments (median 6.5 out of 10, with 10=agree with the item "Someone often has to take me to or from my blood transfusions") or that the time spent in having infusions had an impact on their activities (median 5.5 out of 10 with 10=agree with the item, "Time spent for a blood transfusion has an impact on other activities").

Overall, among the 92% (46/50) of the patients who reported using a management method at the time of the survey, only 34% (16/46) were satisfied with the method used, primarily because of the improvement in CAD symptoms (4 quotations). A total of 31% (14/46) of the patients were dissatisfied with their current management method, including 11% (5/46) who were very dissatisfied. The most common reason for dissatisfaction was the lack of efficacy of the management method applied (5 quotations).

As CAD-related fatigue has a significant impact on the patients' quality of life, approximately all patients who experienced fatigue (44/45, 98%) had devised a coping mechanism to manage the symptoms. However, only 11% (5/45) of the patients sought professional help for fatigue. On average, patients implemented 5 coping mechanisms to deal with this symptom. However, only 28% (12/44) of the patients were satisfied with the coping mechanism that was implemented, among whom only 5% (2/44) were very satisfied. In contrast, 32% (14/44) of the patients were dissatisfied, including 9% (4/44) patients who were very dissatisfied.

Discussion

Principal Findings

This study brings new evidence of the burden of CAD from patients' perspectives.

In particular, our survey provided new data on fatigue experience (importance and type of fatigue). A total of 90% (45/50) of the patients reported having experienced fatigue (fatigue, tiredness, lack of stamina, or weakness). Fatigue was reported daily by 44% (20/45) of the patients, several times a week by 27% (12/45), and only after physical exertion by 16% (7/45) of the patients. When fatigue was experienced, patients mainly mentioned that it usually fluctuated during the course of the day. For example, in the survey, a patient stated the following:

> *Anything I need to routinely do, such as laundry, cooking, and light to moderate housework, leaves me completely drained.*

Fatigue, which is common to many disorders, could also explain the long delay between the first symptoms and the initial diagnosis.

Our findings also reveal that CAD had a great impact on patients' daily activities, physical well-being, emotional well-being, social life, and household finances. For example, a patient stated the following in the results:

> *I have always been a very active person. Not being able to do the same activities is very hard emotionally. We look normal to most people, but we are not and that explanation is very hard for most people to understand and I get frustrated trying to explain CAD to people.*

In total, 38% (19/50) of the patients experienced at least one CAD episode in the 12 months before the study. On average, they experienced 4 to 5 episodes during this time. Patients reported that CAD episodes were most often triggered by cold temperatures, winter, and air conditioning.

Finally, new data on disease severity and progression have been obtained. Most patients (27/50, 54%) considered their disease to be moderate or severe. Fatigue was reported to be the most prominent and most impactful symptom. People who were employed indicated that their work was affected by CAD. In total, 42% (21/50) of the study group reported that their symptoms had worsened since the time of diagnosis. To reduce the impact of CAD symptoms, most patients (46/50, 92%) reported using a management method.

These results confirm and supplement the findings from other studies.

Su et al [12] interviewed patients with CAD using a semistructured interview guide and demonstrated that fatigue, tiredness, or lack of energy; reaction to cold environments; shortness of breath; and trouble thinking or concentrating were the main symptoms experienced by patients and described the impact of CAD on daily activities. Nonetheless, that study included a very limited number of patients (n=16). This study, involving a large group of patients and using a questionnaire built with a patient advocacy group, allowed us to confirm the main findings published by Su et al [12] and provided additional evidence of the burden of CAD.

In comparison with other studies, our results also confirm the great impact of CAD on daily activities [12]. Moreover, the triggers quoted by patients (cold temperatures, winter, and air conditioning) were consistent with the information found in other studies [17]. Finally, the management method most often reported was cold avoidance, which is consistent with previous studies [6,17,18]. However, CAD symptoms have been shown to vary for the first time in a patient survey, implying variation in the impact of CAD. In addition, lack of satisfaction with the management methods used, which was not assessed in the previous studies, was shown.

Limitations

A few limitations of this study should be mentioned.

The self-administered web-based nature of the survey is likely to be biased toward patients who have access to the internet and are comfortable with using computers (ie, bias toward young patients). Approximately 10% of Americans do not use the internet [19], and many others may have access, but decline to

participate in web-based surveys. This bias may partly explain the difference between the study group profile, who were younger in terms of average age at the time of diagnosis (59.2 years) and average age at the time of study (66.7 years), compared with those in the previous studies. A multinational observational study of 232 patients with CAD by Berentsen et al [8] indicated a mean age of 67 years at disease onset, mean age of 68 years at the time of diagnosis, and mean age of 72 years at the time of study. In addition, the men-to-women sex ratio in our study group (0.22) varied from that (0.56) in the study by Berentsen et al [8]. Therefore, the results are likely to be biased toward women's perspective. It has been shown that this reflects the main characteristics of web-based users willing to share their experience with a disease [20].

Another potential limitation of a web-based survey is that it may exclude patients who are non–English-speaking or socioeconomically disadvantaged without internet access, who may have a different subset of symptoms, management methods, and diagnostic characteristics. In addition, if the information was not documented and the patient had to rely solely on memory to complete the questionnaire, certain clinical information may have been underestimated (eg, age at diagnosis, symptoms before diagnosis, number of CAD episodes over the past 12 months, symptoms during the episodes, and perception of disease severity). In addition, the study was only performed in the United States, and treatment, mental health, and so on can be different in other countries.

Patients were recruited from the CAD Unraveled website and CADF, which introduces a selection bias that may not be representative of the entire US population with the disease. This selection bias can be identified in the fact that the frequency of circulatory symptoms is lower in this study than in large studies among people with CAD [8,11]. However, our study population was different, as it was based in the United States (different states and temperature), with secondary CAD accepted. Collection of data was also different, as they were obtained from patients and not from physicians. The acrocyanosis reported in our study is only a part of the cold-induced circulatory symptoms. Other cold-induced circulatory symptoms are observed in CAD. For instance, in our study, 50% (25/50) of the patients reported Raynaud syndrome or acrocyanosis. Another selection bias can be the length of the survey. A total of 30 minutes were needed to complete the whole questionnaire, which could have affected the answer rate. Finally, among the 1155 patients reached, 50 (4.33%) patients completed the questionnaire and no hard criteria for diagnosis of CAD were used, which can bias the sample. The sample may not be representative of the entire US population with CAD. This participation rate can be explained by different factors (no compensation, methodology used, presence of caregivers or people not diagnosed with CAD or with invalid emails in the CAD Unraveled database, overlap between both databases, reduction of the open survey window, etc).

The absence of a question on race and ethnicity is a limitation of our study, as it has historically affected access to care in the United States. In addition, recall bias could have affected the results, especially for the CAD burden, as it was asked at 3 different time points. As many patients (36/50, 72%) were retired or did not work owing to reasons other than CAD, our assessment of the potential impact of CAD on professional life was limited. Circulatory symptoms were not fully addressed in the context of impact on daily life. However, such impact can be substantial.

Finally, even with the purpose of the study being scientifically driven with no patient compensation, voluntary participation, and review by an ethics committee, a bias in information cannot be ruled out, as most of the questions were closed (39/45, 87%), limited in spontaneous reporting, and completed knowing the study was conducted by a pharmaceutical company. In addition, the study may be biased as it was conducted by a pharmaceutical company and Carenity, a company closely linked to the pharmaceutical company conducting the study. However, a patient advocacy group dedicated to CAD was involved to reduce this bias as much as possible.

Conclusions

In conclusion, our study provides new insights into the symptoms associated with CAD and the impact of the disease from the patient's perspective. It particularly highlights the importance and type of fatigue, seriousness of the disease as perceived by patients, impact of CAD even when no symptoms are experienced, and importance of CAD symptom variation, implying that the impact of CAD varies over time. In addition, difficulties in CAD diagnosis have been raised. Patients with CAD experienced difficulty in receiving proper diagnosis and treatment owing to their rare disease; hence, they should be diagnosed earlier and, then, closely monitored and advised by a health care provider with knowledge of how to treat the disease.

Additional studies are necessary to better understand the burden of CAD-related symptoms, especially fatigue, and the patient's needs regarding symptom management. It is also important to raise awareness among health care professionals regarding CAD-related symptoms to appropriately diagnose and support these patients. Specifically, health care professionals should proactively discuss ways to manage fatigue, which is not often discussed during consultations.

Acknowledgments

The authors wish to thank all people living with cold agglutinin disease who participated in this survey and Corinne Forster from the Cold Agglutinin Disease Foundation for her contributions to the design of the web-based survey. Editorial support was provided by Florentina Rafael (freelancer) as well as Tracy Royle and Helen Marshall of Ashfield Medcomms (an Ashfield Health company). Sanofi sponsored the study and the development of this manuscript.

Conflicts of Interest

FJ and LAS are employees and stockholders of Sanofi. EP is an employee and DT is a former employee of Carenity, a company that received funds from Sanofi to conduct the study. PAMW has served as a speaker for Sanofi.

Multimedia Appendix 1
Symptoms experienced by patients before diagnosis (N=50).
[DOCX File , 107 KB - formative_v6i7e34248_app1.docx]

References

1. Berentsen S, Sundic T. Red blood cell destruction in autoimmune hemolytic anemia: role of complement and potential new targets for therapy. Biomed Res Int 2015;2015:363278 [FREE Full text] [doi: 10.1155/2015/363278] [Medline: 25705656]
2. Berentsen S, Tjønnfjord GE. Diagnosis and treatment of cold agglutinin mediated autoimmune hemolytic anemia. Blood Rev 2012 May;26(3):107-115. [doi: 10.1016/j.blre.2012.01.002] [Medline: 22330255]
3. Ulvestad E, Berentsen S, Mollnes TE. Acute phase haemolysis in chronic cold agglutinin disease. Scand J Immunol 2001;54(1-2):239-242 [FREE Full text] [doi: 10.1046/j.1365-3083.2001.00960.x] [Medline: 11439172]
4. Jäger U, Barcellini W, Broome CM, Gertz MA, Hill A, Hill QA, et al. Diagnosis and treatment of autoimmune hemolytic anemia in adults: recommendations from the First International Consensus Meeting. Blood Rev 2020 May;41:100648. [doi: 10.1016/j.blre.2019.100648] [Medline: 31839434]
5. Bylsma LC, Gulbech Ording AG, Rosenthal A, Öztürk B, Fryzek JP, Arias JM, et al. Occurrence, thromboembolic risk, and mortality in Danish patients with cold agglutinin disease. Blood Adv 2019 Oct 22;3(20):2980-2985 [FREE Full text] [doi: 10.1182/bloodadvances.2019000476] [Medline: 31648316]
6. Berentsen S. How I manage patients with cold agglutinin disease. Br J Haematol 2018 May;181(3):320-330. [doi: 10.1111/bjh.15109] [Medline: 29363757]
7. Mullins M, Jiang X, Bylsma LC, Fryzek JP, Reichert H, Chen EC, et al. Cold agglutinin disease burden: a longitudinal analysis of anemia, medications, transfusions, and health care utilization. Blood Adv 2017 May 23;1(13):839-848 [FREE Full text] [doi: 10.1182/bloodadvances.2017004390] [Medline: 29296728]
8. Berentsen S, Barcellini W, D'Sa S, Randen U, Tvedt TH, Fattizzo B, et al. Cold agglutinin disease revisited: a multinational, observational study of 232 patients. Blood 2020 Jul 23;136(4):480-488 [FREE Full text] [doi: 10.1182/blood.2020005674] [Medline: 32374875]
9. Swiecicki PL, Hegerova LT, Gertz MA. Cold agglutinin disease. Blood 2013 Aug 15;122(7):1114-1121 [FREE Full text] [doi: 10.1182/blood-2013-02-474437] [Medline: 23757733]
10. Barcellini W. Current treatment strategies in autoimmune hemolytic disorders. Expert Rev Hematol 2015 Oct;8(5):681-691. [doi: 10.1586/17474086.2015.1073105] [Medline: 26343892]
11. Berentsen S, Ulvestad E, Langholm R, Beiske K, Hjorth-Hansen H, Ghanima W, et al. Primary chronic cold agglutinin disease: a population based clinical study of 86 patients. Haematologica 2006 Apr;91(4):460-466. [Medline: 16585012]
12. Su J, Kosa K, DiBenedetti D. Patient-reported disease burden: in-depth interviews of patients with CAD. Blood 2020 Nov 5;136(1):29-30. [doi: 10.1182/blood-2020-136788]
13. Berentsen S. New insights in the pathogenesis and therapy of cold agglutinin-mediated autoimmune hemolytic anemia. Front Immunol 2020 Apr 7;11:590 [FREE Full text] [doi: 10.3389/fimmu.2020.00590] [Medline: 32318071]
14. Berentsen S. Cold agglutinin disease. Hematology Am Soc Hematol Educ Program 2016 Dec 02;2016(1):226-231 [FREE Full text] [doi: 10.1182/asheducation-2016.1.226] [Medline: 27913484]
15. Broome CM, Cunningham JM, Mullins M, Jiang X, Bylsma LC, Fryzek JP, et al. Increased risk of thrombotic events in cold agglutinin disease: a 10-year retrospective analysis. Res Pract Thromb Haemost 2020 Apr 9;4(4):628-635 [FREE Full text] [doi: 10.1002/rth2.12333] [Medline: 32548562]
16. Hansen DL, Berentsen S, Fattizzo B, Hansen PL, Barcellini W, Frederiksen H. Seasonal variation in the incidence of cold agglutinin disease in Norway, Denmark, and Italy. Am J Hematol 2021 Jul 01;96(7):E262-E265. [doi: 10.1002/ajh.26196] [Medline: 33864697]
17. Berentsen S, Röth A, Randen U, Jilma B, Tjønnfjord GE. Cold agglutinin disease: current challenges and future prospects. J Blood Med 2019 Apr 9;10:93-103 [FREE Full text] [doi: 10.2147/JBM.S177621] [Medline: 31114413]
18. Cold Agglutinin Disease. National Organization for Rare Disorders. URL: https://rarediseases.org/rare-diseases/cold-agglutinin-disease/#:~:text=General%20Discussion,subtype%20of%20autoimmune%20hemolytic%20anemia [accessed 2022-05-17]
19. Keeter S, McGeeney K. Coverage Error in Internet Surveys. Pew Research Center. 2015 Sep 22. URL: https://www.pewresearch.org/methods/2015/09/22/coverage-error-in-internet-surveys/ [accessed 2021-02-17]
20. Raïs S, Radoszycki L, Dourgnon P, Rochaix L, Chekroun M. Accurate representation of patients' opinions for decision-making: are online health communities good candidates? Value Health 2017 Oct 1;20(9):PA760-PA761. [doi: 10.1016/j.jval.2017.08.2153]

Abbreviations

AIHA: autoimmune hemolytic anemia
CAD: cold agglutinin disease
CADF: Cold Agglutinin Disease Foundation

Edited by A Mavragani; submitted 13.10.21; peer-reviewed by M Gilbert, S Berentsen; comments to author 25.02.22; revised version received 05.04.22; accepted 01.06.22; published 22.07.22.

Please cite as:
Joly F, Schmitt LA, Watson PAM, Pain E, Testa D
The Burden of Cold Agglutinin Disease on Patients' Daily Life: Web-Based Cross-sectional Survey of 50 American Patients
JMIR Form Res 2022;6(7):e34248
URL: https://formative.jmir.org/2022/7/e34248
doi:10.2196/34248
PMID:35867390

Original Paper

Virtual Intervention for Caregivers of Persons With Lewy Body Dementia: Pilot Quasi-Experimental Single-Arm Study

Oleg Zaslavsky[1], BSN, MHA, PhD; Jasmine Kaneshiro[1], DNP; Frances Chu[1], MSN, MLiS; Andrew Teng[1], PhD; Kimiko Domoto-Reilly[1], MD; Annie T Chen[1], PhD

University of Washington, Seattle, WA, United States

Corresponding Author:
Oleg Zaslavsky, BSN, MHA, PhD
University of Washington
1959 Pacific Street NE
Seattle, WA, 98195
United States
Phone: 1 2068493301
Email: ozasl@uw.edu

Abstract

Background: Compared to other types of dementia, family caregivers of people with Lewy body dementia (LBD) report higher stress levels and more severe depressive symptoms. Although several digital support interventions for caregivers of persons with dementia exist, few target LBD specifically or leverage a fully remote and asynchronous approach suitable for pandemic circumstances.

Objective: We performed a pilot evaluation of a digital intervention designed to help caregivers of people with LBD address challenges they have experienced, with the end goal of reducing psychological distress in this population.

Methods: We recruited 15 family caregivers of people with LBD to participate in the quasi-experimental, single-arm, mixed methods study titled Virtual Online Communities for Aging Life Experience–Lewy Body Dementia (VOCALE-LBD). The study offers an 8-week web-based intervention that uses a digital discussion platform and involves moderation, peer-to-peer support, didactic training, and problem-solving skill enactment.

Results: Participants' baseline characteristics were the following: mean age 66 (SD 8) years; 14 of 15 (93%) of them were female; all (15/15, 100%) were White; and 8 (53%) of them had at least a postgraduate degree. Throughout the intervention, participants engaged in weekly web-based discussions, generating a total of 434 posts (average 4 posts per week). Attrition was 20% (3/15). Upon study exit, participants showed the following average improvements: 3.0 (SD 6.0) in depression, 8.3 (SD 16.7) in burden, 2.9 (SD 6.8) in stress, and 0.3 (SD 0.8) in loneliness. When looking at the proportion of participants with clinically significant improvement versus those with a worsening of ≥0.5 SD for each outcome, we observed net improvements of 50% (6/12), 33% (4/12), 25% (3/12), and 25% (3/12) in depression, loneliness, burden, and stress, respectively. In terms of the benefits of participation, participants reported that participation helped them "a great deal" to (1) improve their understanding of LBD (9/12, 75%), (2) gain confidence in dealing with difficult behaviors of the care recipient (6/12, 50%), and (3) improve in one's abilities to provide care to the care recipient (4/12, 33%).

Conclusions: The study generated promising feasibility and preliminary efficacy data for a low-cost, web-based intervention designed for caregivers of persons with LBD. Though the study was not powered for significance, we observed nominal average and net improvements in important psychological outcomes. Moreover, many caregivers reported that study participation helped them better understand the disease, feel more confident in dealing with difficult behaviors of the care recipient, and improve their ability to care for the care recipient. If validated in future studies, the intervention could be an accessible, on-demand resource for caregivers, enabling them to engage in moderated remote discussions with peers at their own convenience in terms of location, time of the day, and frequency.

(JMIR Form Res 2022;6(7):e37108) doi:10.2196/37108

KEYWORDS

dementia; caregiver; internet based; digital health; digital intervention; eHealth; feasibility; web-based; peer support; didactic training; caregiving; informal care; spousal care; remote intervention; Lewy body; Lewy bodies; discussion forum; discussion

platform; online support; distress; stress reduction; online discussion; support group; discussion group; burden; depression; depressive symptom; lonely; loneliness; mental health

Introduction

Lewy body dementias (LBDs), referring to both dementia with Lewy bodies and Parkinson disease dementia (PDD), are the second most common type of degenerative dementia in older adults [1]. These are complex disorders in which patients often exhibit disruptive behaviors that make caregiving challenging [1,2]. Compared to other types of dementia, caregivers of people with LBD report higher stress [3] and more severe depressive symptoms [4]. Many were unsatisfied with the support they received, even in pre–COVID-19 contexts [5]. The ongoing COVID-19 pandemic has multiplied the challenges for family caregivers of persons with LBD [6]. First, changes in daily life, such as limited in-person social contact, can exacerbate neurobehavioral symptoms, causing significant distress to caregivers [7]. Second, as older persons with dementia are especially susceptible to COVID-19 complications [8], caregivers of persons with LBD may engage in more restrictive, prolonged protective practices that affect their own opportunities for peer-to-peer connection, socialization, and physical and cognitive stimulation. Finally, many caregivers rely on support services that could be severely disrupted or even permanently discontinued amid the ongoing pandemic. As such, support interventions that conform to "the new normal" realities in the COVID-19 era for caregivers of persons with LBD are warranted.

A proliferation of interventions for informal caregivers of persons living with dementia has occurred in recent years, including digital interventions focusing on caregiver empowerment and the psychological health [9-13]. However, many of the digital interventions do not account for the unique characteristics of their target users, affecting their effectiveness and implementation potential [14], and still include in-person elements that might be suboptimal in pandemic circumstances. Many lack a fully remote, asynchronous approach that allows on-demand engagement in terms of location, time, and frequency and do not focus on LBD specifically. A fully remote, asynchronous intervention specifically designed for family caregivers of persons living with LBD, leveraging peer-to-peer support, may fill the gap and improve psychological health in this population.

This study examined the feasibility of such an intervention, called Virtual Online Communities for Aging Life Experience–Lewy Body Dementia (VOCALE-LBD), in terms of recruitment, retention, and preliminary efficacy concerning the following psychological outcomes: caregiver burden, depressive symptoms, stress, and loneliness using a quasi-experimental, single-arm, mixed methods design. Owing to the preliminary nature of the study, the quantitative analyses were not powered for significance but were used to examine whether the intervention was associated with trends in the expected directions of the outcome measures. The qualitative data, on the other hand, provided opportunities to better understand participants' experiences with the intervention and areas for improvement.

Methods

Study Design

This was a prospective, one-group, pre-post study to assess the feasibility and preliminary efficacy of a new web-based intervention involving moderation, peer-to-peer support, didactic training, and problem-solving skill enactment. Participants were granted access to the VOCALE-LBD intervention hosted on a private website.

Ethical Considerations

The University of Washington's (UW's) institutional review board (IRB) approved the study protocol (approval 13431). At enrollment, participants provided informed consent to be included in the study.

Participants

Participants were recruited from the Memory and Brain Wellness Center (MBWC) at Harborview Medical Center in Seattle, Washington. The MBWC encompasses both the MBWC clinic and the Alzheimer's Disease Research Center (ADRC) and has been designated as a Lewy Body Dementia Association Research Center of Excellence [15]. The MBWC evaluates over 1000 new patients on an annual basis, and it is the only major academic medical center serving the 5-state region of Washington, Wyoming, Alaska, Montana, and Idaho. The ADRC maintains a continually updated, comprehensive contact list of persons who have signed an IRB-approved consent form to be contacted about participation in research studies conducted by UW-affiliated researchers. Many of the Research Registry members have been evaluated at the MBWC clinic and hence have a recent, reliable clinical diagnosis of specific types of dementia such as Alzheimer disease, Parkinson disease dementia (PDD), Dementia with Lewy bodies (DLB), etc. Family members of clinic patients are also eligible to join the registry. All have expressed willingness to be contacted for potential study participation. Eligibility criteria for this study were as follows: being family or informal caregiver of a person with a diagnosis of LBD; being able to read, write, and speak English; having a device that can access the internet and be used for videoconferencing or telephone calls; and being ≥18 years old. Of note, when possible, MBWC clinicians aim to distinguish between DLB and other subtypes of dementia, such as PDD, and to educate patients and their caregivers on the exact diagnosis. As such, most, if not all, study participants cared for people who were highly likely to have a diagnosis of DLB. Participants were compensated up to US $250 depending on their participation.

Intervention

We adapted a prior social networking intervention that was developed for older adults with prefrailty and frailty, VOCALE [16-18], to the needs of caregivers of individuals with LBD. Similar to our previous studies, the refined intervention included training and moderated web-based discussion components. Training sessions were performed remotely; participants were

introduced to the platform (Figure 1) and described activities to learn how to interact with the discussion board. Weekly thematic discussion prompts (Figure 2 shows an example of a prompt concerning hallucinations) allowed participants to respond to a specific topic of interest at their leisure and provided the participants an opportunity to interact with each other. The first 3 weeks were focused on the most salient LBD caregiving experiences featured in previous literature [5,19,20] and our previous formative work. Briefly, in this formative work, we conducted 8 individual interviews and 2 focus groups with caregivers of persons with LBD to identify relevant topics for the intervention. We used open-ended prompts to elicit ideas that caregivers would be interested in discussing during the study. Sleep problems, hallucinations and delusions, and self-care emerged as the topics of greatest interest and thus became our topics for the first 3 weeks.

The next 5 weeks involved psychoeducational materials based on problem-solving therapy (PST), a cognitive behavioral intervention focused on the adoption and application of adaptive problem-solving attitudes and skills [21]. PST has been used with individuals with different types of problems, including caregivers [22,23] and older adults with other health issues [24]. The intervention also incorporated personas or prototypical examples of caregivers of persons with LBD (Figure 3 shows an example of a persona). Personas, or representations or archetypes grounded in real data, are often used in user-centered design to help inform product design [25]. In this context, we used personas to enable participants to practice solving realistic problems that they might face. Working with the personas also provided some other potential benefits, such as being able to practice problem-solving without sharing anything too personal, as well as having an outlet to focus on someone besides themselves.

Two research staff members served as the study's moderators. Both were advanced practice nurses who received moderator training from the study team. Moderators logged in at least 3 times each weekday and at least once per weekend day to review any new comments and add comments to address questions, provide emotional support and validation, encourage dialogue, and redirect discussions as needed to stay on topic. The moderators did not provide any medical advice. Moderators also monitored for comments with inaccurate information and addressed them as needed through private email or discussion board comments. Lastly, moderators sent personalized reminder emails once a week to participants who had not yet posted. They replied to most participants' comments.

Figure 1. Landing page of Virtual Online Communities for Aging Life Experience (VOCALE)–Lewy Body Dementia.

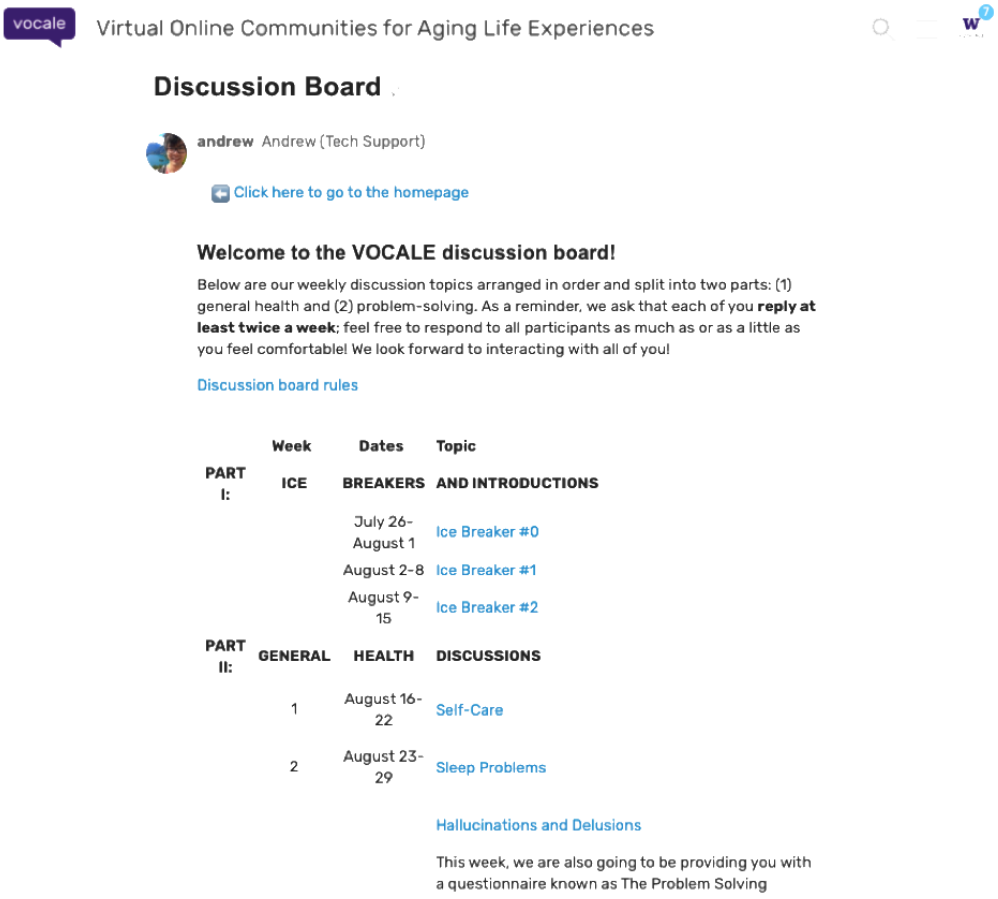


Figure 2. Virtual Online Communities for Aging Life Experience–Lewy Body Dementia sample weekly discussion prompt. LBD: Lewy body dementia.

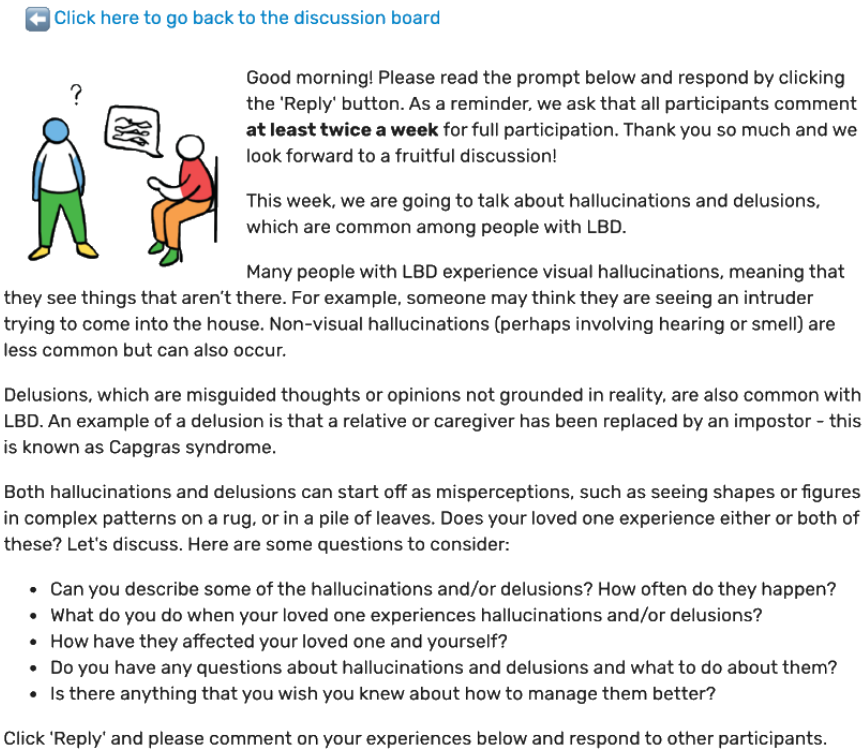
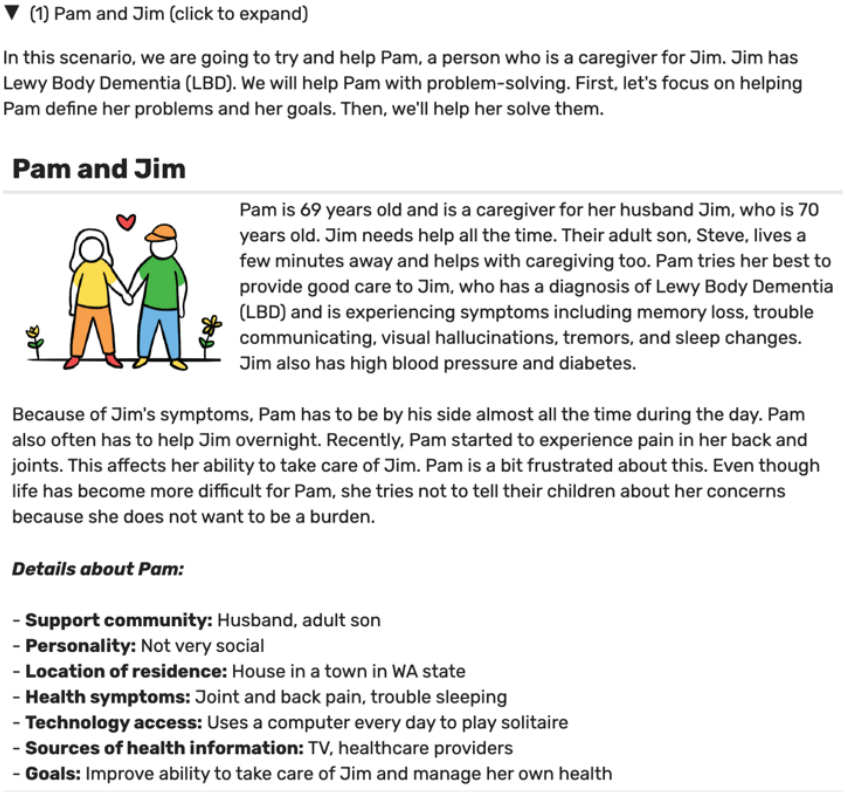


Figure 3. Virtual Online Communities for Aging Life Experience–Lewy Body Dementia persona page.



Data Collection Procedures

All data collection sessions in this pilot were virtual and took place on Zoom. The study team collected data using Research Electronic Data Capture (REDCap) [26], a web-based survey platform, at 2 time points: baseline and post intervention. At baseline, the research coordinator provided training on the discussion board website and administered a set of demographic and clinical questionnaires (described in *Measures*). The second postintervention assessment occurred between weeks 9 and 10 and included another set of questionnaires and exit interviews. Exit interviews lasted approximately an hour each and were

conducted with a semistructured interview guide that asked for caregiver feedback on topics including various program components, whether people felt a sense of community or support, what was learned or gained from the discussion, and motivations to participate. Caregivers were also given the opportunity to provide any additional feedback or suggestions.

Measures

Depressive Symptoms

We used the 20-item version of the Center for Epidemiologic Studies Depression (CES-D) Scale to assess depression [27]. For each statement, respondents indicated how often they felt depressed during the past week using a scale from 0=rarely or never to 3=most or all of the time. Scores could range from 0 through 60, with higher scores indicating an increased presence of depression symptoms.

Caregiver Burden

We used the full 22-item version of the Zarit Caregiver Burden Interview [28]. Caregivers rated each item on a 5-point scale from 0=never to 4=nearly always, yielding a possible range of 0 to 88. Higher values indicated greater levels of caregiver burden.

Perceived Stress

We used the full 10-item version of the Perceived Stress Scale [29]. Caregivers rated each item on a 5-point scale from 0=never to 4=very often, yielding a possible range of 0 to 40. Higher values indicated greater perceived stress.

Loneliness

We used a short 3-item version of the Revised UCLA Loneliness Scale [30]. Its possible range was from 0 to 6 with higher scores indicating greater loneliness.

Social Support

We used a 9-item questionnaire from the Medical Outcomes Study [31]. The questionnaire was designed to assess the amount of social support the participant had available. Responses were scored on a 5-point scale ranging from “none of the time” to “all of the time.” We computed total social support scores by summing the scores for all items after recoding responses. Scores could range from 0 through 45, with higher scores indicating increased social support.

Self-efficacy

We used a 5-item Health Self-Efficacy Measure [32]. For each statement, respondents indicated their level of agreement concerning their health management using a scale from 0=strongly disagree to 4=strongly agree. Scores could range from 0 through 20, with higher scores indicating stronger health self-efficacy.

Benefits of Participation

On conclusion of the study, we asked all caregivers 4 questions adapted from the REACH II study [33] about ways in which they benefited from participating in the study. Specifically, we asked whether their participation helped improve their understanding of LBD, improve their confidence and ability to deal with difficult behaviors of the care recipient, and make their caregiving life easier. The response options for each question were “not at all,” “some,” and “a great deal.”

Data Analysis

Descriptive statistics summarized demographics at baseline and outcome distributions by measurement occasions. We also calculated mean (SD) pre- and postintervention change scores for each of the measurements. Next, for each outcome, we calculated the number and percentage of individuals who improved by at least 0.5 SD from baseline to post intervention, and the number and percentage of those who worsened by at least 0.5 SD, and then we subtracted the former form the latter to calculate a net improvement score. We also calculated the number and percentage of individuals selecting response options from the benefit of the participation questionnaire.

Exit interviews were recorded and transcribed verbatim. We analyzed the data using an inductive method adapted from previous research concerning digital interventions for behavioral change [34]. The research team reviewed transcripts line by line and then created a table to organize interview data at the participant level into categories. Categories were based on the interview guide, such as motivations to participate in the study, what was learned from the experience, and participants’ thoughts on different program components. Suggestions, critiques, and other comments from participants were categorized as actionable, not actionable, or needing further discussion and consideration in the future. After each participant’s interview, data were added to the table, and common responses were tabulated. Data abstraction was performed by JK, and all the coauthors participated in the review of the table.

Results

Participants

Of the 19 potential candidates screened, 15 were eligible, consented to the study, and provided baseline data (Figure 4). A total of 3 participants left the study at weeks 2, 3, and 5, which yielded a total of 12 participants who completed the full protocol, resulting in an 80% retention rate. Table 1 summarizes the baseline data. The mean participant age was 66 (SD 8) years, all participants (n=15, 100%) were White, and 93% (14/15) were female. In total, 8 (53%) participants had at least a postgraduate degree, and 80% (12/15) of the participants provided more than 40 hours of care per week.

Figure 4. Flowchart for study screening, enrollment, and completion. LBD: Lewy body dementia.

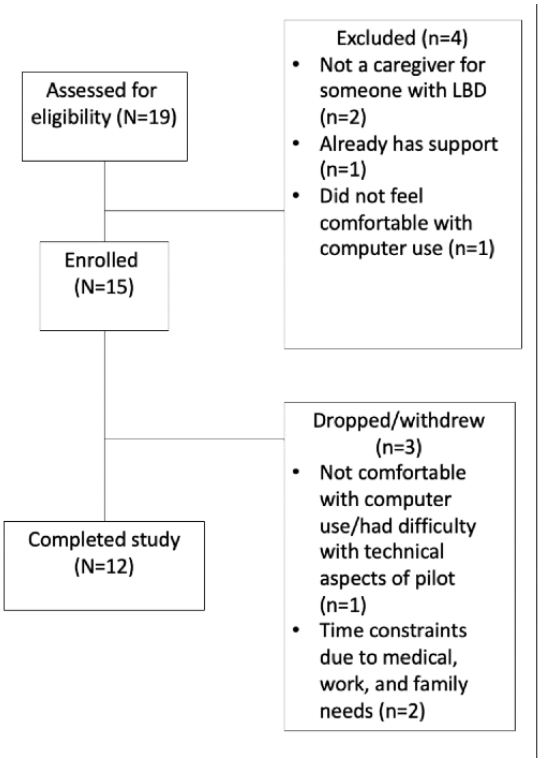


Table 1. Participant demographics (baseline; N=15).

Variables	Values
Sex, n (%)	
Female	14 (93)
Male	1 (7)
Age (years), mean (SD)	65.8 (8.3)
Caregiver relationship: spouse or partner, n (%)	15 (100)
White, not of Hispanic/Latino ethnicity, n (%)	15 (100)
Education level, n (%)	
High school diploma or General Educational Development	3 (20)
Vocational or associate's degree	3 (20)
Baccalaureate degree	1 (7)
Master's degree	5 (33)
Doctoral degree	3 (20)
Time spent as a caregiver (years), mean (SD)	3.8 (2.5)
Average care time per week, n (%)	
Up to 8 hours	1 (7)
20-39 hours	2 (17)
≥40 hours	12 (80)
Computer comfort level, n (%)	
Very uncomfortable	2 (17)
Neutral	1 (7)
Somewhat comfortable	6 (40)
Very comfortable	6 (40)

Data, Measures, and Analysis

Table 2 summarizes engagement metrics over time. Throughout the intervention, participants engaged in weekly web-based discussions, generating a total of 434 posts (average of 4 posts per week). On study exit (Table 3), participants showed the following mean improvements: 3.0 (SD 6.0) in depression, 8.3 (SD 16.7) in burden, 2.9 (SD 6.8) in stress, and 0.3 (SD 0.8) in loneliness. When we calculated differences in the proportion of participants with clinically significant improvement versus those with a worsening of ≥0.5 SD for each outcome (Table 4), we observed net improvements of 50% (6/12), 35% (4/12), 25% (3/12), and 25% (3/12) in depression, loneliness, burden, and stress, respectively. When we assessed the benefits of participation (Table 5), 75% (9/12) of the participants reported that participation helped them "a great deal" to improve their understanding of LBD compared to 25% (3/12) of them selecting the "some" response option. Similarly, in response to the question about confidence in dealing with difficult behaviors of the care recipient, 50% (6/12) of the participants selected "a great deal," 42% (5/12) of them selected "some," and 8% (1/12) of them selected "not at all." Finally, in response to a question about improvement in one's abilities to provide care to the care recipient, 33% (4/12) of the participants selected "a great deal," 58% (7/12) of them selected "some," and 8% (1/12) of them selected "not at all."

Qualitative data from the exit interviews showed that of the 12 participants who completed the pilot, 92% (11/12) of them reported they had a positive experience. All participants noted that they felt supported by other participants. Motivations for participating in the study included the following: wanting to support others (4/12, 33%), wanting to interact with others in similar situations and obtain feedback (6/12, 50%), and contributing to research (4/12, 33%).

When asked what they learned or gained from the pilot, 50% (6/12) of the participants mentioned the opportunity to hear from others who were caregivers of people with LBD, and 83% (10/12) of them discussed learning from others, such as solutions and strategies that others have used. Two-thirds (8/12, 67%) of participants felt that the problem-solving elements of the pilot were helpful. One-fourth (3/12, 25%) of participants felt that their baseline problem-solving skills were already robust; therefore, the intervention did not contribute to improvement. One-third (4/12, 33%) of participants did not feel that the intervention's problem-solving framework was helpful, with 2 participants (17%) mentioning that they felt the framework was difficult to apply in real life. All participants offered constructive feedback regarding the personas. For example, one-third (4/12, 33%) of participants felt that they could not relate to one or both personas, and one-third (4/12, 33%) of them felt that the personas needed to be fleshed out with more details to be useful. Over half (7/12, 58%) of the participants would have preferred to discuss participants' actual problems rather than those of the personas.

Table 2. Participant engagement (posts per week).

Participant ID	Posts, n								Total posts, n
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	
1	3	4	2	4	7	7	5	5	37
2	9	3	2	3	3	6	1	2	29
3	2	5	2	2	2	4	2	3	22
4	2	3	2	2	3	1	0	1	14
5	4	7	3	1	3	0	0	0	18
6	8	5	15	3	1	3	3	6	44
7	4	2	0	0	0	0	0	0	6
8	15	12	8	2	8	10	2	7	64
9	4	6	5	2	4	3	3	2	29
10	3	3	4	2	4	2	2	2	22
11	6	12	5	5	8	10	6	5	57
12	4	5	5	2	4	2	3	4	29
13	3	6	9	4	2	8	2	5	39
14	0	0	0	0	0	0	0	0	0
15	3	3	2	3	3	3	3	4	24
Weekly total	70	76	64	35	52	59	32	46	434

Table 3. Baseline, postintervention, and change scores (N=15).

Parameters	Baseline, mean (SD)	Postintervention (n=12), mean (SD)	Change (n=12), mean (SD)
Depressive symptoms	13.3 (9.0)	10.9 (9.9)	–3.0 (6.0)
Health self-efficacy	21.5 (3.3)	21.8 (3.6)	–0.6 (4.3)
Perceived stress	16.2 (7.2)	13.0 (9.0)	–2.9 (6.8)
Loneliness	5.4 (1.7)	5.0 (1.7)	–0.33 (0.8)
Social support	32.3 (7.9)	33.08 (7.1)	–0.33 (4.2)
Caregiving burden	39.1 (10.8)	36.3 (17.0)	–8.3 (16.6)

Table 4. Clinically significant changes[a] (N=12).

Variable	Improved, n (%)	Worsened, n (%)	Net improvement[b], n (%)
Depressive symptoms	8 (70)	2 (17)	6 (50)
Health self-efficacy	3 (25)	4 (33)	–1 (–8)
Perceived stress	5 (42)	2 (17)	3 (25)
Loneliness	4 (33)	0 (0)	4 (33)
Social support	3 (25)	2 (12)	1 (8)
Caregiving burden	5 (42)	2 (17)	3 (25)

[a]Clinically significant improvement and worsening was defined as an unadjusted standardized change of ≥0.5 SD from baseline to follow-up.

[b]Net improvement = participants who improved – participants who worsened.

Table 5. Benefits of participating in the study (N=12).

Question or response categories	Participants, n (%)
1. Did your participation in the study help you better understand Lewy Body Dementia?	
Not at all	0 (0)
Some	3 (25)
A great deal	9 (75)
2. Did your participation in the study help you feel more confident in dealing with difficult behavior of the care recipient?	
Not at all	1 (8)
Some	5 (42)
A great deal	6 (50)
3. Did your participation in the study help make your life easier?	
Not at all	2 (17)
Some	7 (58)
A great deal	3 (25)
4. Did your participation in the study help your ability to care for care recipient?	
Not at all	1 (8)
Some	7 (58)
A great deal	4 (33)

Discussion

Principal Findings

The study generated promising preliminary feasibility and efficacy data for a fully remote intervention designed specifically for family caregivers of persons with LBD. Enrollment and retention were successful, with the study experiencing only a 20% dropout rate. Our study showed retention rates similar to or better than those reported in other recent studies of digital interventions for family caregivers of persons with dementia [14]. A pooled estimate from recent studies that mainly recruited spousal caregivers [9-13] reported an average 66% retention rate, compared to the 80% retention

rate reported in this study. Moreover, in the exit interviews, participants confirmed that the experience of participating in the study was valuable and motivated them to continue participating. As the study was not powered for statistical significance, we observed nominal average and net improvements in important psychological outcomes. Additionally, many caregivers reported that study participation helped them better understand the disease, feel more confident in dealing with difficult behaviors of the care recipient, and improve their ability to care for the care recipient. These encouraging results concerning improvement in psychological measures are on par with those of the hallmark REACH II trial [35], which, according to the recent AHRQ [36] and NASEM [37] reports, meets the threshold for an evidence-based intervention for family caregivers of a person with dementia. Given that VOCALE-LBD and the REACH II protocol share similar elements, such as the provision of information, didactic instruction, role-playing, problem-solving, and skills training, in this era of increased demand for fully remote interventions, this pilot program seems well suited for further research and evaluation.

In this pilot study, we found nominal average and net improvements in important psychological outcomes, similar to changes observed in the hallmark REACH II trial. Specifically using the same approach to calculate clinically significant net improvements, the REACH II trial reported an approximately 30% improvement in depression and 10% in burden in a comparable intervention subgroup in terms of race and ethnicity. Moreover, in terms of benefits of participation, 58% of the participants in the White/Caucasian REACH II intervention subgroup asserted that the intervention helped them understand memory loss a great deal, and similarly, 59% of them asserted that the intervention helped them to some extent to feel more confident in dealing with the care recipient. Our results of 50% and 30% net improvement in depression and burden, respectively, and 75% and 50% endorsements in benefits for participation fall within the upper bound of the REACH II results. As such, pending further research, one cautious interpretation is that the VOCALE-LBD remote intervention for family caregivers of persons with LBD holds promise for further development and evaluations.

Despite the full-time caregiving demands of our study participants, all but 3 participants completed the 8-week intervention (9/12, 80%). The platform is designed to promote convenience and accessibility, allowing the user to engage within the weekly interval at times of day and frequencies based on their own preferences. We are aware of only 5 web-based intervention studies since 2015 in the dementia caregiving context that recruited mostly spousal caregivers. A study by Pot et al [9] included approximately 60% spousal caregivers and had a 44% retention rate. A study by Gustafson et al [10] included approximately 90% spousal caregivers and had an 84% retention rate. A study by Boots et al [11] included approximately 98% spousal caregivers and had a 61% retention rate. A study by Blom et al [12] included approximately 58% spousal caregivers and had a 70% retention rate. A study by Griffiths et al [13] included approximately 73% spousal caregivers and had a 73% retention rate. There may be various factors accounting for the 80% retention rate and engagement throughout this study. At the outset, because participants were recruited from a study registry, they are likely to demonstrate greater motivation and desire to participate. However, participants also commented that they appreciated the sense of community and connection among their peers, as well as the local context of the intervention (all participants were from the Pacific Northwest, with almost all from Washington state). On several occasions, participants discussed local resources and programs and even suggested plans to meet after the end of the study.

Limitations

This study has some limitations, which also suggest potential future directions. The effects of the intervention, although promising, must be interpreted with caution as the study was not powered for significance and did not include a control group. Moreover, the study focused on immediate health benefits without evaluating whether the changes were sustained over time. Eligible participants were part of a research registry and needed to have access to the internet. Although today, almost 70% of older adults have access to the internet [38], those who do not may respond differently to the intervention. A lack of racial or ethnic and gender diversity is a notable limitation, and there is a need to explore how caregivers of persons with LBD with other racial or ethnic and gender characteristics might experience VOCALE-LBD. Some participants had trouble relating to the personas and expressed a desire for skill enactment to be more focused on their real-life experiences, suggesting a need in future work to incorporate skill enactment activities more into participants' own lives. Finally, this study was subject to the common limitations of research involving self-report.

Conclusions

In this study, we performed a pilot evaluation of an innovative digital intervention that empowers caregivers of persons with LBD by providing a support network and enabling them to develop and improve problem-solving skills that they can use for daily challenges. If validated in future studies, the intervention could be an accessible, on-demand resource for caregivers to engage in moderated remote discussions with their peers at their own convenience in terms of location, time of the day, and frequency. The intervention could be used in conjunction with caregiver usual care and as a stand-alone module in circumstances, such as current and future pandemic emergencies when routine professional interventions might not be readily available.

Acknowledgments

This study was supported by the innovation grant program of the University of Washington Garvey Institute for Brain Health Solutions. The ADRC Research Registry was supported by the National Institute on Aging Grant P30AG066509.

Conflicts of Interest

KDR receives funding from the Lewy Body Dementia Association Research Centers of Excellence.

References

1. Walker Z, Possin KL, Boeve BF, Aarsland D. Lewy body dementias. The Lancet 2015 Oct;386(10004):1683-1697. [doi: 10.1016/s0140-6736(15)00462-6]
2. Taylor J, McKeith IG, Burn DJ, Boeve BF, Weintraub D, Bamford C, et al. New evidence on the management of Lewy body dementia. The Lancet Neurology 2020 Feb;19(2):157-169. [doi: 10.1016/s1474-4422(19)30153-x]
3. Svendsboe E, Terum T, Testad I, Aarsland D, Ulstein I, Corbett A, et al. Caregiver burden in family carers of people with dementia with Lewy bodies and Alzheimer's disease. Int J Geriatr Psychiatry 2016 Sep;31(9):1075-1083. [doi: 10.1002/gps.4433] [Medline: 26765199]
4. Lowery K, Mynt P, Aisbett J, Dixon T, O'Brien J, Ballard C. Depression in the carers of dementia sufferers: a comparison of the carers of patients suffering from dementia with Lewy bodies and the carers of patients with Alzheimer's disease. J Affect Disord 2000 Jul;59(1):61-65. [doi: 10.1016/s0165-0327(99)00123-8] [Medline: 10814772]
5. Galvin JE, Duda JE, Kaufer DI, Lippa CF, Taylor A, Zarit SH. Lewy body dementia: caregiver burden and unmet needs. Alzheimer Dis Assoc Disord 2010;24(2):177-181 [FREE Full text] [doi: 10.1097/WAD.0b013e3181c72b5d] [Medline: 20505434]
6. Killen A, Macaskill A. Using a Gratitude Intervention to Enhance Well-Being in Older Adults. J Happiness Stud 2014 Jun 20;16(4):947-964. [doi: 10.1007/s10902-014-9542-3]
7. Boutoleau-Bretonnière C, Pouclet-Courtemanche H, Gillet A, Bernard A, Deruet AL, Gouraud I, et al. The Effects of Confinement on Neuropsychiatric Symptoms in Alzheimer's Disease During the COVID-19 Crisis. J Alzheimers Dis 2020;76(1):41-47. [doi: 10.3233/JAD-200604] [Medline: 32568211]
8. Docherty AB, Harrison EM, Green CA, Hardwick H, Pius R, Norman L, ISARIC4C Investigators, et al. Features of 16,749 hospitalised UK patients with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol. medRxiv Preprint posted online April 28, 2020. [doi: 10.1101/2020.04.23.20076042]
9. Pot AM, Blom MM, Willemse BM. Acceptability of a guided self-help Internet intervention for family caregivers: mastery over dementia. Int Psychogeriatr 2015 Aug;27(8):1343-1354. [doi: 10.1017/S1041610215000034] [Medline: 25648589]
10. Gustafson DH, Gustafson DH, Cody OJ, Chih M, Johnston DC, Asthana S. Pilot Test of a Computer-Based System to Help Family Caregivers of Dementia Patients. J Alzheimers Dis 2019;70(2):541-552. [doi: 10.3233/JAD-190052] [Medline: 31256126]
11. Boots LM, de Vugt ME, Smeets CM, Kempen GI, Verhey FR. Implementation of the Blended Care Self-Management Program for Caregivers of People With Early-Stage Dementia (Partner in Balance): Process Evaluation of a Randomized Controlled Trial. J Med Internet Res 2017 Dec 19;19(12):e423 [FREE Full text] [doi: 10.2196/jmir.7666] [Medline: 29258980]
12. Blom MM, Zarit SH, Groot Zwaaftink RBM, Cuijpers P, Pot AM. Effectiveness of an Internet intervention for family caregivers of people with dementia: results of a randomized controlled trial. PLoS One 2015;10(2):e0116622 [FREE Full text] [doi: 10.1371/journal.pone.0116622] [Medline: 25679228]
13. Griffiths PC, Whitney MK, Kovaleva M, Hepburn K. Development and Implementation of Tele-Savvy for Dementia Caregivers: A Department of Veterans Affairs Clinical Demonstration Project. Gerontologist 2016 Feb;56(1):145-154. [doi: 10.1093/geront/gnv123] [Medline: 26566806]
14. Wu K, Su Y, Chu F, Chen AT, Zaslavsky O. Behavioral Change Factors and Retention in Web-Based Interventions for Informal Caregivers of People Living With Dementia: Scoping Review. J Med Internet Res 2022 Jul 07;24(7):e38595 [FREE Full text] [doi: 10.2196/38595] [Medline: 35797100]
15. Peterson B, Armstrong M, Galasko D, Galvin JE, Goldman J, Irwin D, et al. Lewy Body Dementia Association's Research Centers of Excellence Program: Inaugural Meeting Proceedings. Alzheimers Res Ther 2019 Mar 13;11(1):23 [FREE Full text] [doi: 10.1186/s13195-019-0476-1] [Medline: 30867052]
16. Teng AK, Han S, Lin S, Demiris G, Zaslavsky O, Chen AT. Using an Innovative Discussion Platform to Give Voice to Aging-Related Experiences: A Pilot Study. J Gerontol Nurs 2019 Dec 01;45(12):33-40. [doi: 10.3928/00989134-20191105-05] [Medline: 31755541]
17. Chen AT, Chu F, Teng AK, Han S, Lin S, Demiris G, et al. Promoting Problem Solving About Health Management: A Mixed-Methods Pilot Evaluation of a Digital Health Intervention for Older Adults With Pre-Frailty and Frailty. Gerontol Geriatr Med 2021;7:2333721420985684 [FREE Full text] [doi: 10.1177/2333721420985684] [Medline: 33457461]

18. Chen AT, Ge S, Cho S, Teng AK, Chu F, Demiris G, et al. Reactions to COVID-19, information and technology use, and social connectedness among older adults with pre-frailty and frailty. Geriatr Nurs 2021 Jan;42(1):188-195 [FREE Full text] [doi: 10.1016/j.gerinurse.2020.08.001] [Medline: 32863038]
19. Armstrong MJ, Alliance S, Taylor A, Corsentino P, Galvin JE. End-of-life experiences in dementia with Lewy bodies: Qualitative interviews with former caregivers. PLoS ONE 2019 May 30;14(5):e0217039. [doi: 10.1371/journal.pone.0217039]
20. Killen A, Flynn D, De Brún A, O'Brien N, O'Brien J, Thomas AJ, et al. Support and information needs following a diagnosis of dementia with Lewy bodies. Int Psychogeriatr 2016 Mar;28(3):495-501. [doi: 10.1017/S1041610215001362] [Medline: 26328546]
21. Nezu A, Nezu C, D'Zurilla T. Problem-Solving Therapy. In: Kazantzis N, Reinecke MA, Freeman A, editors. Cognitive and Behavioral Theories in Clinical Practice. New York, NY: Guilford Publications; 2010:76-114.
22. Demiris G, Parker Oliver D, Wittenberg-Lyles E, Washington K, Doorenbos A, Rue T, et al. A noninferiority trial of a problem-solving intervention for hospice caregivers: in person versus videophone. J Palliat Med 2012 Jun;15(6):653-660 [FREE Full text] [doi: 10.1089/jpm.2011.0488] [Medline: 22536989]
23. Washington KT, Demiris G, Parker Oliver D, Albright DL, Craig KW, Tatum P. Delivering problem-solving therapy to family caregivers of people with cancer: A feasibility study in outpatient palliative care. Psychooncology 2018 Oct;27(10):2494-2499 [FREE Full text] [doi: 10.1002/pon.4859] [Medline: 30107070]
24. Kirkham JG, Choi N, Seitz DP. Meta-analysis of problem solving therapy for the treatment of major depressive disorder in older adults. Int J Geriatr Psychiatry 2016 May 05;31(5):526-535. [doi: 10.1002/gps.4358] [Medline: 26437368]
25. Pruitt J, Grudin J. Personas: practice and theory. 2003 Presented at: DUX03: Designing the User Experience; June 6-7, 2003; San Francisco, CA. [doi: 10.1145/997078.997089]
26. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
27. Radloff LS. The CES-D Scale. Applied Psychological Measurement 2016 Jul 26;1(3):385-401. [doi: 10.1177/014662167700100306]
28. Bédard M, Molloy D, Squire L, Dubois S, O'Donnell M. Zarit Burden Interview--Screening Version [Database record]. APA PsycTests 2001. [doi: 10.1037/t35628-000]
29. Ezzati A, Jiang J, Katz MJ, Sliwinski MJ, Zimmerman ME, Lipton RB. Validation of the Perceived Stress Scale in a community sample of older adults. Int J Geriatr Psychiatry 2014 Jun 03;29(6):645-652 [FREE Full text] [doi: 10.1002/gps.4049] [Medline: 24302253]
30. Russell D, Peplau LA, Cutrona CE. The revised UCLA Loneliness Scale: Concurrent and discriminant validity evidence. Journal of Personality and Social Psychology 1980 Sep;39(3):472-480. [doi: 10.1037/0022-3514.39.3.472]
31. Sherbourne CD, Stewart AL. The MOS social support survey. Soc Sci Med 1991 Jan;32(6):705-714. [doi: 10.1016/0277-9536(91)90150-B]
32. Lee SY, Hwang H, Hawkins R, Pingree S. Interplay of Negative Emotion and Health Self-Efficacy on the Use of Health Information and Its Outcomes. Commun Res 2008 Apr 14;35(3):358-381. [doi: 10.1177/0093650208315962]
33. Lykens K, Moayad N, Biswas S, Reyes-Ortiz C, Singh KP. Impact of a community based implementation of REACH II program for caregivers of Alzheimer's patients. PLoS One 2014 Feb 27;9(2):e89290 [FREE Full text] [doi: 10.1371/journal.pone.0089290] [Medline: 24586664]
34. Bradbury K, Morton K, Band R, van Woezik A, Grist R, McManus RJ, et al. Using the Person-Based Approach to optimise a digital intervention for the management of hypertension. PLoS One 2018;13(5):e0196868 [FREE Full text] [doi: 10.1371/journal.pone.0196868] [Medline: 29723262]
35. Belle SH, Burgio L, Burns R, Coon D, Czaja SJ, Gallagher-Thompson D, Resources for Enhancing Alzheimer's Caregiver Health (REACH) II Investigators. Enhancing the quality of life of dementia caregivers from different ethnic or racial groups: a randomized, controlled trial. Ann Intern Med 2006 Nov 21;145(10):727-738 [FREE Full text] [doi: 10.7326/0003-4819-145-10-200611210-00005] [Medline: 17116917]
36. Butler M, Gaugler J, Talley K, Abdi H, Desai P, Duval S, et al. Care Interventions for People Living With Dementia and Their Caregivers. Agency for Healthcare Research and Quality. 2020 Aug 31. URL: https://effectivehealthcare.ahrq.gov/products/care-interventions-pwd/report [accessed 2022-07-15]
37. National Academies of Sciences, Engineering, and Medicine. Meeting the Challenge of Caring for Persons Living with Dementia and Their Care Partners and Caregivers: A Way Forward. Washington, DC: The National Academies Press; 2021.
38. Anderson M, Perrin A. 1. Technology use among seniors. Pew Research Center. 2017 May 17. URL: https://www.pewresearch.org/internet/2017/05/17/technology-use-among-seniors/ [accessed 2019-10-30]

Abbreviations

ADRC: Alzheimer's Disease Research Center
CES-D: Center for Epidemiologic Studies Depression

DLB: dementia with Lewy bodies
IRB: institutional review board
LBD: Lewy body dementia
MBWC: Memory and Brain Wellness Center
PDD: Parkinson disease dementia
PST: problem-solving therapy
REDCap: Research Electronic Data Capture
UW: University of Washington
VOCALE-LBD: Virtual Online Communities for Aging Life Experience–Lewy Body Dementia

Edited by A Mavragani; submitted 07.02.22; peer-reviewed by A Ramamurthy, A Killen; comments to author 27.06.22; revised version received 13.07.22; accepted 13.07.22; published 29.07.22.

Please cite as:
Zaslavsky O, Kaneshiro J, Chu F, Teng A, Domoto-Reilly K, Chen AT
Virtual Intervention for Caregivers of Persons With Lewy Body Dementia: Pilot Quasi-Experimental Single-Arm Study
JMIR Form Res 2022;6(7):e37108
URL: https://formative.jmir.org/2022/7/e37108
doi:10.2196/37108
PMID:35904843

Original Paper

Comparison of the Impact of Insulin Degludec U100 and Insulin Glargine U300 on Glycemic Variability and Oxidative Stress in Insulin-Naive Patients With Type 2 Diabetes Mellitus: Pilot Study for a Randomized Trial

Pavle Vrebalov Cindro[1], MD; Mladen Krnić[2,3], MD, PhD; Darko Modun[4], MD, PhD; Jonatan Vuković[1], MD, PhD; Tina Tičinović Kurir[2,3], MD, PhD; Goran Kardum[5], MD, PhD; Doris Rušić[4], PharmD; Ana Šešelja Perišin[4], PharmD; Josipa Bukić[4], PharmD

[1]Department of Gastroenterology, University Hospital Split, Split, Croatia

[2]Department of Endocrinology, University Hospital Split, Split, Croatia

[3]Department of Pathophysiology, School of Medicine, University of Split, Split, Croatia

[4]Department of Pharmacy, School of Medicine, University of Split, Split, Croatia

[5]Department of Psychology, Faculty of Humanities and Social Sciences, University of Split, Split, Croatia

Corresponding Author:
Mladen Krnić, MD, PhD
Department of Endocrinology
University Hospital Split
Šoltanska 1
Split, 21000
Croatia
Phone: 385 981772113
Email: mladen.krnic@gmail.com

Abstract

Background: There is an ongoing discussion about possible differences between insulin degludec (IDeg-100) and glargine U300 (IGlar-300). There is little data and head-to-head comparison of IDeg-100 and IGlar-300 regarding their simultaneous impact on glycemic variability and oxidative stress in patients with type 2 diabetes mellitus (T2DM).

Objective: In our randomized, open-label, crossover study, we compared the impact of IDeg-100 and IGlar-300 on glycemic variability and oxidative stress in insulin-naive patients with T2DM.

Methods: We recruited a total of 25 adult patients with T2DM (7 females) whose diabetes was uncontrolled (HbA_{1c} ≥7.5%) on two or more oral glucose-lowering drugs; a total of 22 completed the study. Mean age was 57.3 (SD 6.99) years and duration of diabetes was 9.94 (SD 5.01) years. After the washout period, they were randomized alternately to first receive either IDeg-100 or IGlar-300 along with metformin. Each insulin was administered for 12 weeks and then switched. At the beginning and end of each phase, biochemical and oxidative stress parameters were analyzed. On 3 consecutive days prior to each control point, patients performed a 7-point self-monitoring of blood glucose profile. Oxidative stress was assessed by measuring thiol groups and hydroperoxides (determination of reactive oxygen metabolites test) in serum.

Results: IGlar-300 reduced mean glucose by 0.02-0.13 mmol/L, and IDeg-100 reduced glucose by 0.10-0.16 mmol/L, with no significant difference. The reduction of the coefficient of glucose variation also did not show a statistically significant difference. IGlar-300 increased thiols by 0.08 µmol/L and IDeg-100 increased thiols by 0.15 µmol/L, with no significant difference (*P*=.07) between them. IGlar-300 reduced hydroperoxides by 0.040 CARR U and IDeg-100 increased hydroperoxides by 0.034 CARR U, but the difference was not significant (*P*=.12).

Conclusions: The results of our study do not show a significant difference regarding glycemic variability between patients receiving either insulin IDeg-100 or IGlar-300, although IGlar-300 showed greater dispersion of data. No significant difference in oxidative stress was observed. In a larger study, doses of insulins should be higher to achieve significant impact on glycemic parameters and consequently on glycemic variability and oxidative stress.

Trial Registration: ClinicalTrials.gov, NCT04692415; https://clinicaltrials.gov/ct2/show/NCT04692415

(JMIR Form Res 2022;6(7):e35655) doi:10.2196/35655

KEYWORDS

type two diabetes mellitus; type 2 diabetes mellitus; insulin degludec; insulin glargine U300; glucose variability; oxidative stress; insulin; diabetes; diabetic; glycemic variability; glycaemic variability; RCT; pilot; control trial; clinical trial

Introduction

Background

Global diabetes prevalence in 2019 is estimated to be 9.3% (463 million people), rising to 10.2% (578 million) by 2030 and 10.9% (700 million) by 2045 [1].

The main feature of diabetes mellitus of all types is dysglycemia, which consists of two main components: chronic sustained hyperglycemia and acute glycemic fluctuations from peaks to nadirs. Although disputed by some authors [2], it is generally considered that both components contribute to diabetes complications through two main mechanisms—excessive protein glycation and activation of oxidative stress [3]—with glycemic variability being more specific in having an effect on oxidative stress than chronic sustained hyperglycemia [4], as both upward (postprandial glucose increments) and downward (interprandial glucose decrements) changes activate the oxidative stress pathway [5].

Glucose fluctuations gradually increase from normal glucose metabolism to impaired glucose regulation and diabetes mellitus. Intraday glucose variability occurs at the early stage of abnormal glucose tolerance. In addition to elevated intraday glucose fluctuations, newly diagnosed, drug-naive patients with type 2 diabetes mellitus (T2DM) also demonstrate increased postprandial glucose excursions, higher glucose levels overnight, and more interday fluctuations [6].

The main purpose of insulin therapy in diabetes mellitus is to control glucose—in other words, to combat dysglycemia. Long-acting basal insulin analogues (insulin glargine U100, insulin detemir) significantly improved diabetes management, providing longer duration, flatter profiles of action, lower risk of hypoglycemia, and less glycemic variability compared to NPH (Neutral Protamine Hagedorn) insulin [7,8].

The second generation of basal insulin analogues—insulin degludec 100 units/mL (IDeg-100) and insulin glargine 300 units/mL (IGlar-300)—have even smoother pharmacokinetic/pharmacodynamic profiles than insulin glargine U100, are longer acting, and further lower glycemic variability, at least in patients with T1DM [9,10].

Although several studies [11-13] have compared the impact of these two second-generation basal insulin analogues on glycemic variability in patients with type 1diabetes mellitus (T1DM), there is little data and head-to-head comparison of IDeg-100 and IGlar-300 regarding their simultaneous impact on glycemic variability and oxidative stress in patients with T2DM. In addition, the results from the T1DM studies are inconsistent [12,13].

Aim of the Study

In this initial study, we compared the impact of IDeg-100 and IGlar-300 on glucose variability and oxidative stress (represented through its surrogate markers) in insulin-naive patients with T2DM. The main research question was whether there is any difference between the two insulins regarding these parameters. The results of this study should inform a larger study comparing these two insulins.

Methods

Ethics Approval and Consent to Participate

This randomized, open-label, crossover study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the University of Split School of Medicine (number 2181-198-03-04-17-0045). All subjects gave written consent prior to their participation in the study.

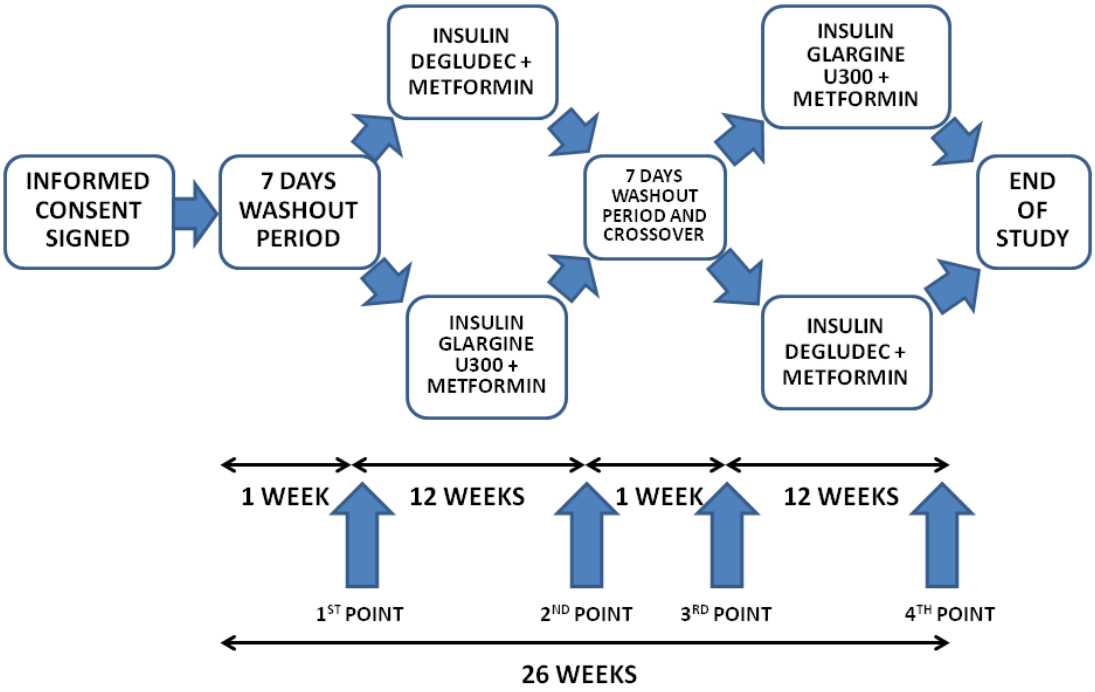
Study Protocol and Population

Between December 2018 and May 2019, we recruited 25 outpatient insulin-naive patients with uncontrolled T2DM (glycated hemoglobin [HbA_{1c}] ≥7.5% on two or more oral glucose-lowering drugs) and assigned them to either degludec insulin or glargine U300 insulin combined with metformin. All patients were recruited and treated at University Hospital Split, Croatia. All patients finished the study, but only 22 were analyzed (3 patients were excluded from data analysis—one patient did not perform his 7-point self-monitoring of blood glucose (SMBG) profile 3 days prior to control points, one patient decided to continue with his oral glucose-lowering agents only, and one patient left the study for personal reasons). Basal characteristics of the participants studied are shown in Table 1. The protocol of the study is shown in Figure 1. The study adheres to CONSORT (Consolidated Standards of Reporting Trials) guidelines (Multimedia Appendix 1 displays the CONSORT flow diagram). Patients and the public were not involved in the design, conduct, reporting, or dissemination plans of our research. The trial was retrospectively registered on ClinicalTrials.gov on December 31, 2020 (NCT04692415).

Table 1. Basal clinical characteristics of patients (N=25).

Parameter	Mean (SD)
Age (years)	56.23 (8.09)
Duration of diabetes (years)	8.90 (5.05)
Body weight (kg)	89.02 (13.82)
Body height (cm)	176.20 (10.03)
BMI (kg/m^2)	28.92 (3.89)
Waist circumference (cm)	102.84 (8.56)
Glycated hemoglobin (HbA_{1c}), %	9.66 (1.65)
Fasting glucose (mmol/L)	13.02 (4.47)
Serum creatinine (µmol/L)	68.24 (13.15)
Serum uric acid (µmol/L)	310.04 (60.75)
Total cholesterol (mmol/L)	5.28 (1.46)

Figure 1. The study protocol.



Patients who were eligible for the study fulfilled all of the following inclusion criteria: history of T2DM for at least 1 year, aged between 45 and 65 years (women obligatory in postmenopause and with no hormonal replacement), uncontrolled glycemia on two or more oral antidiabetic drugs, no prior use of insulin, HbA_{1c} ≥7.5%, receiving statins (if not on statins, they were put on it), not on antiaggregant therapy (if on antiaggregants, they were temporarily excluded from therapy).

Exclusion criteria were the following: the use of glitazones or anticoagulant therapy, renal impairment with creatinine clearance <60 mL/s, presence of malignant disease, chronic liver disease, severe cardiovascular disease and history of cardiovascular incidents (stroke, myocardial infarction, peripheral amputation), and rheumatic and autoimmune diseases.

All participants were asked to avoid the consumption of vitamin supplements, coffee, wine, and Coca-Cola and similar beverages, especially in the days before each control point. Patients were also asked to avoid intensive physical activity up to two days before each control point. All subjects were told to report any side effects immediately, and were given the telephone numbers of the study conductors for this purpose. At each control point, participants were asked about possible side effects.

At baseline, all patients discontinued their previous therapy and were given metformin alone (2 g/day) for 7 days (washout period). After the 7-day washout period, they were randomized alternately by investigators (1:1 ratio) to first receive either IDeg-100 or IGlar-300 subcutaneously according to the order they were included in the study. In phase one, they received either IDeg-100 or IGlar-300 combined with metformin for 12

weeks. Phase one was followed by a second washout period in which patients received metformin alone again for 7 days. Finally, in phase two, which also lasted for 12 weeks, patients were switched from IDeg-100 to IGlar-300 and vice versa (and metformin was continued). The initial dose of both insulins was 0.2 IU/kg. We did not change the dose of insulin during the study period to avoid hypoglycemia, which could significantly influence the results [14-16].

At the beginning and end of each phase, blood samples were collected for the analysis of standard biochemical and oxidative stress parameters (control points 1-4). On 3 consecutive days prior to each control point (at the beginning and the end of each phase), patients completed the 7-point SMBG profile. All patients were already experienced with the use of SMBG [17].

Glucose Measurement

To standardize results, all patients received a standard Bionime GM550 glucose meter. They were asked to regularly check their blood glucose 1-2 times per day during the entire study and, in the 3 consecutive days prior to each control point, to perform the 7-point SMBG profile. The 7-point blood glucose profile consisted of seven measurements: (1) before breakfast, (2) 2 hours after breakfast, (3) before lunch, (4) 2 hours after lunch, (5) before dinner, (6) 2 hours after dinner, and (7) before sleeping. Glucose variability was determined by calculating mean glycemia, SD, and coefficient of variation (CV) for each control point [17,18].

Standard Laboratory Measurement

Serum uric acid concentrations, serum creatinine, total cholesterol, serum bilirubin values, and other basic biochemical laboratory values were determined by Olympus AU 600 Chemistry Analyzer (Olympus Michima Co Ltd) and enzymatic laboratory kit.

Oxidative Stress Measurement

Thiol groups were assayed according to the Ellman assay [19], modified by Hu [20]. In detail, 100 µL of plasma was diluted with 2 mL of Tris-EDTA buffer (0.1 mol/L Tris,1 mmol/L EDTA, pH 8.2), and mixed with 100 µL of 10 mM DTNB (5,5′-dithiobis(2-nitrobenzoic acid)), previously prepared in methanol. To subtract the absorbance of plasma and DTNB at 412 nm, two parallel blank samples were assembled. The first one (“blank sample”) was prepared by mixing 2.1 mL of Tris-EDTA buffer with 100 µL of plasma and the second one (“blank reagent”) was prepared by mixing 2.1 mL of Tris-EDTA buffer with 100 µL of DTNB. All measurements were performed in triplicate and blanks were run for each sample. Readings were taken spectrophotometrically (Lambda 25; Perkin Elmer) at 412 nm after 15 minutes of reaction at 25 °C. Results were compared with a standard curve prepared daily with different concentrations of glutathione and expressed as µmol/L of glutathione.

The determination of reactive oxygen metabolites (d-ROM) assay measures the concentration of total hydroperoxides in serum or heparin plasma. The method was first described by Alberti et al in 1999 [21] and modified by Verde et al [22], and this modified assay was used to determine d-ROM values in plasma in this study. Each sample was prepared by mixing 2 mL of 0.1 M sodium acetate buffer (pH 4.8) and 20 µL of 0.1 M DMPD (N,N-diethyl-p-phenylenediamine) with 10 µL of plasma. After preparation, tubes with samples were vortexed for 15 seconds and incubated in a thermomixer (Thermomixer comfort, Eppendorf) at 37 °C and 1000 revolutions per minute for 75 minutes. All measurements were performed in triplicate and blanks were run for each sample. Readings were taken spectrophotometrically (Lambda 25; Perkin Elmer) at 505 nm after incubation. Results were compared with a standard curve prepared daily with different concentrations of H_2O_2. The results are expressed in CARR U (Carratelli units), where 1 CARR U corresponds to 0.08 mg/100 mL H_2O_2.

Statistical Analysis

The number of subjects to include in the protocol was selected according to previously published literature [4,23]. Statistical analyses were performed using Statistica 6.0 (StatSoft Inc). Two-way ANOVA for repeated measures was used to evaluate changes in plasma glucose levels, CV, plasma thiols, and hydroperoxides due to IDeg-100 and IGlar-300.

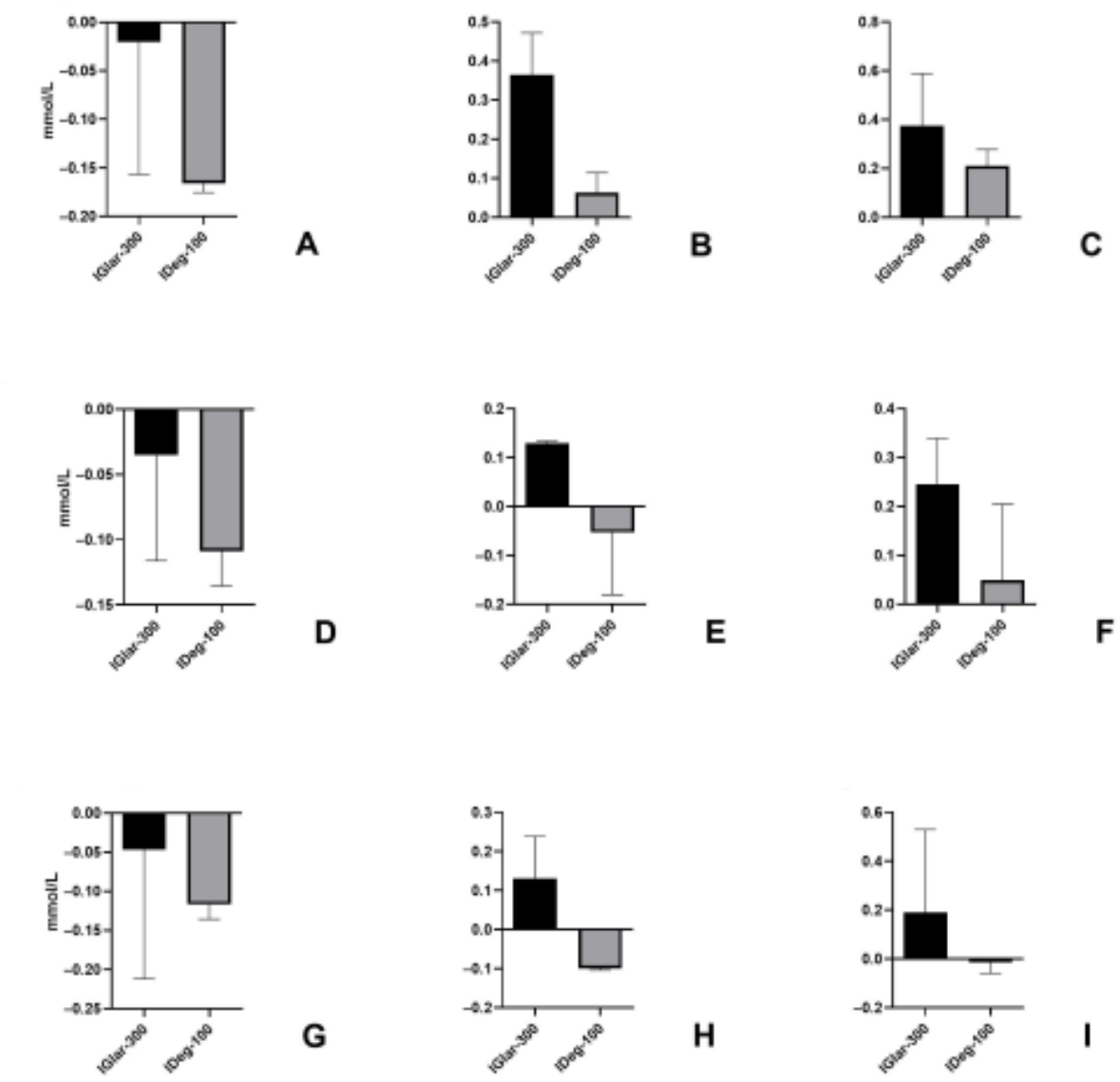
Results

Out of 25 randomized patients, 3 patients were excluded from data analysis—one patient did not perform his 7-point SMBG profile 3 days prior to control points, one patient decided to continue with his oral glucose-lowering agents only, and one patient left the study for personal reasons. No adverse reactions or unintended effects were noticed.

A total of 22 patients (7 females) successfully completed the trial, and their mean basal values were as follows: age, 57.3 (SD 6.99) years; duration of diabetes, 9.94 (SD 5.01) years; body weight, 88.25 (SD 13.57) kg; body height, 174.95 (SD 9.67) cm; BMI, 29.10 (SD 3.80) kg/m^2; waist circumference, 102.73 (SD 8.02) cm; HbA_{1c}, 9.60% (SD 1.68%); fasting glucose, 13.20 (SD 4.48) mmol/L; serum creatinine, 66.0 (SD 2.09) µmol/L; serum uric acid, 305.32 (SD 62.60) µmol/L; total cholesterol, 5.05 (SD 1.12) mmol/L.

On the first of 3 consecutive days of 7-point SMBG, performed at the end of the observed period, IGlar-300 and IDeg-100 reduced mean glucose values by 0.02 and 0.16 mmol/L, respectively, which was statistically insignificant (P=.06; 95% CI 0.003 to 0.28); there was also no significant difference between the two insulins (P=.17; 95% CI –0.10 to 0.46) (Figure 2A).

Figure 2. Mean glucose change, SD, and CV on the first, second, and third day. CV: coefficient of variation; IDeg-100: insulin degludec; IGlar-300: insulin glargine U300. A) Mean glucose change on the first day. B) SD of glucose change on the first day. C) CV of glucose change on the first day. D) Mean glucose change on the second day. E) SD of glucose change on the second day. F) CV of glucose change on the second day. G) Mean glucose change on the third day. H) SD of glucose change on the third day. I) CV of glucose change on the third day.



The SD of glucose excursions was 0.36 for IGlar-300 and 0.06 for IDeg-100, which was not significant (*P*=.20; 95% CI –0.27 to 0.87); in addition, there was no statistically significant difference between the two insulins (*P*=.07; 95% CI –1.07 to 1.22) (Figure 2B)

The CV on the first day was 0.37 (37%) for IGlar-300 and 0.21 for IDeg-100, which was statistically insignificant (*P*=.22; 95% CI –0.30 to 0.63). When compared, CV for these two insulins was not significantly different (*P*=.20; 95% CI –0.73 to 1.15) (Figure 2C).

On the second of the 3 days of 7-point SMBG, performed at the end of the observed period, IGlar-300 and IDeg-100 reduced mean glucose by 0.03 and 0.10 mmol/L, respectively, which was statistically insignificant (*P*=.08; 95% CI –0.09 to 0.24); there was also no significant difference between the two insulins (*P*=.07; 95% CI –0.41 to 0.26) (Figure 2D).

The SD of glucose excursions on the second day was 0.12 for IGlar-300 and –0.05 for IDeg-100, which was insignificant (*P*=.19; 95% CI –0.23 to 0.59); when we compared the SD of the two insulins, there was no significant difference (*P*=.17; 95% CI –1.00 to 0.65) (Figure 2E).

The CV for the second day was 0.24 (24%) for IGlar-300 and 0.04 for IDeg-100 and that was statistically insignificant (*P*=.20; 95% CI –0.24 to 0.63). When compared, CV for the two insulins was not significantly different (*P*=.08; 95% CI –0.96 to 0.78) (Figure 2F).

On the third (last) day of the SMBG, the insulins reduced mean glucose levels by 0.04 (IGlar-300) and 0.11 mmol/L (IDeg-100), which was statistically insignificant (*P*=.08; 95% CI –0.10 to 0.24), again without a significant difference between the two insulins (*P*=.20; 95% CI –0.55 to 0.14) (Figure 2G).

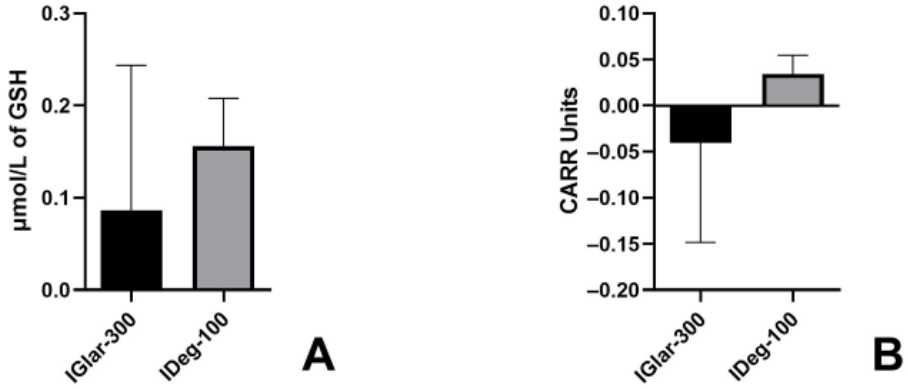
The SD of glucose concentrations on the third day was 0.13 for IGlar-300 and –0.09 for IDeg-100, which was insignificant (*P*=.18; 95% CI –0.15 to 0.61), and comparison of the SD of the two insulins revealed no statistical difference (*P*=.14; 95% CI –0.62 to 0.91) (Figure 2H).

The CV on the third day was 0.19 for IGlar-300 and –0.01 for IDeg-100, which was statistically insignificant (*P*=.16; 95% CI –0.14 to 0.55). When compared, CV for these two insulins was not statistically different (*P*=.54; 95% CI –1.23 to 0.14) (Figure 2I).

IGlar-300 increased thiol levels by 0.08 µmol/L and IDeg-100 increased thiol levels by 0.15 µmol/L (*P*=.07; 95% CI –0.21 to 0.08). No significant difference was found between the two insulins regarding the increase of thiols (*P*=.14; 95% CI –0.15 to 0.44) (Figure 3A).

Although IGlar-300 decreased hydroperoxides by 0.04 CARR U, and IDeg-100 increased hydroperoxides by 0.034 CARR U (*P*=.06; 95% CI –0.19 to 0.05), this impact was not statistically significant, and there was no significant difference between the two insulins (*P*=.12; 95% CI –0.13 to 0.37) (Figure 3B).

Figure 3. Changes in plasma thiols and hydroperoxides. GSH: glutathione; IDeg-100: insulin degludec; IGlar-300: insulin glargine U300. A) Changes in plasma thiols. B) Changes in plasma hydroperoxides.



Discussion

Principal Findings

Since the new generation of insulin analogues, degludec and glargine U300, appeared on the market, there has been an ongoing discussion about the possible advantages of one insulin over the other. The majority of comparisons related to the incidence of hypoglycemia [7,19,21-23], as well as the absorption stability and profile flatness, as possible causes of differences in hypoglycemia tendency were compared [7,13,24].

Absorption stability and profile flatness, if different, should lead to a difference in glycemic variability between the two insulins as variability in insulin absorption represents an important source of glucose variability in these subjects. Variability can relate to the insulin preparation, the injection technique, and the individual [25].

Glycemic variability and the incidence of hypoglycemia are the elements that "upgraded" the diabetological paradigm from the "diabetic triad" (HbA_{1c}, fasting glycemia, and postprandial glycemia) to the "diabetic pentad" (HbA_{1c}, fasting glycemia, postprandial glycemia, hypoglycemia, and glycemic variability) or even "hexad," if quality of life is included [26].

The final consequences of increased glycemic variability and long-lasting hyperglycemia, as mentioned in the Introduction section, are diabetic complications, both micro- and macrovascular. As glycemic variability contributes to an increase in oxidative stress (one of two main mechanisms leading to the development of diabetic complications) to a greater extent, the need for the comparison of glycemic variability measures and oxidative stress markers in patients exposed to IDeg-100 and IGlar-300 emerges.

Two large, recently published studies (BRIGHT and DELIVER D+), although different in design, found "more similarities than differences" using IDeg-100 and IGlar-300. On the other hand, another large study (CONFIRM) attributed some advantages to IDeg-100. However, these studies, with the exception of the BRIGHT study, focused primarily on hypoglycemia rate and HbA_{1c} outcomes [27-29]. In this study, we wanted to associate glycemic variability with its ultimate consequence—oxidative stress.

Oxidative stress causes the development of diabetic complications through 5 main molecular mechanisms: the polyol pathway, the hexosamine pathway, increased formation of advanced glycation end products, increased expression of the receptors for advanced glycation end products and their activating ligands, and activation of protein kinase C isoforms. In addition, oxidative stress negatively influences the antiatherosclerotic endothelial enzymes: endothelial nitric oxide synthase and prostacyclin synthase. The intracellular reactive oxygen species increase by these mechanisms, then lead to defective angiogenesis in response to ischemia, and activate proinflammatory and epigenetic mechanisms after the normalization of glycemia ("hyperglycemic memory") [30].

We used a basal-supported oral therapy variant in this study to emphasize the impact of the insulins studied. However, the results of this initial study did not show statistically significant differences both in glycemic variability and in the expression of oxidative stress in our patients. The primary reason for this was that the decrease in glycemic parameters was too small and,

consequently, the impact on glycemic variability and oxidative stress was too weak. The too-small decrease in glycemic parameters was a consequence of using a low dose of insulin. Namely, we administered both insulins at a dose of 0.2 IU per kg of body weight, and we did not titrate the dose for two reasons: to avoid hypoglycemia, which could significantly influence the oxidative stress and glycemic variability results, and to eschew the difference in dosing of two insulins, which could also affect the results. Many studies have shown that hypoglycemia can worsen oxidative stress through, among other mechanisms, a decrease in nitric oxide and a "reperfusion-like" effect [14-16]. If an examinee has experienced hypoglycemia, he or she should be excluded from the study. Nevertheless, in future studies, the dose of insulins administered must be higher. The question is if "treat-to-target" is the best therapy approach, as treating to target inevitably lowers blood sugar toward hypoglycemia. We think that an initial dose of 0.4 IU/kg with very careful titration of the dose will achieve a more desirable glucose level, but still far from the hypoglycemic zone. A smaller-scale titration study to optimize the dose of insulin and the optimal time for oxidative marker readout would be a good in-between step.

Some studies showed a lower incidence of hypoglycemia with IDeg-100 versus IGlar-300, and that also could contribute to the lower levels of variability and oxidative stress observed with degludec [27], although other studies showed no difference between IDeg-100 and IGlar-300 in that regard [28,29]. Previous research has also suggested a somewhat greater potency of IDeg-100, thus titration to target would probably lead to differences in the final doses used and, consequently, make the comparison more difficult [24,28].

A longer exposure period (longer than 12 weeks) would allow for a more expressed impact of each insulin, presumed positive, although some studies showed a negative effect of chronic insulin therapy on oxidative stress [31]. The longer exposure would possibly explain the difference in the simultaneous increase in thiols and hydroperoxides produced by IDeg-100. The increase in thiol group concentrations represents protein oxidative stress reduction and d-ROM gives insight into the acute changes of lipid peroxide oxidation. This should be considered in the context of the negative effect of chronic insulin treatment on oxidative stress, as mentioned above.

We assessed within-day glycemic variability through changes in average glucose levels, SD of glycemic excursions, and the CV, derived from the SD. The 7-point SMBG profile represents the standard method of glucose monitoring, and we used it for the assessment of glycemic variability [17,32]. Diagnostic CGMS (continuous glucose monitoring systems) and the Libre Flash monitoring system would be more precise tools for glycemic variability measurement but, unfortunately, at the time, due to financial reasons, they could not be employed in this initial study. CGMS or the Flash monitoring system would allow the detection of a greater number of daily peaks and nadirs and give a better insight into glycemic variability [33]. Using CGMS, it would be possible to use the mean amplitude of glucose excursions (MAGE) index as an assessment tool for glycemic variability as well. Hence, in future studies, we highly recommend the use of CGMS.

In this study, we recruited insulin-naive patients who experienced the failure of oral glucose-lowering therapy and needed insulin introduction. We excluded those who were previously on pioglitazone therapy because of the prolonged action of this drug (which was impossible to remove during the 7-day washout period) and its possible influence on the results. Moreover, we standardized the concomitant therapy by introducing the same dose of atorvastatin in all patients and by temporarily removing salicylic acid from the therapy.

Conclusion

The results of this study do not show a statistically significant difference in glycemic variability between IDeg-100 and IGlar-300. An insufficient dose of insulin was the main reason for the lack of impact on glycemic parameters and, consecutively, on glycemic variability. Probably due to the absence of a difference in glycemic variability, no difference in the oxidative stress level was noticed. A full-scale study should use larger doses of insulins (at least 0.4 IU/kg), and an optimized and adjusted "treat-to-target" algorithm. CGMS should be used instead of the 7-point SMBG profile. The MAGE index derived from the CGMS should be used for the assessment of glycemic variability. Another small titration study could be performed for optimization of the insulin dose and calculation of the sample size for the main study.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors wish to thank Salvus d.o.o. Croatia for financial support and donation of glucose meters and test strips. The authors also thank Ms Tanja Vrebalov Cindro for language review.

Data Availability Statement

The data sets used and/or analyzed in this study are available from the corresponding author on reasonable request.

Authors' Contributions

MK contributed to study concept and design. PVC, MK, and TTK contributed to the acquisition of data. DM, DR, ASP, and JB contributed to laboratory analysis. MK, DM, PVC, and JV contributed to analysis and interpretation of data. MK, DM, and PVC contributed to drafting of the manuscript. MK, PVC, DM, and JV critically revised the manuscript for important intellectual

content. GK, DM, and JV contributed to statistical analysis. MK supervised the study. All authors have read and approved the final manuscript.

Conflicts of Interest

MK received honoraria from Sanofi, NovoNordisk, Eli Lilly, Abbott, MSD, Takeda, Novartis, Boehringer Ingelheim, Servier, Lifescan, and AstraZeneca as a speaker and for attendance at advisory boards. MK author has no conflicts of interest associated with this research.
TTK received honoraria from Sanofi, NovoNordisk, Eli Lilly, Abbott, MSD, Takeda, Novartis, Boehringer Ingelheim, Servier, Lifescan, and AstraZeneca as a speaker and for attendance at advisory boards. TTK author has no conflicts of interest associated with this research.
PVC received honoraria from Abbott, Teva, and Takeda as a speaker and a case presenter. PVC author has no conflicts of interest associated with this research.
JV, DM, GK, DR, ASP, and JB declare no conflicts of interest associated with this research.

Multimedia Appendix 1
CONSORT flow diagram.
[PDF File (Adobe PDF File), 420 KB - formative_v6i7e35655_app1.pdf]

Multimedia Appendix 2
CONSORT 2010 checklist.
[PDF File (Adobe PDF File), 145 KB - formative_v6i7e35655_app2.pdf]

References

1. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, IDF Diabetes Atlas Committee. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9 edition. Diabetes Res Clin Pract 2019 Nov;157:107843. [doi: 10.1016/j.diabres.2019.107843] [Medline: 31518657]
2. Siegelaar SE, Barwari T, Kulik W, Hoekstra JB, DeVries JH. No relevant relationship between glucose variability and oxidative stress in well-regulated type 2 diabetes patients. J Diabetes Sci Technol 2011 Jan 01;5(1):86-92 [FREE Full text] [doi: 10.1177/193229681100500112] [Medline: 21303629]
3. Monnier L, Colette C. Glycemic variability: should we and can we prevent it? Diabetes Care 2008 Feb;31 Suppl 2:S150-S154. [doi: 10.2337/dc08-s241] [Medline: 18227477]
4. Monnier L, Mas E, Ginet C, Michel F, Villon L, Cristol J, et al. Activation of oxidative stress by acute glucose fluctuations compared with sustained chronic hyperglycemia in patients with type 2 diabetes. JAMA 2006 Apr 12;295(14):1681-1687. [doi: 10.1001/jama.295.14.1681] [Medline: 16609090]
5. Monnier L, Colette C, Owens DR. Glycemic variability: the third component of the dysglycemia in diabetes. Is it important? How to measure it? J Diabetes Sci Technol 2008 Nov;2(6):1094-1100 [FREE Full text] [doi: 10.1177/193229680800200618] [Medline: 19885298]
6. Wang C, Lv L, Yang Y, Chen D, Liu G, Chen L, et al. Glucose fluctuations in subjects with normal glucose tolerance, impaired glucose regulation and newly diagnosed type 2 diabetes mellitus. Clin Endocrinol (Oxf) 2012 Jun 23;76(6):810-815. [doi: 10.1111/j.1365-2265.2011.04205.x] [Medline: 21854404]
7. Heise T, Mathieu C. Impact of the mode of protraction of basal insulin therapies on their pharmacokinetic and pharmacodynamic properties and resulting clinical outcomes. Diabetes Obes Metab 2017 Jan 26;19(1):3-12 [FREE Full text] [doi: 10.1111/dom.12782] [Medline: 27593206]
8. Home P, Bartley P, Russell-Jones D, Hanaire-Broutin H, Heeg JE, Abrams P, Study to Evaluate the Administration of Detemir Insulin Efficacy, Safety and Suitability (STEADINESS) Study Group. Insulin detemir offers improved glycemic control compared with NPH insulin in people with type 1 diabetes: a randomized clinical trial. Diabetes Care 2004 May;27(5):1081-1087. [doi: 10.2337/diacare.27.5.1081] [Medline: 15111525]
9. Becker RHA, Dahmen R, Bergmann K, Lehmann A, Jax T, Heise T. New insulin glargine 300 Units · mL-1 provides a more even activity profile and prolonged glycemic control at steady state compared with insulin glargine 100 Units · mL-1. Diabetes Care 2015 Apr;38(4):637-643. [doi: 10.2337/dc14-0006] [Medline: 25150159]
10. Heise T, Hermanski L, Nosek L, Feldman A, Rasmussen S, Haahr H. Insulin degludec: four times lower pharmacodynamic variability than insulin glargine under steady-state conditions in type 1 diabetes. Diabetes Obes Metab 2012 Sep 07;14(9):859-864 [FREE Full text] [doi: 10.1111/j.1463-1326.2012.01627.x] [Medline: 22594461]
11. Hamasaki H. Comparison of Glycemic Variability by Using Insulin Glargine and Insulin Degludec in Japanese Patients With Type 1 Diabetes, Monitored by Continuous Glucose Monitoring: A Preliminary Report. J Endocrinol Metab 2013;3(6):138-146. [doi: 10.4021/jem193w]

12. Bailey T, Pettus J, Roussel R, Schmider W, Maroccia M, Nassr N, et al. Morning administration of 0.4U/kg/day insulin glargine 300U/mL provides less fluctuating 24-hour pharmacodynamics and more even pharmacokinetic profiles compared with insulin degludec 100U/mL in type 1 diabetes. Diabetes Metab 2018 Feb;44(1):15-21 [FREE Full text] [doi: 10.1016/j.diabet.2017.10.001] [Medline: 29153485]
13. Heise T, Kaplan K, Haahr HL. Day-to-Day and Within-Day Variability in Glucose-Lowering Effect Between Insulin Degludec and Insulin Glargine (100 U/mL and 300 U/mL): A Comparison Across Studies. J Diabetes Sci Technol 2018 Mar 26;12(2):356-363 [FREE Full text] [doi: 10.1177/1932296817731422] [Medline: 28946756]
14. Wang J, Alexanian A, Ying R, Kizhakekuttu TJ, Dharmashankar K, Vasquez-Vivar J, et al. Acute Exposure to Low Glucose Rapidly Induces Endothelial Dysfunction and Mitochondrial Oxidative Stress. ATVB 2012 Mar;32(3):712-720. [doi: 10.1161/atvbaha.111.227389]
15. Kahal H, Halama A, Aburima A, Bhagwat AM, Butler AE, Graumann J, et al. Effect of induced hypoglycemia on inflammation and oxidative stress in type 2 diabetes and control subjects. Sci Rep 2020 Mar 16;10(1):4750 [FREE Full text] [doi: 10.1038/s41598-020-61531-z] [Medline: 32179763]
16. Ceriello A, Novials A, Ortega E, La Sala L, Pujadas G, Testa R, et al. Evidence that hyperglycemia after recovery from hypoglycemia worsens endothelial function and increases oxidative stress and inflammation in healthy control subjects and subjects with type 1 diabetes. Diabetes 2012 Nov 16;61(11):2993-2997 [FREE Full text] [doi: 10.2337/db12-0224] [Medline: 22891214]
17. Schnell O, Barnard K, Bergenstal R, Bosi E, Garg S, Guerci B, et al. Clinical Utility of SMBG: Recommendations on the Use and Reporting of SMBG in Clinical Research. Diabetes Care 2015 Sep;38(9):1627-1633. [doi: 10.2337/dc14-2919] [Medline: 26294772]
18. Umpierrez GE, P Kovatchev B. Glycemic Variability: How to Measure and Its Clinical Implication for Type 2 Diabetes. Am J Med Sci 2018 Dec;356(6):518-527 [FREE Full text] [doi: 10.1016/j.amjms.2018.09.010] [Medline: 30447705]
19. Ellman GL. Tissue sulfhydryl groups. Archives of Biochemistry and Biophysics 1959 May;82(1):70-77. [doi: 10.1016/0003-9861(59)90090-6]
20. Hu ML. Measurement of protein thiol groups and glutathione in plasma. In: Methods in Enzymology. Amsterdam: Elsevier; 1994.
21. Alberti A, Bolognini L, Macciantelli D, Caratelli M. The radical cation of N,N-diethyl-para-phenylendiamine: A possible indicator of oxidative stress in biological samples. Res Chem Intermed 2000 Jan;26(3):253-267. [doi: 10.1163/156856700x00769]
22. Verde V, Fogliano V, Ritieni A, Maiani G, Morisco F, Caporaso N. Use of N,N-dimethyl-p-phenylenediamine to evaluate the oxidative status of human plasma. Free Radic Res 2002 Aug 07;36(8):869-873. [doi: 10.1080/1071576021000005302] [Medline: 12420745]
23. Ceriello A, Esposito K, Piconi L, Ihnat M, Thorpe J, Testa R, et al. Glucose 'peak' and glucose 'spike': Impact on endothelial function and oxidative stress. Diabetes Research and Clinical Practice 2008 Nov;82(2):262-267. [doi: 10.1016/j.diabres.2008.07.015]
24. Heise T, Nørskov M, Nosek L, Kaplan K, Famulla S, Haahr HL. Insulin degludec: Lower day-to-day and within-day variability in pharmacodynamic response compared with insulin glargine 300 U/mL in type 1 diabetes. Diabetes Obes Metab 2017 Jul 23;19(7):1032-1039 [FREE Full text] [doi: 10.1111/dom.12938] [Medline: 28295934]
25. Gradel AKJ, Porsgaard T, Lykkesfeldt J, Seested T, Gram-Nielsen S, Kristensen NR, et al. Factors Affecting the Absorption of Subcutaneously Administered Insulin: Effect on Variability. J Diabetes Res 2018 Jul 04;2018:1205121-1205117 [FREE Full text] [doi: 10.1155/2018/1205121] [Medline: 30116732]
26. Kalra S, Baruah M, Sahay R, Kishor K. Pentads and hexads in diabetes care: Numbers as targets; Numbers as tools. Indian J Endocr Metab 2017;21(6):794. [doi: 10.4103/ijem.ijem_281_17]
27. Tibaldi J, Hadley-Brown M, Liebl A, Haldrup S, Sandberg V, Wolden ML, et al. A comparative effectiveness study of degludec and insulin glargine 300 U/mL in insulin-naïve patients with type 2 diabetes. Diabetes Obes Metab 2019 Apr;21(4):1001-1009 [FREE Full text] [doi: 10.1111/dom.13616] [Medline: 30552800]
28. Rosenstock J, Cheng A, Ritzel R, Bosnyak Z, Devisme C, Cali AM, et al. More Similarities Than Differences Testing Insulin Glargine 300 Units/mL Versus Insulin Degludec 100 Units/mL in Insulin-Naive Type 2 Diabetes: The Randomized Head-to-Head BRIGHT Trial. Diabetes Care 2018 Oct 13;41(10):2147-2154. [doi: 10.2337/dc18-0559] [Medline: 30104294]
29. Sullivan SD, Bailey TS, Roussel R, Zhou FL, Bosnyak Z, Preblick R, et al. Clinical outcomes in real-world patients with type 2 diabetes switching from first- to second-generation basal insulin analogues: Comparative effectiveness of insulin glargine 300 units/mL and insulin degludec in the DELIVER D+ cohort study. Diabetes Obes Metab 2018 Sep 25;20(9):2148-2158 [FREE Full text] [doi: 10.1111/dom.13345] [Medline: 29938887]
30. Giacco F, Brownlee M. Oxidative Stress and Diabetic Complications. Circ Res 2010 Oct 29;107(9):1058-1070. [doi: 10.1161/circresaha.110.223545]
31. Ge X, Yu Q, Qi W, Shi X, Zhai Q. Chronic insulin treatment causes insulin resistance in 3T3-L1 adipocytes through oxidative stress. Free Radic Res 2008 Jun 07;42(6):582-591. [doi: 10.1080/10715760802158448] [Medline: 18569016]
32. Czupryniak L, Barkai L, Bolgarska S, Bronisz A, Broz J, Cypryk K, et al. Self-monitoring of blood glucose in diabetes: from evidence to clinical reality in Central and Eastern Europe--recommendations from the international Central-Eastern

European expert group. Diabetes Technol Ther 2014 Jul;16(7):460-475 [FREE Full text] [doi: 10.1089/dia.2013.0302] [Medline: 24716890]
33. Jung HS. Clinical Implications of Glucose Variability: Chronic Complications of Diabetes. Endocrinol Metab 2015;30(2):167. [doi: 10.3803/enm.2015.30.2.167]

Abbreviations

CGMS: continuous glucose monitoring system
CV: coefficient of variation
d-ROM: determination of reactive oxygen metabolites
HbA_{1c}: glycated hemoglobin
IDeg-100: insulin degludec 100 IU/mL
IGlar-300: insulin glargine 300 IU/mL
MAGE: mean amplitude of glucose excursions
NPH: Neutral Protamine Hagedorn
SMBG: self-monitoring of blood glucose
T1DM: type 1 diabetes mellitus
T2DM: type 2 diabetes mellitus

Edited by T Leung; submitted 12.12.21; peer-reviewed by M Bakula, D Lewis; comments to author 11.05.22; revised version received 30.05.22; accepted 01.06.22; published 08.07.22.

Please cite as:
Vrebalov Cindro P, Krnić M, Modun D, Vuković J, Tičinović Kurir T, Kardum G, Rušić D, Šešelja Perišin A, Bukić J
Comparison of the Impact of Insulin Degludec U100 and Insulin Glargine U300 on Glycemic Variability and Oxidative Stress in Insulin-Naive Patients With Type 2 Diabetes Mellitus: Pilot Study for a Randomized Trial
JMIR Form Res 2022;6(7):e35655
URL: https://formative.jmir.org/2022/7/e35655
doi:10.2196/35655
PMID:35802405

Original Paper

The Association Between COVID-19 Information Sources and Stigma Against Health Care Workers Among College Students: Cross-sectional, Observational Study

Miharu Nakanishi[1,2], PhD; Mai Sakai[1], PhD; Gen Takagi[3], PhD; Keita Toshi[1], MSc; Koubun Wakashima[4], PhD; Hatsumi Yoshii[1], PhD

[1]Department of Psychiatric Nursing, Tohoku University Graduate School of Medicine, Sendai-shi, Japan

[2]Research Center for Social Science & Medicine, Tokyo Metropolitan Institute of Medical Science, Setagaya-ku, Japan

[3]Department of Welfare Psychology, Faculty of General Welfare, Tohoku Fukushi University, Sendai-shi, Japan

[4]Department of Clinical Psychology & Family Psychology, Graduate School of Education, Tohoku University, Sendai-shi, Japan

Corresponding Author:
Miharu Nakanishi, PhD
Department of Psychiatric Nursing
Tohoku University Graduate School of Medicine
2-1 Seiryo-machi
Aoba-ku
Sendai-shi, 980-8575
Japan
Phone: 81 227178179
Email: nakanishi-mh@tohoku.ac.jp

Abstract

Background: The COVID-19 pandemic has triggered stigmatic attitudes against health care workers. Some forms of social media may play a role in disseminating stigmatizing messages.

Objective: We aimed to investigate the association between COVID-19 information sources and stigma against health care workers among college students during the pandemic.

Methods: A cross-sectional, observational study was conducted using a web-based platform in the Tohoku region of Japan. College students aged ≥20 years were asked to complete the questionnaire between August 18 and October 31, 2020. Stigma against health care workers was evaluated using a modified Japanese version of the Social Distance Scale. Participants were also asked to rate their perceived vulnerability to infection using the Japanese version of the Perceived Vulnerability to Disease scale.

Results: A total of 281 students from 8 colleges completed the web-based survey. There were 139 (49.5%) participants who used Twitter, 187 (66.5%) who used news websites, and 46 (16.4%) who used the websites of public health agencies as COVID-19 information sources. After adjusting for age, sex, department, and Perceived Vulnerability to Disease scores, the level of stigma did not differ between students who used Twitter and those who did not. Students who used the websites of public health agencies showed a significantly less stigmatic attitude than those who did not.

Conclusions: Fact-checking and directing visitors to credible information sources from public health agencies may have prevented the formation of stigmatic attitudes toward health care workers. An effective strategy to enable easy access to information provided by public agencies should be integrated into widespread web-based platforms.

(JMIR Form Res 2022;6(7):e35806) doi:10.2196/35806

KEYWORDS

health personnel; social media; social stigma; young adult; Twitter; public health; COVID-19; health care workers; students; information accuracy; information credibility; dissemination; information source; misinformation; information spread; infodemiology

Introduction

The COVID-19 pandemic is a serious threat to public health. Moreover, it is characterized by widespread fear, worry, and uncertainty, as many COVID-19 infections are contracted through presymptomatic and asymptomatic transmissions [1]. This health emergency has triggered discriminatory behavior and stigma against health care workers, despite their vital role in caring for people with COVID-19 [2,3]. Stigma is defined as an undesirable characteristic that results in discrimination against an individual [4]. Several incidents of stigmatization of health care workers have been reported; these include avoidance by family members or their community, being denied access to public transport, and even being subjected to physical assault [3,5]. The psychological challenges entailing stigmatization may amplify the negative consequences, such as emotional burnout [6], posttraumatic stress disorder [7], and turnover [8] of working with COVID-19 patients as frontline care providers. A reduction of stigma against health care workers is thus warranted during the COVID-19 pandemic.

The use of unreliable forms of social media as information sources for COVID-19 can lead to the spread of misinformation and increased stigma against health care workers. This is especially true because various COVID-19–related rumors, stigma, and conspiracy theories have been circulating on the internet [9]. Most people rely on the internet for COVID-19 information [10,11]. Particularly, Twitter conversations pertaining to COVID-19 are characterized by the dissemination of stigmatizing messages [12,13]. Although the media, including newspapers, television, and websites, are important sources that can be used to promote health education and literacy, mass media and even health agencies have contributed to the spread of health misinformation that could circulate stigma on the internet [14]. In addition to the media, Twitter allows users to post short messages (tweets), “retweet” messages (reposts), send replies, and “like” messages by other users. Therefore, Twitter users are more likely to be exposed to misinformation and stigmatizing messages, which in turn may exacerbate fear and anxiety and result in the further circulation of such messages [9]. Adolescents and young adults have been using social networking services more frequently since the COVID-19 pandemic, as pandemic-related restrictions have substantially changed their social lives due to school closures [15]. A previous study indicated that the level of anxiety about COVID-19 differed according to the type of preferred news source [16]. However, no studies have examined the association between the types of COVID-19 information sources and stigma against health care workers. An understanding of this association will provide a basis for developing strategies aimed at reducing stigma.

This study, thus, investigated the association between COVID-19 information sources and stigma against health care workers among college students during the pandemic. We hypothesized that college students who used Twitter as a COVID-19 information source would show a more stigmatic attitude toward health care workers than those who do not use Twitter.

Methods

Study Design

This cross-sectional, observational study was conducted using Google Forms, which is a web-based tool that allows data collection through personalized surveys. An anonymous questionnaire was uploaded and shared through an invitation email to potential participants.

Setting

On August 18, 2020, the survey link was shared with teachers from 8 colleges in the Tohoku region of Japan. The link provided a detailed explanation of the study purpose instructions. Subsequently, the teachers emailed the invitation link with the explanatory documents to the students. Participants aged ≥20 years were requested to complete the questionnaire on October 31, 2020.

The first page of the website contained details on the voluntary nature of the participation and protection of personal information. After reading the introduction, students indicated their consent to participate by clicking on the link to start the survey. The “Limit to 1 response” function of Google Forms was enabled to prevent respondents from completing the form more than once.

Participants

We used purposive and convenience sampling to select colleges and departments with (1) nursing students who underwent on-the-job training in hospitals and (2) students from other departments. The other departments were selected to ensure gender ratios similar to those of students in the nursing departments.

Nursing and other students were recruited to compare the level of stigma against health care workers. We assumed that nursing students would show less stigmatic attitudes than other students, because they are expected to become professional nurses and, therefore, have more psychological proximity with other health care workers.

Measurements

The questionnaire included questions regarding stigmatic attitudes against health care workers, COVID-19 information sources, perceived vulnerability to infection, department, age, sex, and contact with COVID-19 patients. The questions and response options of stigma against health care workers were developed for this study (Table S1 in Multimedia Appendix 1).

Stigma against health care workers was evaluated using a modified version of the Japanese version of the Social Distance Scale (SDSJ) [17]. The original Social Distance Scale was developed based on the Keyed Favorable Response and Ego-Involvement Ratings of Scale [18], which is used to assess the level of stigmatic attitude toward patients with schizophrenia. It contains 8 stigma-related items, which are rated on a 4-point Likert scale, and the scale has good reliability and validity [17]. For this study, we replaced “patients with schizophrenia” in each item with “health care workers and their families who are performing infectious disease management with an unestablished

treatment regimen." The total score on the 8 items ranged from 0 to 24; higher scores indicated more stigmatic attitudes.

The 8 types of COVID-19 information sources were listed in the questionnaire. Participants were asked to check any source, if applicable. The list was developed by a research panel to contain a range of social media platforms, including newspapers, television news streams, television tabloid talk shows, news websites, Twitter, the websites of public health agencies (eg, Ministry of Health, Labour and Welfare), Instagram, and Facebook.

Perceived vulnerability to infection was evaluated using the Perceived Vulnerability to Disease (PVD) scale [19]. The PVD contains 15 items, which are rated on a 7-point Likert scale from 1 to 7. It comprises 2 subscales: Perceived Infectability and Germ Aversion. Perceived Infectability refers to beliefs about immunological functioning and personal susceptibility to infectious diseases. Germ Aversion refers to aversive affective responses to situations that connote a relatively high likelihood of pathogen transmission. The Japanese version of the PVD is reported to have good reliability and validity [20]. In this study, the Cronbach α coefficient was .78 (95% CI .74-.82) for Perceived Infectability and .65 (95% CI .58-.71) for Germ Aversion. Although Perceived Infectability refers to one's susceptibility to infection, Germ Aversion covers behaviors exerting emotional discomfort in a high-pathogen context, which in turn deters from the source of infection. Therefore, we assumed that stigma against health care workers would show a moderate positive correlation with Germ Aversion but not with Perceived Infectability.

Participants also answered items pertaining to their age, sex, department, and contact with COVID-19 patients. The presence of contact was assessed if (1) the respondent or their family members or friends had been infected with COVID-19; and if (2) the respondent or their family members or friends had close contact with infected persons. Based on the contact hypothesis [21,22], we assumed that individuals who directly interacted with COVID-19 patients would be less likely to hold stigmatic attitudes against health care workers. However, the number of individuals with direct contact was small (n=18). Therefore, we used the presence of contact for sensitivity analysis instead of including it in bivariate and multivariate analyses.

Study Size

The required sample size was calculated using G*Power (version 3.1.9.7; Faul et al [23,24]). Based on a recent report on the use of Twitter among Japanese college students [25], we assumed the prevalence of Twitter as a COVID-19 information source to be 50% in this study. Assuming an α level of 5%, 95% power, and medium effect size (Cohen *d*=0.5) for a 2-tailed test, the minimum sample size was determined to be 210.

Statistical Analysis

The validity and reliability of the modified SDSJ were examined. To test concurrent validity, the mean difference between nursing and other students was examined. The normality of distribution was assumed for total SDSJ score (Table S2 in Multimedia Appendix 1). The equality of variances between the 2 groups was also assumed (Table S3 in Multimedia Appendix 1). Therefore, 2-sided testing using 2-tailed students' *t* test was used. To test convergent validity, Pearson correlation coefficient was calculated between the total SDSJ score and the subscales scores of the PVD. To test internal reliability, the Cronbach α coefficient and 95% CI [26] were calculated for the total SDSJ score.

Within the types of COVID-19 information sources, overlaps between Twitter and other major web-based platforms (news websites and the websites of public health agencies) were examined. To investigate the association between Twitter use and stigma against health care workers, the mean difference in SDSJ scores was examined by performing a 2-tailed *t* tests between participants who used Twitter as a COVID-19 information source and those who did not. Furthermore, a multiple Ordinary Least Square (OLS) regression analysis was performed using the SDSJ scores as the dependent variable and the types of information sources as independent variables. Participant characteristics, including age, sex, department, and PVD scores, were included as covariates. Diagnostic tests were conducted to test the assumptions of OLS regression, including the normality of the residuals, homoskedasticity, the absence of outliers, and low multicollinearity (Table S4 in Multimedia Appendix 1). Since some potential outliers were found, a sensitivity analysis of the multivariate model was performed by excluding them. Another sensitivity analysis was also conducted by excluding individuals with social contacts.

All analyses were conducted using Stata statistical software (version 17.0; StataCorp). The significance level was set at low (α=.1), medium (α=.05), and high (α=.01) [27]. As our primary endpoint was to assess the association between use of Twitter and SDSJ under multiple regression analysis adjusting for covariates, we did not apply *P* value adjustments for multiple hypothesis testing [28].

Ethics Approval

The study protocol was approved by the Ethics Board of Tohoku University (2021-1-733) and was conducted in accordance with the Helsinki Declaration of 1975 (as revised in 2013).

Results

Participant Characteristics

A total of 281 participants completed the survey and were included in the final sample. The majority (86.7%, n=238) were women and the most frequent (50.9%, n=143) age was 21 years. There were 18 participants who reported coming into contact with COVID-19 (Table 1).

Table 1. Participant characteristics.

Variable, category	Participant (N=281), n (%)
Sex	
Female	238 (86.7)
Age (year)	
20	86 (30.6)
21	143 (50.9)
≥22	52 (18.5)
Department	
Nursing	104 (37)
Rehabilitation	70 (24.9)
Psychology	63 (22.4)
Other	44 (15.7)
Contact with COVID-19	
Any of the experiences below	18 (6.4)
Family members or friends had close contact with an infected person	13 (4.6)
I had close contact with an infected person	1 (0.4)
Family members or friends had been infected with COVID-19	5 (1.8)
I had been infected with COVID-19	0 (0)

Stigma Against Health Care Workers

The mean total SDSJ score in overall sample was 7.9 (SD 4.7). Nursing students had a significantly lower mean total score (mean 6.0, SD 4.5) than other students (mean 8.9, SD 4.5; $t_{218.48}$=5.18, P<.001). The mean total SDSJ score did not differ between participants with social contact (mean 8.6, SD 1.0) and those without (mean 7.8, SD 0.3; $t_{20.16}$=0.73, P=.47).

Pearson correlation coefficients were –0.03 (P=.62) between the SDSJ and PVD Perceived Vulnerability scores and 0.33 (P<.001) between the SDSJ and PVD Germ Aversion scores. The Cronbach α coefficient of the SDSJ score was high (α=.83; 95% CI .80-.86). In summary, the modified SDSJ demonstrated satisfactory concurrent, convergent, and internal validity.

COVID-19 Information Sources

Half (49.5%, 139/281) of the participants used Twitter as a COVID-19 information source (Table 2). Since less than 10% of participants used newspapers (8.5%, n=24), Instagram (3.2%, n=9), or Facebook (0.4%, n=1), we excluded these sources from the following multivariate analysis.

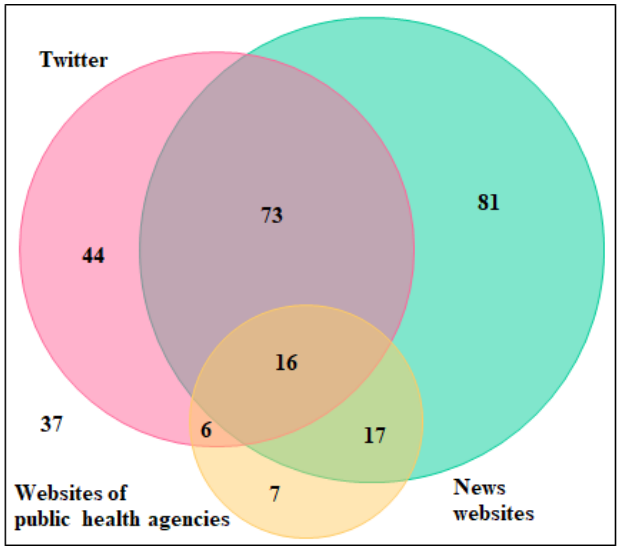
In total, 89 participants reported using both Twitter and news websites, of which 16 participants also used the websites of public health agencies. Further, 44 participants used Twitter, but not news websites or the websites of public health agencies (Figure 1).

The total SDSJ score did not differ between Twitter users (mean 8.1, SD 4.4) and those who did not use Twitter (mean 7.6, SD 5.0; $t_{276.63}$=0.97, P=.33).

Table 2. Prevalence of COVID-19 information sources.

Type, category	Participant (N=281), n (%)
Newspaper	24 (8.5)
Television	
News stream	218 (77.6)
Tabloid talk show	71 (25.3)
Web-based source	
News website	187 (66.5)
Twitter	139 (49.5)
Website of public health agencies (eg, Ministry of Health, Labour and Welfare)	46 (16.4)
Instagram	9 (3.2)
Facebook	1 (0.4)
Other social networking services	13 (4.6)

Figure 1. Venn diagram of overlaps between Twitter and other major web-based platforms (news websites and the websites of public health agencies).



Association Between Stigma and Information Sources

A multiple regression analysis showed that the total SDSJ scores were lower among participants using the websites of public health agencies (P=.008), nursing students (P<.001), and those with lower Germ Aversion scores (P<.001; Table 3). The use of Twitter was not associated with SDSJ scores (P=.58).

The results of OLS diagnostic tests showed that the following assumptions were met: the normality of the residuals, homoskedasticity, and low multicollinearity (Table S4 in Multimedia Appendix 1). However, there were 2 potential outliers (Table S4 in Multimedia Appendix 1). A sensitivity analysis that excluded the 2 individuals did not change the results (Table S5 in Multimedia Appendix 1).

Another sensitivity analysis that excluded individuals (n=18) who came into close contact with a patient with COVID-19 did not alter the results (Table S5 in Multimedia Appendix 1).

Table 3. Multiple linear regression analysis of stigma against health care workers[a,b].

Variable, category	Coefficient (95% CI)	*P* value
COVID-19 information source		
Television news stream	0.86 (–0.40 to 2.12)	.18
Television tabloid show	0.51 (–0.70 to 1.72)	.41
News website	0.60 (–0.45 to 1.66)	.26
Twitter	0.29 (–0.74 to 1.31)	.58
Websites of public health agencies	–1.84 (–3.20 to –0.49)	.008
Sex		
Male	1.07 (–0.38 to 2.52)	.15
Age (year)		
20	0.33 (–1.21 to 1.88)	.67
21	–0.25 (–1.64 to 1.14)	.72
≥22	reference	reference
Department		
Nursing	–3.04 (–4.11 to –1.98)	<.001
Perceived vulnerability to infection		
Perceived Infectability	–0.13 (–0.63 to 0.37)	.61
Germ Aversion	1.85 (1.31-2.39)	<.001

[a]Stigma against health care workers was evaluated using the modified Japanese language version of the Social Distance Scale; the total score ranges from 0 to 24.

[b]Perceived vulnerability to infection was evaluated using the Japanese version of the Perceived Vulnerability to Disease scale; total scores range from 1 to 7 for both Perceived Infectability and Germ Aversion.

Discussion

With increasing concerns about the stigma against health care workers during the COVID-19 pandemic, strategies to reduce stigma are the need of the hour, considering the prevalence of Twitter and other information sources for information related to COVID-19. Thus, this study examined the association between the types of information source and stigma against health care workers among college students.

Principal Findings

Contrary to our hypothesis, Twitter use was not associated with the stigma against health care workers. The survey was conducted between August and October 2020 when the daily number of new COVID-19 cases in Japan ranged from 219 to 1178. This period is a few months after the onset of the outbreak, and our participants may have had a lower level of fear than that from February to April 2020 [9,12,13]. In the acute phase of the psychological response to a crisis, such as the Great East Japan Earthquake on March 11, 2011, Twitter messages diffused rumors and misinformation [29]. However, the level of anxiety expressed in Twitter messages appears to return to normal over time [30]. This shift to Twitter for content has also been observed for COVID-19 [31]. In addition, several students aged ≥20 years in the Tohoku region might have experienced the March 11 earthquake when they were children, resulting in their learned mindset to treat Twitter messages with caution.

College students who used the websites of public health agencies as information sources reported significantly less stigmatic attitudes than those who did not (*P*=.008) in the multiple regression analysis. Fact-checking and directing users to credible information sources from the websites of public health agencies can prevent the further spread of misinformation [32]. News websites also disseminate official announcements from public health agencies. In this study, substantial overlaps were observed between the users of the websites of public health agencies, news websites, and Twitter. This finding is consistent with a previous study [10]. Furthermore, public health agencies also have Twitter accounts, and each message is limited to 140 Japanese characters. Their messages usually include a URL to official website pages that contain longer texts. The accuracy of the information held by Twitter users may vary between those who only read the message and those who click on the link to access the website. Adequate communication strategies should be embedded in reliable information from trusted sources [31,33]. Considering that Twitter was a popular COVID-19 information source, access to information curated by public agencies may help reduce exposure to misinformation and stigma against health care workers. In addition to accurate information, "hero" messaging [34] should be reserved for health care workers in public policy. Health care workers were deemed essential frontline heroes during the COVID-19 crisis. Such perceptions can mitigate stigmatic attitudes toward health care workers. Furthermore, positive video messages embedded in tweets may enable young people to engage in parasocial interactions with health care workers, which in turn may help

change their negative beliefs about health care workers based on the parasocial contact hypothesis [22,35].

Stigma against health care workers, as measured by the modified SDSJ, was significantly lower among nursing students (P<.001) and students with lower germ aversion (P<.001). Unlike Germ Aversion, we found very little statistical evidence that Perceived Infectability was associated with stigma against health care workers using the total SDSJ score. These associations were consistent with our assumption that the modified scale is valid.

Strengths and Limitations

The strength of this study lies in the examination of stigma against health care workers in the context of COVID-19 and social media, thus addressing a gap in the literature. A limitation was the lack of information accuracy and limited data on the intensity of social media use among participants. In addition, the cross-sectional design precludes causal inferences between stigma and the types of information sources. The use of websites of public health agencies might indicate the high health literacy and low stigmatic attitudes of students.

Conclusions

A few months after the onset of the COVID-19 pandemic, nearly half of the college student population used Twitter as an information source. Our findings showed that the level of stigma against health care workers did not differ according to Twitter use. Students who used the websites of public health agencies reported less stigmatic attitudes than those who did not. These results imply that directing people to credible COVID-19 information sources from public agencies may prevent the formation of stigmatic attitudes against health care workers. An effective strategy for the induction of access to credible information sources should be explored for integration into Twitter and other widespread web-based platforms.

Acknowledgments

This study was supported by the Department of Psychiatric Nursing, Tohoku University Graduate School of Medicine. The funder was not involved in the design or study course; data collection, management, analysis, or interpretation; or monitoring of the manuscript for presentation, review, or approval.

Authors' Contributions

MS, KW, and HY collected the data. MN, MS, GT, and KT analyzed and interpreted the data and drafted the manuscript. KW and HY were involved in the study design and setup, the supervision of data analysis, and finalizing the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary tables on (1) questions and response options used in the online survey; (2) test of the normality of scores; (3) Ordinary Least Square regression diagnostic tests; (4) sensitivity analysis: multiple linear regression analysis of stigma against health care workers, excluding possible outliers (N=279); and (5) sensitivity analysis: multiple linear regression analysis of stigma against health care workers, excluding individuals who came in close contact with a patient with COVID-19 (N=263).
[DOCX File , 28 KB - formative_v6i7e35806_app1.docx]

References

1. Moghadas SM, Fitzpatrick MC, Sah P, Pandey A, Shoukat A, Singer BH, et al. The implications of silent transmission for the control of COVID-19 outbreaks. Proc Natl Acad Sci U S A 2020 Jul 28;117(30):17513-17515 [FREE Full text] [doi: 10.1073/pnas.2008373117] [Medline: 32632012]
2. Bagcchi S. Stigma during the COVID-19 pandemic. Lancet Infect Dis 2020 Jul;20(7):782 [FREE Full text] [doi: 10.1016/S1473-3099(20)30498-9] [Medline: 32592670]
3. Larkin H. Navigating attacks against health care workers in the COVID-19 era. JAMA 2021 May 11;325(18):1822-1824. [doi: 10.1001/jama.2021.2701] [Medline: 33881489]
4. Link BG, Phelan JC. Conceptualizing stigma. Annu Rev Sociol 2001 Aug;27(1):363-385. [doi: 10.1146/annurev.soc.27.1.363]
5. Dye TD, Alcantara L, Siddiqi S, Barbosu M, Sharma S, Panko T, et al. Risk of COVID-19-related bullying, harassment and stigma among healthcare workers: an analytical cross-sectional global study. BMJ Open 2020 Dec 30;10(12):e046620 [FREE Full text] [doi: 10.1136/bmjopen-2020-046620] [Medline: 33380488]
6. Ramaci T, Barattucci M, Ledda C, Rapisarda V. Social stigma during COVID-19 and its impact on HCWs outcomes. Sustainability 2020 May 08;12(9):3834. [doi: 10.3390/su12093834]
7. Carmassi C, Foghi C, Dell'Oste V, Cordone A, Bertelloni CA, Bui E, et al. PTSD symptoms in healthcare workers facing the three coronavirus outbreaks: what can we expect after the COVID-19 pandemic. Psychiatry Res 2020 Oct;292:113312 [FREE Full text] [doi: 10.1016/j.psychres.2020.113312] [Medline: 32717711]

8. Al-Mansour K. Stress and turnover intention among healthcare workers in Saudi Arabia during the time of COVID-19: can social support play a role? PLoS One 2021 Oct 7;16(10):e0258101 [FREE Full text] [doi: 10.1371/journal.pone.0258101] [Medline: 34618851]
9. Islam MS, Sarkar T, Khan SH, Mostofa Kamal AH, Hasan SMM, Kabir A, et al. COVID-19-related infodemic and its impact on public health: a global social media analysis. Am J Trop Med Hyg 2020 Oct;103(4):1621-1629 [FREE Full text] [doi: 10.4269/ajtmh.20-0812] [Medline: 32783794]
10. Wang PW, Lu WH, Ko NY, Chen YL, Li DJ, Chang YP, et al. COVID-19-related information sources and the relationship with confidence in people coping with COVID-19: Facebook survey study in Taiwan. J Med Internet Res 2020 Jun 05;22(6):e20021 [FREE Full text] [doi: 10.2196/20021] [Medline: 32490839]
11. Tran BX, Dang AK, Thai PK, Le HT, Le XTT, Do TTT, et al. Coverage of health information by different sources in communities: implication for COVID-19 epidemic response. Int J Environ Res Public Health 2020 May 20;17(10):3577 [FREE Full text] [doi: 10.3390/ijerph17103577] [Medline: 32443712]
12. Xue J, Chen J, Hu R, Chen C, Zheng C, Su Y, et al. Twitter discussions and emotions about the COVID-19 pandemic: machine learning approach. J Med Internet Res 2020 Nov 25;22(11):e20550 [FREE Full text] [doi: 10.2196/20550] [Medline: 33119535]
13. Kumble S, Diddi P. Twitter conversations about COVID-19 during pre-pandemic period: stigma and information format cues. Stigma Health 2021 Aug;6(3):251-262. [doi: 10.1037/sah0000324]
14. Rovetta A, Castaldo L. A new infodemiological approach through Google Trends: longitudinal analysis of COVID-19 scientific and infodemic names in Italy. BMC Med Res Methodol 2022 Jan 30;22(1):33 [FREE Full text] [doi: 10.1186/s12874-022-01523-x] [Medline: 35094682]
15. Preusting LC, Raadsen MP, Abourashed A, Voeten HACM, Wagener MN, de Wit E, et al. COVID-19 related stigma and health-protective behaviours among adolescents in the Netherlands: an explorative study. PLoS One 2021 Jun 22;16(6):e0253342 [FREE Full text] [doi: 10.1371/journal.pone.0253342] [Medline: 34157046]
16. Meltzer GY, Chang VW, Lieff SA, Grivel MM, Yang LH, Des Jarlais DC. Behavioral correlates of COVID-19 worry: stigma, knowledge, and news source. Int J Environ Res Public Health 2021 Oct 30;18(21):11436 [FREE Full text] [doi: 10.3390/ijerph182111436] [Medline: 34769952]
17. Makita K. Development and reliability of the Japanese-language version of Social Distance Scale (SDSJ). Article in Japanese. Jpn Bull Soc Psychiat 2006;14(3):231-241.
18. Whatley CD. Social attitudes toward discharge mental patients. Soc Probl 1959 Apr;6(4):313-320. [doi: 10.2307/799364]
19. Duncan LA, Schaller M, Park JH. Perceived vulnerability to disease: development and validation of a 15-item self-report instrument. Pers Individ Dif 2009 Oct;47(6):541-546. [doi: 10.1016/j.paid.2009.05.001]
20. Fukukawa Y, Oda R, Usami H, Kawahito J. [Development of a Japanese version of the Perceived Vulnerability to Disease Scale]. Article in Japanese. Shinrigaku Kenkyu 2014 Jun;85(2):188-195. [doi: 10.4992/jjpsy.85.13206] [Medline: 25016839]
21. Allport GW. The Nature of Prejudice. Cambridge, MA: Perseus Books; 1954.
22. Cho H, Li W, Cannon J, Lopez R, Song C. Testing three explanations for stigmatization of people of Asian descent during COVID-19: maladaptive coping, biased media use, or racial prejudice? Ethn Health 2021 Jan 15;26(1):94-109. [doi: 10.1080/13557858.2020.1830035] [Medline: 33059486]
23. Faul F, Erdfelder E, Buchner A, Lang A. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
24. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
25. Sakurai R, Nemoto Y, Mastunaga H, Fujiwara Y. Who is mentally healthy? mental health profiles of Japanese social networking service users with a focus on LINE, Facebook, Twitter, and Instagram. PLoS One 2021 Mar 3;16(3):e0246090 [FREE Full text] [doi: 10.1371/journal.pone.0246090] [Medline: 33657132]
26. Feldt LS, Woodruff DJ, Salih FA. Statistical inference for coefficient alpha. Appl Psychol Meas 1987 Mar 01;11(1):93-103. [doi: 10.1177/014662168701100107]
27. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol 2016 Apr 21;31(4):337-350 [FREE Full text] [doi: 10.1007/s10654-016-0149-3] [Medline: 27209009]
28. Greenland S, Hofman A. Multiple comparisons controversies are about context and costs, not frequentism versus Bayesianism. Eur J Epidemiol 2019 Sep 14;34(9):801-808 [FREE Full text] [doi: 10.1007/s10654-019-00552-z] [Medline: 31522327]
29. Takayasu M, Sato K, Sano Y, Yamada K, Miura W, Takayasu H. Rumor diffusion and convergence during the 3.11 earthquake: a twitter case study. PLoS One 2015 Apr 1;10(4):e0121443 [FREE Full text] [doi: 10.1371/journal.pone.0121443] [Medline: 25831122]
30. Doan S, Vo BKH, Collier N. An analysis of Twitter messages in the 2011 Tohoku Earthquake. In: Lecture notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 91. Berlin, Heidelberg: Springer; 2012 Presented at: eHealth 2011: Electronic Healthcare; November 21-23, 2011; Málaga, Spain p. 58-66. [doi: 10.1007/978-3-642-29262-0_8]

31. Gallotti R, Valle F, Castaldo N, Sacco P, de Domenico M. Assessing the risks of 'infodemics' in response to COVID-19 epidemics. Nat Hum Behav 2020 Dec 29;4(12):1285-1293. [doi: 10.1038/s41562-020-00994-6] [Medline: 33122812]
32. Gruzd A, Mai P. Going viral: how a single tweet spawned a COVID-19 conspiracy theory on Twitter. Big Data Soc 2020 Jul 20. [doi: 10.1177/2053951720938405]
33. Goodwin R, Wiwattanapantuwong J, Tuicomepee A, Suttiwan P, Watakakosol R. Anxiety and public responses to covid-19: early data from Thailand. J Psychiatr Res 2020 Oct;129:118-121 [FREE Full text] [doi: 10.1016/j.jpsychires.2020.06.026] [Medline: 32912591]
34. Mejia C, Pittman R, Beltramo JM, Horan K, Grinley A, Shoss MK. Int J Hosp Manag 2021 Feb;93:102772. [doi: 10.1016/j.ijhm.2020.102772]
35. Schiappa E, Gregg PB, Hewes DE. The parasocial contact hypothesis. Commun Monogr 2005 Mar;72(1):92-115. [doi: 10.1080/0363775052000342544]

Abbreviations

OLS: Ordinary Least Square
PVD: Perceived Vulnerability to Disease
SDSJ: Japanese version of the Social Distance Scale

Edited by A Mavragani; submitted 19.12.21; peer-reviewed by A Rovetta, C Zielinski; comments to author 08.05.22; revised version received 25.05.22; accepted 08.06.22; published 07.07.22.

Please cite as:
Nakanishi M, Sakai M, Takagi G, Toshi K, Wakashima K, Yoshii H
The Association Between COVID-19 Information Sources and Stigma Against Health Care Workers Among College Students: Cross-sectional, Observational Study
JMIR Form Res 2022;6(7):e35806
URL: https://formative.jmir.org/2022/7/e35806
doi:10.2196/35806
PMID:35797105

Original Paper

Physicians' Perceptions of Telemedicine Use During the COVID-19 Pandemic in Riyadh, Saudi Arabia: Cross-sectional Study

Sarah Idriss[1*], BA; Abdullah Aldhuhayyan[2*], BA; Ahmad Abdullah Alanazi[3*], BA; Walaa Alasaadi[4*], BA; Reem Alharbi[5*], BA; Ghadah Alshahwan[5*], BA; Mohammad Baitalmal[6*], PhD; Wadi Alonazi[7], PhD

[1]Ministry of Health, Riyadh, Saudi Arabia

[2]Emergency Medical Services, Saudi Red Crescent Authority, Riyadh, Saudi Arabia

[3]Respiratory Therapy Department, King Abdulaziz Medical City, Riyadh, Saudi Arabia

[4]Health Administration Department, King Saud University, College of Business, Riyadh, Saudi Arabia

[5]Home Healthcare Department, King Fahad Medical City, Riyadh, Saudi Arabia

[6]Quantitative Analysis Department, King Saud University, College of Business Administration, Riyadh, Saudi Arabia

[7]Health Administration Department, King Saud University, College of Business Administration, Riyadh, Saudi Arabia

[*]these authors contributed equally

Corresponding Author:
Sarah Idriss, BA
Ministry of Health
Prince Abdulrahman Bin Abdulaziz St
Riyadh
Saudi Arabia
Phone: 966 555070628
Email: s.idriss@hotmail.com

Abstract

Background: The term "telemedicine" refers to the use of communication technology to deliver health care remotely. The COVID-19 pandemic had substantial impacts on health care delivery from 2020 onward, and it was necessary to adapt high-quality care in a manner that limited the potential for viral exposure of both patients and health care workers. Physicians employed video, phone, and electronic written (e-consultation) visits, all of which provided quality of care comparable to that of face-to-face visits while reducing barriers of adopting telemedicine.

Objective: This study sought to assess physicians' perspectives and attitudes regarding the use of telemedicine in Riyadh hospitals during the COVID-19 pandemic. The main objects of assessment were as follows: (1) physicians' experience using telemedicine, (2) physicians' willingness to use telemedicine in the future, (3) physicians' perceptions of patient experiences, and (4) the influence of telemedicine on burnout.

Methods: This study employed SurveyMonkey to develop and distribute an anonymous 28-question cross-sectional survey among physicians across all specialty disciplines in Riyadh hospitals. A chi-square test was used to determine the level of association between variables, with significance set to $P<.05$.

Results: The survey was distributed among 500 physicians who experienced telemedicine between October 2021 and December 2021. A total of 362 doctors were included, of whom 28.7% (n=104) were consultants, 30.4% (n=110) were specialists, and 40.9% (n=148) were residents. Male doctors formed the majority 56.1% (n=203), and female doctors accounted for 43.9% (n=159). Overall, 34% (n=228) agreed or somewhat agreed that the "quality of care during telemedicine is comparable with that of face-to-face visits." Approximately 70% (n=254) believed that telemedicine consultation is cost-effective. Regarding burnout, 4.1% (n=15), 7.5% (n=27), and 27.3% (n=99) of the doctors reported feeling burnout every day, a few times a week, and a few times per month, respectively.

Conclusions: The physicians had generally favorable attitudes toward telemedicine, believing that its quality of health care delivery is comparable to that of in-person care. However, further research is necessary to determine how physicians' attitudes toward telemedicine have changed since the pandemic and how this virtual technology can be used to improve physicians' professional and personal well-being.

(JMIR Form Res 2022;6(7):e36029) doi:10.2196/36029

KEYWORDS

burnout; COVID-19; patient experience; perception; physicians; telemedicine; virtual

Introduction

Background

"Telemedicine" refers to the use of IT in the provision of health care via electronic devices [1]. COVID-19, a highly infectious virus causing respiratory illness, has caused an ongoing global pandemic. Saudi Arabia experienced more than half a million infections and more than 7000 fatalities by June 2021 [1]. This outbreak has led to innovative practices such as Telemedicine among health providers. Thus, physicians have often opted to use telemedicine as an alternative to provide fast and safe care away from outpatient clinics, which require physical contact [2]. This research investigates physicians' perceptions of telemedicine during the COVID-19 pandemic in Riyadh, Saudi Arabia.

Although various forms of telemedicine such as phone calls and electronic messaging have been used in Saudi Arabia for years, COVID-19 has led to the implementation and successful use of audiovisual technologies for patients across the country.

Previous limited evidence suggested that telemedicine could provide generally effective, comparable, and satisfactory quality of care as well as improvements in clinical outcomes [3]. However, there is a lack of larger studies of perception and attitude regarding patient-physician interactions, satisfaction with services and convenience of using telemedicine, preference for face-to-face communication, and support for technology infrastructure [4].

Furthermore, few studies have assessed whether the use of telemedicine affects physician well-being and burnout, as telemedicine theoretically provides more flexibility in terms of physician time and geographical location while performing virtual visits [5]. This paper is organized as follows: first, the research team examined the literature for studies related to physician perceptions of telemedicine; second, the method and the survey design were explained; third, the statistical analysis and results were presented; and finally, discussion, limitations, and conclusions and recommendations are presented.

Physicians' Experience Using Telemedicine

The United Kingdom's National Health Service immediately implemented telemedicine as a substitute for face-to-face consultations during the COVID-19 pandemic [6]. Despite widespread unfamiliarity with telemedicine prior to the pandemic, there was a rapid rise in its usage among health care providers [6,7]. Moreover, web-based platforms became indispensable for boosting public health awareness and disseminating information about the pandemic [8].

The implementation of telemedicine, however, has encountered various barriers in Saudi Arabia. These issues include bureaucracy, lack of expertise, inadequate IT infrastructure, absence of guidelines, and insufficient institutional support [9]. Therefore, Saudi Arabia's Ministry of Health must assess health care workers' knowledge and perceptions of telemedicine to facilitate its future implementation while considering patient privacy and confidentiality concerns [10]. To the best of the authors' knowledge, this is the first study to do so.

Perceptions of Patient Experience

The success of any health care delivery system, including telemedicine, is heavily reliant on patient perceptions and satisfaction. Patients are the primary source of information that tells us whether health care is being delivered properly and whether the care they receive meets their expectations [11]. A study conducted by Power [12] on health care consumer satisfaction indicates that 66% of patients are either generally unaware that they can use telemedicine for consultations, or else it is not available to them. Among those who are aware, however, male patients, particularly those aged 18-59 years, are often satisfied with their telemedicine experiences [13]. Female patients, interestingly, are more likely to feel rather neutral toward telemedicine, that is, neither satisfied nor dissatisfied [13]. It is worth noting that almost half of the study's respondents opined that the quality of telemedicine will never be on par with that of traditional in-person care [13]. Many telemedicine applications are available; however, due to a lack of knowledge about telemedicine technology, patients feel uncomfortable using and adapting to it. Therefore, user-friendly telemedicine apps should be developed to improve favorability among patients who feel that in-person care cannot be matched. Additionally, apps should be available in local languages to ensure that patients both learn about telemedicine and experience it in a positive and comfortable manner [14]. Moreover, policy makers around the world should work to boost patients' awareness of telemedicine to ensure correct care during this pandemic and those to come [15].

In one study, most patients who had used telemedicine noted that the practice's convenience and effectiveness helped them to seek treatment from remote areas and optimize their management of type 2 diabetes mellitus [15]. Another study found that the vast majority of patients believe that telemedicine had optimized their type 2 diabetes mellitus management; however, most of these patients noted that improvements could be made to aspects such as user-friendliness, interaction with medical team, and time required for recording or transferring data [16].

Willingness to Use Telemedicine in the Future

Technological advancements have improved existing practices and paved the way for the expansion of telemedicine in the future. Such advancements in telemedicine have increased dependability, lowered costs, improved audiovisual quality, and emulated clinical settings more successfully (eg, by implementing virtual waiting rooms) [17]. There is significant potential in the scalability of telemedicine visits [18]. However, Florea et al [17] revealed that approximately 77% of professionals believe that continual training is essential for health care providers to stay up-to-date with advancements in telemedicine.

Of course, health care institutions are likely to face new ethical challenges arising from the use of telemedicine. For example, they must be capable of protecting patients' private information against potential cyberattacks [17]. Telemedicine could also lead to a rise in malpractice claims stemming from a lack of appropriate guidelines and, in turn, problems with reimbursement [17]. Evidently, significant changes are necessary to fully incorporate telemedicine services into the health care landscape and to fully reap its benefits in advance of future pandemics [17].

Through technological advancements and appropriately oriented policy developments, telemedicine could become a sustainable mainstream solution for both public health emergencies and routine care [9]. Telemedicine may be a reasonable choice for physicians in the future if properly used by patients and if legal guidelines for telemedicine are implemented to address the aforementioned concerns [18].

Effects of Telemedicine on Burnout

Burnout has been defined as a psychological syndrome involving emotional exhaustion, depersonalization, and a sense of reduced personal accomplishment [19]. A 2020 study conducted by Jha et al [19] assessed how COVID-19 has placed several physical and emotional stressors on physicians, which increased physician burnout. The demanding role of primary care physicians (PCPs) in pandemic mitigation measures has made them susceptible to psychological distress. Some PCPs fear being infected by COVID-19; this fear is exacerbated by a lack of personal protective equipment and extended shifts on the front line. All of these issues are in addition to PCPs' existing anxiety stemming from a fast-paced, efficiency-oriented work environment. Many are concerned about errors of omissions and complaints of community residents. Additionally, the documentation process for reporting instances of COVID-19, which is perceived as time-consuming and not conducive to delivering high-quality care, is a source of frustration for most PCPs [20].

Methods

Ethics Approval

This study was approved by the institutional ethical committee of the King Fahad Medical City (IRB log:21-458).

Study Design and Population

This paper presents a cross-sectional study conducted between October and November 2021 among physicians in Riyadh hospitals. A web link survey conducted through SurveyMonkey distributed an anonymous 28-question survey to 500 physicians. After obtaining institutional ethical committee approval, the questionnaire was sent through social media platforms such as WhatsApp, Twitter, and LinkedIn. A total of 362 participants returned the questionnaire, yielding a response rate of 60%. The questionnaire consisted of the following five sections: (1) demographic characteristics; (2) familiarity with telemedicine; (3) perceptions of patients' experiences; (4) willingness to use telemedicine in the future; and (5) the effects of telemedicine on burnout. The responses were measured using a Likert scale. Each respondent's current academic position (consultant, specialist, or intern), specialty, years of postresidency experience, age, sex, frequency of telemedicine use prior to the COVID-19 pandemic, and length of time using telemedicine were all variables of interest.

Survey Design

The instrument tool used in this study was inspired by Malouff et al [21]. The tool was developed to evaluate the perception of physician perceptions and attitudes toward telemedicine. The anonymous survey consisted of 38 questions. The survey was developed in consultation with an expert panel with a consensus informed by elements from existing evidence and models, including the unified theory of acceptance and use of technology, Technology Acceptance Model 2, and diffusion of innovation frameworks [21].

Statistical Analysis

The data were checked for completeness, and all errors were corrected. All of the variables were categorical; therefore, they are presented as frequencies and percentages. The responses of consultants, specialists, and residents—as well as those of men and women—were compared using chi-square tests. The analyses were performed at a 95% confidence interval using SPSS (v.23.0, IBM Corp). Physicians completed the survey including 28 items (sample item: "I find telemedicine has been easy to navigate and use"). The items were rated on a 5-point scale ranging from 1 (strongly agree) to 5 (strongly disagree). The Cronbach alpha for this study was .82, which is considered a good indication for internal consistency [22].

Results

The survey was sent to physicians who experienced telemedicine between October 2021 and December 2021. A total of 362 doctors were included, of whom 28.7% (n=104) were consultants, 30.4% (n=110) were specialists, and 40.9% (n=148) were residents. Male doctors formed the majority 56.1%, (n=203), and female doctors accounted for 43.9% (n=159). When asked about the frequency with which they use telemedicine during the pandemic, 41.4% (n=150) answered "frequently," 26% (n=49) responded "occasionally," and 32.6% (n=118) said "never" (Table 1). Only 25% (n=89) of doctors specified their specialty. Figure 1 shows the specialty distribution; although the physicians were from almost all specialties, they were most frequently from emergency medicine (n=10, 12%) and pediatrics (n=10, 12%).

Moreover, 34% (n=228) agreed or somewhat agreed that the "quality of care during telemedicine is comparable with that of face-to-face visits." Approximately 70% (n=254) believed that telemedicine consultation is a cost-effective means of providing care relative to traditional face-to-face visits. Most of the doctors were skilled at delivering telemedicine 70% (n=163) and capable of independently solving technological issues during telemedicine visits 54% (n=195). Overall, the physicians felt that their patients view telemedicine positively; 68% (n=246) said that their patients felt comfortable using telemedicine, and 76% (n=273) stated that their patients would assert that telemedicine saves time. Regarding burnout, 4.1% (n=15), 7.5% (n=27), and 27.3% (n=99) of the doctors reported feeling

burnout every day, a few times a week, and a few times per month, respectively.

The physicians' responses to the Likert scale prompts are presented in Multimedia Appendix 1. In all, 31% (113/362) of the physicians agreed or somewhat agreed that "telemedicine's quality of care is generally comparable to that which I deliver during face-to-face visits." Approximately 55% (199/362) believed that telemedicine consultations are cost-effective relative to in-person visits. Moreover, 74% (268/362) believed that telemedicine gives them more flexibility or control over how they deliver patient care (Multimedia Appendix 1). Most of the physicians (254/362, 70%) asserted that they are skilled at telemedicine, and are capable of independently solving technological issues during telemedicine visits (195/362, 54%). Furthermore, most of the physicians felt that their patients view telemedicine in a positive manner; 68% (246/362) said that their patients felt comfortable using telemedicine, and 76% (275/362) said that their patients would assert that telemedicine saves time (Figures 2 and 3).

Regarding burnout, 4.1% (n=15), 7.5% (n=27), and 27.3% (n=99) of the physicians felt it every day, a few times per week, and a few times per month, respectively. When asked about the role of telemedicine in burnout, 23.5% (n=85) of physicians said that it alleviated their burnout symptoms. However, 11.6% (n=42) believed that telemedicine contributed to their burnout, and an additional 6.6% (n=24) thought that it substantially contributed to their burnout (Figures 4 and 5).

This study found no statistically significant difference between the burnout frequencies of men and women (P=.57). However, there was a statistically significant difference in burnout frequency between those in different academic positions (P=.002). Whereas 6.1% (9/148) of residents felt burnout every day, only 2.7% (3/110) of specialists and 2.9% (3/104) of consultants felt the same. Interestingly, 18.3% (19/104) of consultants asserted that telemedicine alleviated their burnout symptoms to a sizable degree, whereas only 7.3% (8/110) of specialists and 8.1% (12/148) of residents felt the same (P=.001).

Table 1. Distribution of all physicians by their characteristics and previous experiences with telemedicine.

Characteristics	Values, n (%)
Current position	
Consultant	104 (28.7)
Specialist	110 (30.4)
Resident	148 (40.9)
Clinical experience following residency training (years)	
<5	142 (39.2)
5-10	114 (31.5)
11-20	77 (21.3)
21-30	25 (6.9)
>30	4 (1.1)
Current age (years)	
<30	114 (31.5)
31-40	140 (38.7)
41-50	85 (23.5)
51-60	23 (6.4)
Gender	
Male	203 (56.1)
Female	159 (43.9)
Frequency of telemedicine use before the COVID-19 pandemic	
Frequently (>1 time per month or >12 times per year)	150 (41.4)
Occasionally (1-12 times per year)	94 (26.0)
Never	118 (32.6)
Time spent using telemedicine in any medical capacity (years)	
<1	154 (42.5)
2-3	106 (29.3)
>3	102 (28.2)

Figure 1. Specialty distribution. ENT: ear, nose, and throat; ICU: intensive care unit; Obgyn: obstetrics and gynecology.

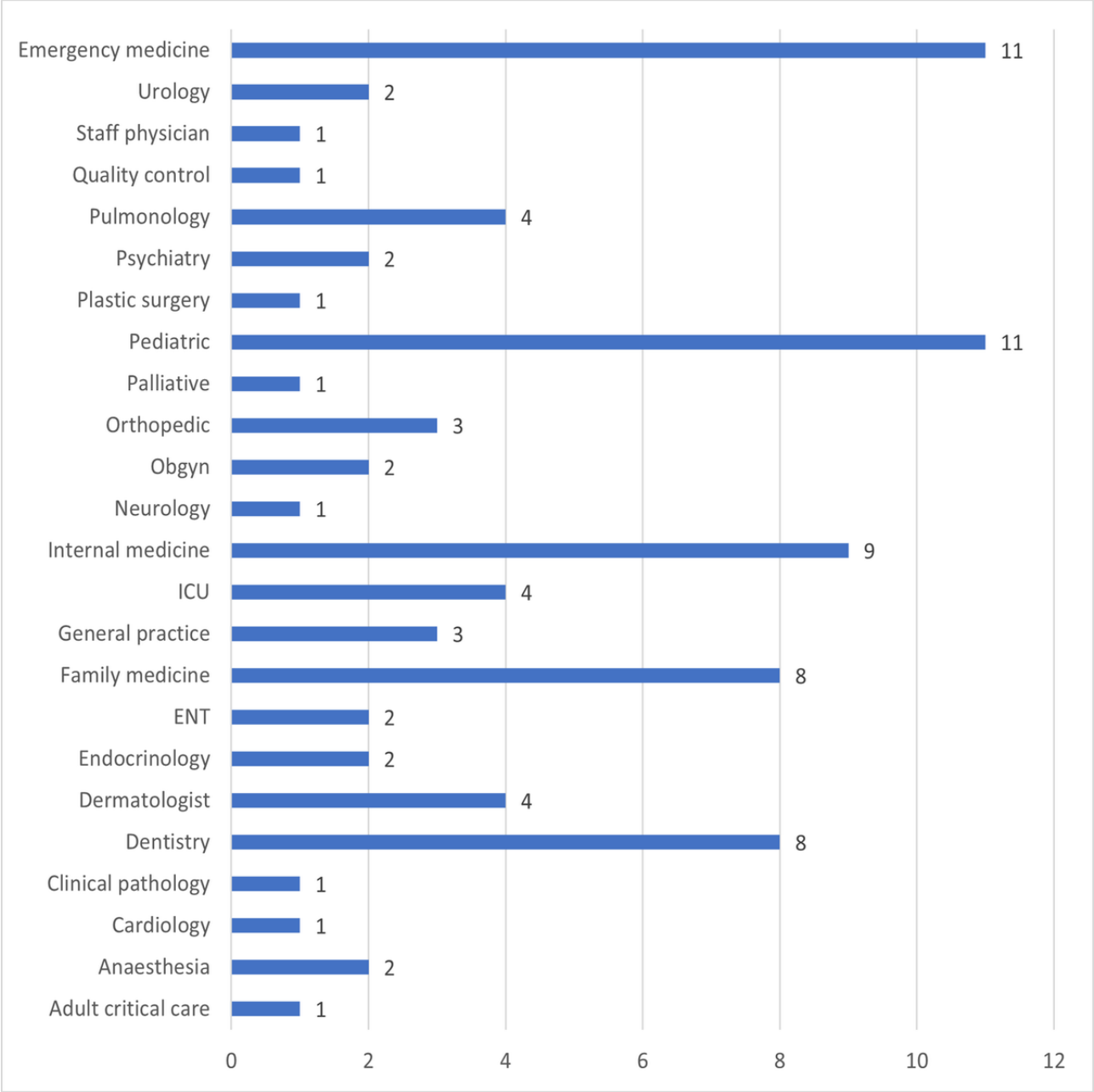


Figure 2. Physicians' attitudes toward the quality and potential advantages of telemedicine.

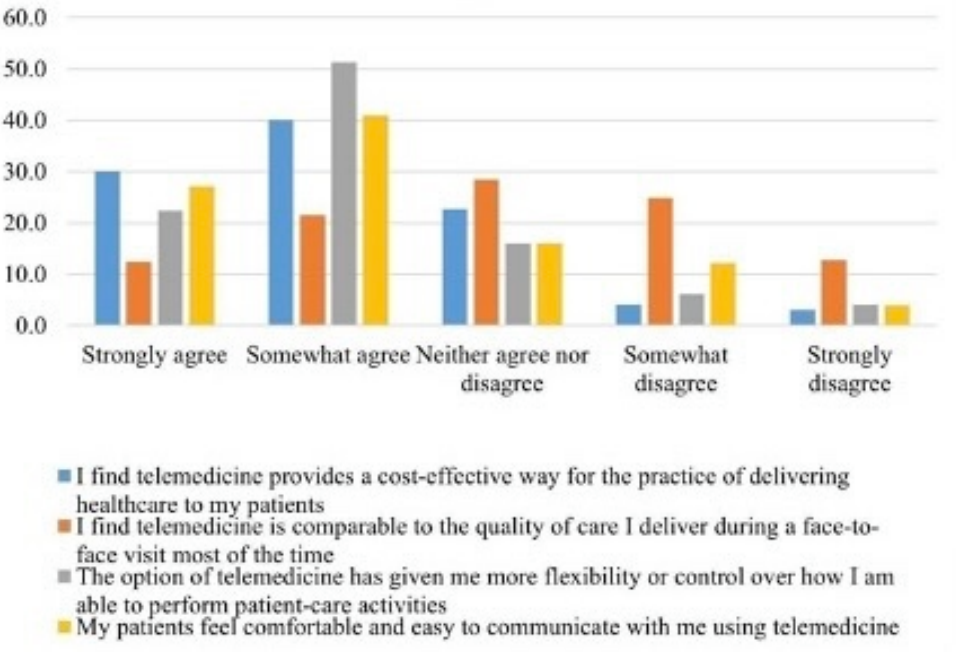


Figure 3. Physicians' openness to using telemedicine after the COVID-19 pandemic.

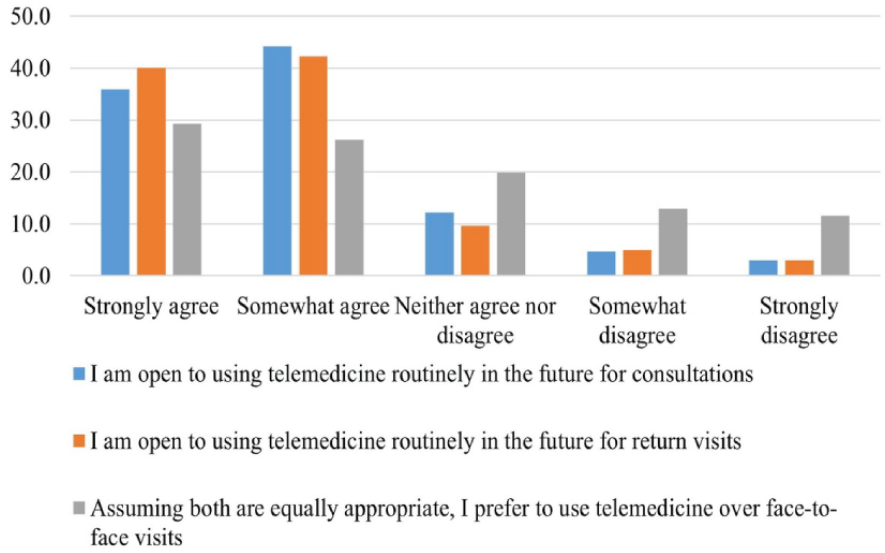


Figure 4. Responses regarding physician burnout and the influence of telemedicine: "What role has telemedicine played in your experience of burnout?" N/A: not applicable.

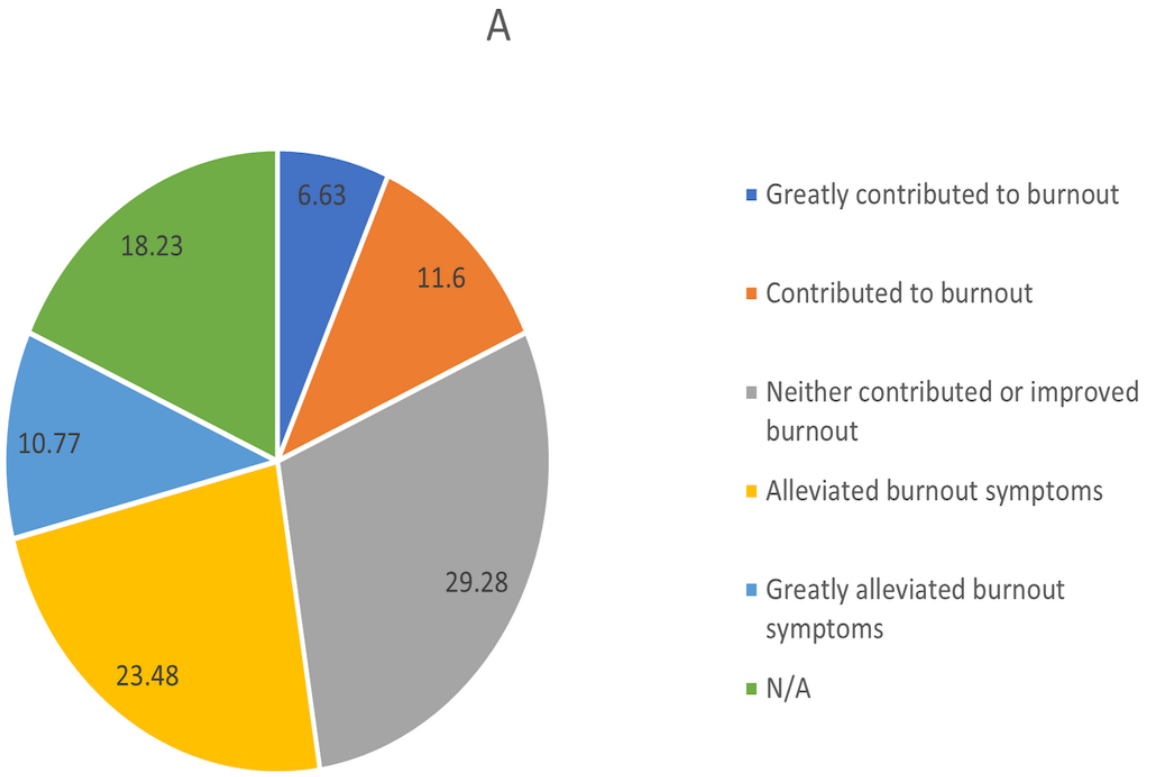
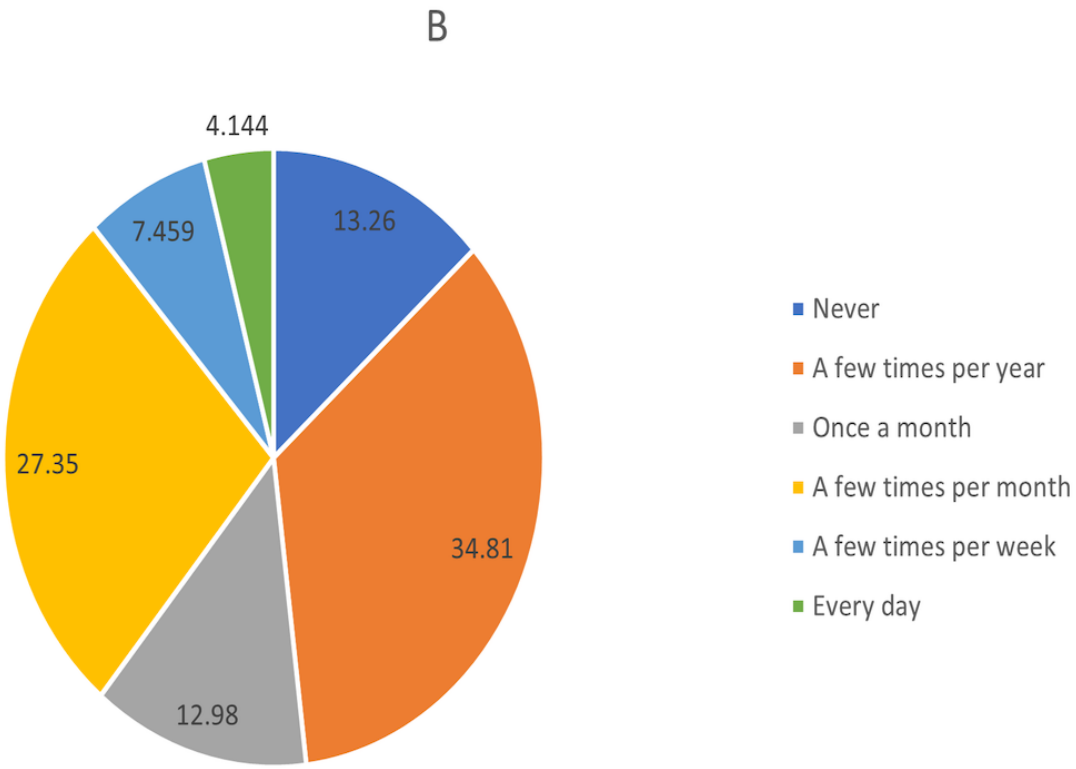


Figure 5. Responses regarding physician burnout and the influence of telemedicine: "I feel burnout from work".



Discussion

Principal Findings

Our data suggest that physicians in Riyadh have adapted to the implementation of telemedicine. Most of the physicians assessed in this study, regardless of their specialty, considered telemedicine to be easy to navigate and use. Furthermore, almost all respondents were open to using telemedicine for routine consultations and follow-ups even after the pandemic ends; in fact, more than half of them actively preferred telemedicine to in-person visits. The findings of this work align with those of previous studies. For example, Gillman-Wells et al [6] found that 70% of plastic surgeons surveyed in the United Kingdom have embraced the use of telemedicine. A survey by Srinivasan et al [7] found that personnel at Stanford University's primary care clinics strongly believe that telemedicine visits should be an ongoing element of health care delivery after the pandemic.

There are several potential reasons for why physicians are willing to adapt to this new technology, including cost-effectiveness, time efficiency for physicians and patients alike, as well as flexibility in scheduling telemedicine visits, all of which may improve physicians' quality of life [21]. These advantages balance out concerns over lower quality of care. According to a study conducted during the first year of the pandemic, one major shortfall of telemedicine that physicians should be aware of is the absence of physical examination [23].

As mentioned in the literature review, the cost-effectiveness of telemedicine relative to traditional in-person care is a significant driver of positive attitudes toward telemedicine; in this study, 70% (254/362) of the physicians agreed that it is a cost-effective way to deliver health care (in a study by the Mayo Clinic, 80% of the respondents agreed with this sentiment) [21]. To assess the verity of this perception, a study evaluated postoperative visit costs and found that patients who used telemedicine services saved an average of US $888 per return visit (increasing to US $1501 when accounting for travel and accommodation costs). The authors reported savings of US $256 per visit, even for patients who did not require accommodation [24]. Furthermore, the Pediatric Cardiology Service at the Coimbra University Hospital Center analyzed telemedicine use in Portugal since 1998 and found that it had saved the country's health care system about EUR €1.1 million (US $1.3 million) overall, equating to savings of approximately EUR €419 (US $500) per patient [25]

Telemedicine also offers considerable efficiency for both physicians and patients. In this study, nearly three-quarters of the respondents agreed that telemedicine enhanced flexibility. Telemedicine visits can be conducted anywhere (including from home), and this ability helps physicians to balance their professional and personal needs, particularly during the COVID-19 pandemic. According to a study conducted by Chaudhry et al [26], telemedicine visits may also reduce some time-consuming activities that are common at clinics, such as waiting for rooms to become available, checking patients in, and moving patients from one room to another. Orthopedic patients reported that telemedicine saved them time, both when including (180 minutes) and excluding (17 minutes) travel time.

Regardless of the benefits that many physicians derive from telemedicine, more than one-third of this survey's respondents did not agree that telemedicine is equivalent to in-person visits in terms of quality of care. This finding aligns with those of previous studies. Although telemedicine can be efficient and cost-effective, physicians lose the ability to conduct physical examinations, which are often crucial in order to meet patients' needs and deliver appropriate preventative care. Furthermore, the patient-physician relationship largely depends on face-to-face visits. However, because telemedicine is even more common at the time of writing this paper compared to when this study was conducted, further surveys are necessary to understand physicians' concerns and their relative importance.

Indeed, the most common concerns over telemedicine pertained to an inability to provide care on the same level as traditional in-person care. Zhang et al [5] considered this concern when examining the Memorial Sloan Kettering Cancer Center during the COVID-19 pandemic. They found that 92% of the center's radiation oncology visits were conducted via telemedicine at the peak of the pandemic [5]. Overall, 71% of the providers reported no difference in their ability to treat cancer appropriately, and 55% of the patients reported no difference in their overall visit quality [5].

Although our study did not directly examine patients' experiences of differences between virtual and in-person consultations regarding lab tests, imaging exams, and prescribed medications, a study on Stanford's ClickWell Care clinic evaluated practice patterns for both telemedicine and in-person visits. It found no difference in laboratory tests, imaging tests, or prescriptions ordered between virtual and in-person visits for 17 of the most common diagnoses. However, overall, there were more laboratory and imaging tests ordered following in-person visits for all diagnoses; this increase may have affected general medical examinations [27].

Another widely reported concern when using telemedicine was the absence of a physical evaluation, which remains an essential element of follow-up care, mainly when assessing adverse posttherapy events for patients with physical disabilities [28]. To counter this limitation, some physicians are working to develop evaluations that can function via telemedicine, such as a neurosurgical spine examination that can be conducted remotely [29]. Early evidence on the feasibility and comparability of such examinations is promising. In addition, Laskowski et al [28] developed a specific set of guidelines to enhance evaluations of the musculoskeletal system when performing virtual examination.

Physical examinations are also correlated with patient satisfaction, which is critical in telemedicine because cost and time savings are meaningless if patients do not believe that they are receiving high-quality care [21]. Although this study did not directly survey patients, approximately 75% of the physicians felt that their patients were at ease communicating with them via telemedicine, with half agreeing that their patients found the technology easy to use and comparable in quality of care with face-to-face visits. These findings align with those of previous studies. For example, a study by Elawady et al [30] found that 73% of the physician respondents felt that their

patients understood their medical conditions and the corresponding recommendations given to them over the phone. In addition, physicians were asked whether videoconference consultation would improve patient care over telephone consultation alone, and 70% of the respondents agreed that it would [30]. Moreover, according to a meta-analysis conducted by Chaudhry et al [26], there were no differences in surgeon satisfaction or patient-reported outcome measurements when comparing telemedicine visits with in-person visits [26].

Although patient satisfaction with telemedicine is critical, reduction of burnout among physicians by boosting flexibility in care delivery is one potential benefit of telemedicine. Telemedicine may reduce transportation time, granting physicians more time for sleep, family life, and social activities, all of which are key factors in avoiding burnout [21]. This study found that more than one-third of the surveyed physicians' burnout symptoms were alleviated or substantially alleviated due to telemedicine. However, it is important to note that these figures may have been influenced by other stress factors related to the COVID-19 pandemic [21]. Further studies may be needed to evaluate whether—and to what degree—burnout can be mitigated through the postpandemic use of telemedicine.

According to Malouff et al [21], looking ahead, telemedicine will be an essential element of post–COVID-19 crises. Virtual reality has been proposed as a means of improving feelings of physical presence during examinations to compensate for the lack of physical presence during virtual visits [21]. In addition, telemedicine may be beneficial for people who are fearful of visiting clinics or hospitals; patients with anxiety or depression may prefer telemedicine to in-person visits. Finally, telemedicine may offer an opportunity for underrepresented populations to participate in clinical trials because follow-ups and toxicity supervision can be conducted virtually [27].

Limitations

Despite this study's finding that physicians' perceptions of telemedicine are generally favorable, it comes with certain limitations. For instance, this was a survey-based study, meaning that it is subject to the typical limitations of survey-based evaluations, including incomplete responses and a low response rate. It was unable to precisely determine response rates because the survey link was partially distributed through social media platforms. Another major limitation of this study is that it does not directly survey patients. Furthermore, the results of this study are not generalizable to a wider health care population, given its small sample size and the sample size of the subgroups that were examined when comparing survey responses.

Conclusions and Recommendations

This study shows that physicians in Riyadh, Saudi Arabia, have generally favorable attitudes toward the adoption of telemedicine because they believe that the quality of care delivered using telemedicine is comparable to that delivered using traditional methods. However, further research is necessary to correctly assess how the COVID-19 pandemic influenced physicians' attitudes toward telemedicine and how telemedicine can be used to advance care delivery and improve patient outcomes in the future.

Acknowledgments

The authors extend their appreciation to the Researchers Supporting Project number (RSP-2021/332), King Saud University, Riyadh, Saudi Arabia.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Physicians' thoughts, attitudes, and capacities pertaining to telemedicine.
[DOCX File , 20 KB - formative_v6i7e36029_app1.docx]

References

1. Altulaihi B, Alharbi K, Alhassan A, Altamimi A, Al Akeel MA. Physician's Perception Toward Using Telemedicine During COVID-19 Pandemic in King Abdulaziz Medical City, Riyadh, Saudi Arabia. Cureus 2021 Jul;13(7):e16107 [FREE Full text] [doi: 10.7759/cureus.16107] [Medline: 34350074]
2. Bhaskar S, Bradley S, Chattu VK, Adisesh A, Nurtazina A, Kyrykbayeva S, et al. Telemedicine as the New Outpatient Clinic Gone Digital: Position Paper From the Pandemic Health System REsilience PROGRAM (REPROGRAM) International Consortium (Part 2). Front Public Health 2020 Sep 7;8:410 [FREE Full text] [doi: 10.3389/fpubh.2020.00410] [Medline: 33014958]
3. Ekeland AG, Bowes A, Flottorp S. Effectiveness of telemedicine: a systematic review of reviews. Int J Med Inform 2010 Nov;79(11):736-771. [doi: 10.1016/j.ijmedinf.2010.08.006] [Medline: 20884286]
4. 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]

5. Zhang H, Cha EE, Lynch K, Cahlon O, Gomez DR, Shaverdian N, et al. Radiation Oncologist Perceptions of Telemedicine from Consultation to Treatment Planning: A Mixed-Methods Study. Int J Radiat Oncol Biol Phys 2020 Oct 01;108(2):421-429 [FREE Full text] [doi: 10.1016/j.ijrobp.2020.07.007] [Medline: 32890525]
6. Gillman-Wells CC, Sankar TK, Vadodaria S. COVID-19 Reducing the Risks: Telemedicine is the New Norm for Surgical Consultations and Communications. Aesthetic Plast Surg 2021 Feb 03;45(1):343-348 [FREE Full text] [doi: 10.1007/s00266-020-01907-8] [Medline: 32885319]
7. Srinivasan M, Asch S, Vilendrer S, Thomas SC, Bajra R, Barman L, et al. Qualitative Assessment of Rapid System Transformation to Primary Care Video Visits at an Academic Medical Center. Annals of Internal Medicine 2020 Oct 06;173(7):527-535. [doi: 10.7326/m20-1814]
8. Li X, Liu Q. Social Media Use, eHealth Literacy, Disease Knowledge, and Preventive Behaviors in the COVID-19 Pandemic: Cross-Sectional Study on Chinese Netizens. J Med Internet Res 2020 Oct 09;22(10):e19684 [FREE Full text] [doi: 10.2196/19684] [Medline: 33006940]
9. El Kheir DYM, Boumarah DN, Bukhamseen FM, Masoudi JH, Boubshait LA. The Saudi Experience of Health-Related Social Media Use: A Scoping Review. Saudi J Health Syst Res 2021 Aug 17;1(3):81-92. [doi: 10.1159/000516473]
10. Albarrak AI, Mohammed R, Almarshoud N, Almujalli L, Aljaeed R, Altuwaijiri S, et al. Assessment of physician's knowledge, perception and willingness of telemedicine in Riyadh region, Saudi Arabia. J Infect Public Health 2021 Jan;14(1):97-102 [FREE Full text] [doi: 10.1016/j.jiph.2019.04.006] [Medline: 31060975]
11. Berger S, Saut AM, Berssaneti FT. Using patient feedback to drive quality improvement in hospitals: a qualitative study. BMJ Open 2020 Oct 23;10(10):e037641 [FREE Full text] [doi: 10.1136/bmjopen-2020-037641] [Medline: 33099495]
12. Telehealth: Best Consumer Healthcare Experience You've Never Tried, Says J.D. Power Study. J.D. Power. 2019. URL: https://www.jdpower.com/business/press-releases/2019-us-telehealth-satisfaction-study [accessed 2022-04-27]
13. Alharbi K, Aldosari M, Alhassan A, Alshallal K, Altamimi A, Altulaihi B. Patient satisfaction with virtual clinic during Coronavirus disease (COVID-19) pandemic in primary healthcare, Riyadh, Saudi Arabia. J Fam Community Med 2021;28(1):48. [doi: 10.4103/jfcm.jfcm_353_20]
14. Yamin MAY, Alyoubi BA. Adoption of telemedicine applications among Saudi citizens during COVID-19 pandemic: An alternative health delivery system. J Infect Public Health 2020 Dec;13(12):1845-1855 [FREE Full text] [doi: 10.1016/j.jiph.2020.10.017] [Medline: 33172819]
15. Arumugam S, Ramadoss D, Brindhadevi P, Easwaran S, Kumari SS. COVID-19 Lockdown Period: Perception of Doctors regarding Telemedicine Use for General Practice. JCDR 2021:FC09-FC12. [doi: 10.7860/jcdr/2021/46549.14516]
16. Rodríguez-Fortúnez P, Franch-Nadal J, Fornos-Pérez JA, Martínez-Martínez F, de Paz HD, Orera-Peña ML. Cross-sectional study about the use of telemedicine for type 2 diabetes mellitus management in Spain: patient's perspective. The EnREDa2 Study. BMJ Open 2019 Jun 22;9(6):e028467 [FREE Full text] [doi: 10.1136/bmjopen-2018-028467] [Medline: 31230025]
17. Florea M, Lazea C, Gaga R, Sur G, Lotrean L, Puia A, et al. Lights and Shadows of the Perception of the Use of Telemedicine by Romanian Family Doctors During the COVID-19 Pandemic. IJGM 2021 Apr;Volume 14:1575-1587. [doi: 10.2147/ijgm.s309519]
18. Bashshur R, Doarn CR, Frenk JM, Kvedar JC, Woolliscroft JO. Telemedicine and the COVID-19 Pandemic, Lessons for the Future. Telemed J E Health 2020 May 01;26(5):571-573. [doi: 10.1089/tmj.2020.29040.rb] [Medline: 32275485]
19. Soin A. The Effect of COVID-19 on Interventional Pain Management Practices: A Physician Burnout Survey. Pain Phys 2020 Aug 15;4S;23(8;4S):S271-S282. [doi: 10.36076/ppj.2020/23/s271]
20. Xu Z, Ye Y, Wang Y, Qian Y, Pan J, Lu Y, et al. Primary Care Practitioners' Barriers to and Experience of COVID-19 Epidemic Control in China: a Qualitative Study. J Gen Intern Med 2020 Nov 31;35(11):3278-3284 [FREE Full text] [doi: 10.1007/s11606-020-06107-3] [Medline: 32869200]
21. Malouff TD, TerKonda SP, Knight D, Abu Dabrh AM, Perlman AI, Munipalli B, et al. Physician Satisfaction With Telemedicine During the COVID-19 Pandemic: The Mayo Clinic Florida Experience. Mayo Clin Proc Innov Qual Outcomes 2021 Aug;5(4):771-782 [FREE Full text] [doi: 10.1016/j.mayocpiqo.2021.06.006] [Medline: 34226884]
22. Gliem JA, Gliem RR. Calculating, Interpreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales. ScholarWorks. 2003. URL: https://scholarworks.iupui.edu/bitstream/handle/1805/344/Gliem%20%26%20Gliem.pdf?sequence=1&isAllowed=y [accessed 2022-06-20]
23. Iyengar K, Jain VK, Vaishya R. Pitfalls in telemedicine consultations in the era of COVID 19 and how to avoid them. Diabetes Metab Syndr 2020 Sep;14(5):797-799 [FREE Full text] [doi: 10.1016/j.dsx.2020.06.007] [Medline: 32534432]
24. Demaerschalk BM, Cassivi SD, Blegen RN, Borah B, Moriarty J, Gullerud R, et al. Health Economic Analysis of Postoperative Video Telemedicine Visits to Patients' Homes. Telemed J E Health 2021 Jun 01;27(6):635-640. [doi: 10.1089/tmj.2020.0257] [Medline: 32907513]
25. Maia MR, Castela E, Pires A, Lapão LV. How to develop a sustainable telemedicine service? A Pediatric Telecardiology Service 20 years on - an exploratory study. BMC Health Serv Res 2019 Sep 23;19(1):681. [doi: 10.1186/s12913-019-4511-5] [Medline: 31547824]
26. Chaudhry H, Nadeem S, Mundi R. How Satisfied Are Patients and Surgeons with Telemedicine in Orthopaedic Care During the COVID-19 Pandemic? A Systematic Review and Meta-analysis. Clin Orthop Relat Res 2020 Sep 28;479(1):47-56. [doi: 10.1097/corr.0000000000001494]

27. Norden JG, Wang JX, Desai SA, Cheung L. Utilizing a novel unified healthcare model to compare practice patterns between telemedicine and in-person visits. Digit Health 2020 Sep 17;6:2055207620958528 [FREE Full text] [doi: 10.1177/2055207620958528] [Medline: 32995039]
28. Laskowski ER, Johnson SE, Shelerud RA, Lee JA, Rabatin AE, Driscoll SW, et al. The Telemedicine Musculoskeletal Examination. Mayo Clin Proc 2020 Aug;95(8):1715-1731 [FREE Full text] [doi: 10.1016/j.mayocp.2020.05.026] [Medline: 32753146]
29. Piche J, Butt BB, Ahmady A, Patel R, Aleem I. Physical Examination of the Spine Using Telemedicine: A Systematic Review. Global Spine J 2021 Sep 22;11(7):1142-1147 [FREE Full text] [doi: 10.1177/2192568220960423] [Medline: 32959711]
30. Elawady A, Khalil A, Assaf O, Toure S, Cassidy C. Telemedicine during COVID-19: a survey of Health Care Professionals' perceptions. Monaldi Arch Chest Dis 2020 Sep 22;90(4):576-581 [FREE Full text] [doi: 10.4081/monaldi.2020.1528] [Medline: 32959627]

Abbreviations

PCP: primary care physicians

Edited by A Mavragani; submitted 29.12.21; peer-reviewed by A Ren, V Castillo; comments to author 10.04.22; revised version received 15.05.22; accepted 05.06.22; published 12.07.22.

Please cite as:
Idriss S, Aldhuhayyan A, Alanazi AA, Alasaadi W, Alharbi R, Alshahwan G, Baitalmal M, Alonazi W
Physicians' Perceptions of Telemedicine Use During the COVID-19 Pandemic in Riyadh, Saudi Arabia: Cross-sectional Study
JMIR Form Res 2022;6(7):e36029
URL: https://formative.jmir.org/2022/7/e36029
doi:10.2196/36029
PMID:35714293

Original Paper

eHealth Literacy and its Associated Factors Among Health Professionals During the COVID-19 Pandemic in Resource-Limited Settings: Cross-sectional Study

Getanew Aschalew Tesfa[1], MPH; Delelegn Emwodew Yehualashet[1], MPH; Helen Ali Ewune[1], MSc; Addisu Getnet Zemeskel[2], MSc; Mulugeta Hayelom Kalayou[3], MPH; Binyam Tariku Seboka[1], MPH

[1]School of Public Health, College of Medicine and Health Science, Dilla University, Dilla, Ethiopia

[2]Department of Midwifery, College of Medicine and Health Science, Dilla University, Dilla, Ethiopia

[3]Department of Health Informatics, School of Public Health, College of Medicine and Health Sciences, Wollo University, Wollo, Ethiopia

Corresponding Author:
Getanew Aschalew Tesfa, MPH
School of Public Health
College of Medicine and Health Science
Dilla University
15th Road
Dilla, 419
Ethiopia
Phone: 251 31167972
Email: getanewaschalew@gmail.com

Abstract

Background: The COVID-19 pandemic has wreaked havoc on health care systems and governments worldwide. Although eHealth literacy is acknowledged as a critical component of public health, it was overlooked during the pandemic. To assist patients and their families, health professionals should be knowledgeable about online health information resources and capable of evaluating relevant online information. In a resource-constrained situation, the level of eHealth literacy among health professionals is not well documented.

Objective: The aim of this study was to assess the eHealth literacy level and its associated factors among health professionals working in Amhara regional state teaching hospitals, Ethiopia.

Methods: A self-administered questionnaire was used in an institutional-based cross-sectional study design. Descriptive statistics were calculated to describe eHealth literacy statements and key variables using SPSS v.24. Bivariable and multivariable logistic regression models were fit to identify factors related to eHealth literacy. Variables with $P<.05$ were declared to be statistically significant predictors.

Results: A total of 383 participants completed and returned the questionnaire with a response rate of 90.5%. Health professionals demonstrated a moderate level of eHealth literacy (mean 29.21). Most of the professionals were aware of the available health resources located on the internet, and know how to search and locate these resources. However, they lack the ability to distinguish high-quality health resources from low-quality resources. Factors that were significantly associated with eHealth literacy were computer access, computer knowledge, perceived ease of use, and perceived usefulness of eHealth information resources.

Conclusions: It is crucial to provide training and support to health care workers on how to find, interpret, and, most importantly, evaluate the quality of health information found on the internet to improve their eHealth literacy level. Further research is needed to explore the role of eHealth literacy in mitigating pandemics in developing countries.

(JMIR Form Res 2022;6(7):e36206) doi:10.2196/36206

KEYWORDS

eHealth literacy; COVID-19; health information; health literacy

Introduction

The internet is currently one of the most widely used tools for obtaining information about health care and medical conditions [1], providing health professionals with unprecedented access to a massive volume of relevant and high-quality current health care information [2]. Moreover, the internet has a significant impact on health and health care since it can improve health care delivery and help decision-making for health care workers [3]. For modern health care recipients, the internet provides a useful and accessible source of health-related information [4].

Health literacy is vital for people who interact with the digital world, with its diverse information and sources [5], and plays an important role in evaluating online health information [6], especially during pandemics such as COVID-19, the disease caused by infection with the novel coronavirus SARS-CoV-2, which has posed unprecedented challenges worldwide [7,8]. The COVID-19 pandemic has been accompanied by the rapid spread of disinformation or fake news via social media platforms and other outlets. The dissemination of this misinformation could lead to people acting inappropriately, thereby jeopardizing governments' and health authorities' efforts to manage the pandemic [9]. To control this worldwide health catastrophe, eHealth literacy is recommended among the interdisciplinary and multidimensional techniques [9]. Individuals with adequate eHealth literacy are more likely to utilize the internet to obtain health-related information and believe they are capable of applying web-based knowledge to improve their health [10].

eHealth literacy is defined as the ability to seek out, find, interpret, evaluate, and appraise health information from electronic sources, as well as apply that knowledge to address or solve a health problem. This composite skill necessitates the ability to interact with technology, think critically about media and science concerns, and navigate a large assortment of information tools and sources to obtain the information needed to make decisions [5].

The eHealth Literacy Scale (eHEALS) was developed in response to the need to assess eHealth literacy in a variety of populations and settings. The eHEALS is a self-report tool that can be administered by a health professional and is based on a person's perception of their skills and knowledge in each of the measured domains. The test is intended to offer a broad estimate of consumer eHealth abilities that can be used to guide clinical decision-making and health-promotion planning with individuals or groups. The eHEALS has potential to be used in a clinical setting to identify individuals who may or may not benefit from referrals to an eHealth intervention or resource [1,5].

Controlling pandemics requires strong enabling environments as well as modern and digitized health information systems. Beyond the COVID-19 pandemic, it is critical to promote digital solutions [11]. Countries with health information systems that combine data from the health and long-term care sectors are likely to be better prepared to deal with this challenge [12]. The use of online consultations from hospitals and health care centers has been found to be a safe and effective way to mitigate the pandemic's negative effects [13].

The outbreak of COVID-19 has made eHealth literacy more vital than ever [9]. However, this has been an underestimated issue during the pandemic [14]. eHealth literacy has a direct effect on health-promoting behaviors by improving health information–seeking behavior, which can ultimately lead to health-promoting behavior and health outcomes [15]. eHealth literacy has the potential to increase adherence to infection prevention and control measures, promote healthy habits, and maintain health care workers' health. This would help contain the COVID-19 pandemic and further mitigate its impacts [16]. The world learned via COVID-19 that eHealth is not an optional or excessive approach to health care but rather a crucial, safer, and effective means of providing health care to both individuals with underlying problems and others during such periods. Because it is critical to obtain accurate information from reputable sources when self-managing diseases, individuals' health perceptions and behaviors are negatively influenced by misinformation [17,18].

Aside from having a basic understanding of how to utilize the internet and eHealth literacy, health care practitioners should be knowledgeable on how to evaluate sources of information as reputable sources of information [19]. According to a study conducted at the University of Gondar (UOG) specialized hospital, patients have a poor level of eHealth literacy, implying that there is a need to bridge the skill gap [20]. This study could imply that this is where health care professionals can help patients make health-related decisions, specifically to assist patients and families in getting up-to-date, reliable, and quality health information, and identifying and analyzing suitable web sources for such decisions. However, health care practitioners must be eHealth literate to provide this assistance [21]. The potential predictor variables of this study were identified based on the review of findings from other related literature. Previous studies showed that variables such as age, sex, professional background, work experience, training about information retrieval techniques, computer knowledge, perception toward web resources, and computer accessibility were significantly associated with eHealth literacy [16,22-26].

The expansion of smartphone penetrations, the growing number of internet users, and information needs in developing nations are the most compelling reasons for assessing eHealth literacy and its associated factors among health care professionals to maximize eHealth benefits. However, there is limited information on eHealth literacy among health care workers in Ethiopia. Therefore, the aim of this study was to assess eHealth literacy among health professionals at Amhara regional state teaching hospitals, as well as to identify factors that influence eHealth literacy. We postulated the following three key hypotheses.

Hypothesis 1: computer knowledge positively correlates with eHealth literacy.

Hypothesis 2: there is a significant positive association between eHealth literacy and perception (perceived usefulness and perceived ease of use).

Hypothesis 3: The accessibility of computers is positively linked with eHealth literacy.

Methods

Study Area, Design, and Period

The study was conducted among health professionals working at Amhara regional state teaching hospitals in Ethiopia from February 23 to May 10, 2020, using an institutional-based cross-sectional study design. Ethiopia is divided into nine regions and two city administrations. The Amhara region is the country's second-largest and most populous among these regions [27]. The Tibebe Ghion specialized teaching hospital in Bahir Dar city and UOG specialized teaching hospital in the town of Gondar are the two specialized teaching hospitals in the region. At the time of the survey, the UOG specialized teaching hospital had 1076 permanent employees, whereas the Tibebe Ghion specialized teaching hospital had 738 permanent employees.

Study Procedure

The study's sample size was calculated using the following single-population proportion formula by assuming that 50% of health professionals have a high degree of eHealth literacy and a 10% nonresponse rate.

$$[Za^2 \times p(1-p)]/d^2=[(1.96)^2 \times 0.5(1-0.5)]/0.05^2=384.4+38.4=423$$

Where Za^2 is the Z statistic (value of the standard normal distribution) at 95% confidence, d is the margin of error at 5%, and p is the single-population proportion.

Participants of the study were selected from the UOG and Tibebe Ghion specialized teaching hospitals. To guarantee a fair distribution of the entire sample, health care workers from each hospital were stratified based on profession and then study participants were chosen using a simple random sampling technique.

Measures

eHealth Literacy

eHealth literacy was measured using the eHEALS, which was introduced by Norman and Skinner [5] to determine consumers' combined knowledge and perceived skill in finding, evaluating, and applying eHealth information to health problems. The eHEALS has eight items that are scored on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree), and the total score ranges between 5 and 40. This scale was reported as a reliable tool with a Cronbach α coefficient of .88 in the original study [5] and had a Cronbach α of .847 in this study. A high eHEALS score indicates a high eHealth literacy level, whereas a low eHEALS score indicates low eHealth literacy.

Computer Access

Computer access was measured by a yes-or-no question; health professionals were asked whether they can access computers in their working area/hospital or not.

Perceived Usefulness

Perceived usefulness was measured using five closed questions, and study participants who scored the mean and above on the 5-point Likert scale questions were categorized according to whether they considered that using eHealth information resources to be useful or not useful.

Perceived Ease of Use

Perceived ease of use refers to the degree to which a person believes that using eHealth information resources is free of effort, which was measured using four closed questions. Study participants who scored the mean and above on the 5-point Likert scale questions were categorized as to whether or not they considered that using eHealth information resources is easy.

Computer Knowledge

Computer knowledge was assessed based on participants' self-perceived reports; respondents were asked to rate their level of computer knowledge according to five categories (none, beginner, below average, average, above average) with respect to basic computer skills and internet navigation. According to these responses, the participants were dichotomized as having poor or good computer knowledge for further analysis.

Data Collection

Self-administered questionnaires were used to collect the data. The questionnaire included 29 questions divided into four sections: sociodemographic characteristics (six items), the eHEALS (eight items), perceived usefulness of eHealth information resources (five items), and perceived ease of use (four items).

Data collectors received training on data collection techniques, maintaining the confidentiality of health professionals' information, obtaining informed consent, study participants' rights, and all study protocols. Before the actual data collection, a pretest was performed among 5% of the entire sample population outside the study area, and any required corrections and revisions to the questionnaire were implemented. The completeness of the questionnaire was checked daily.

Data Processing and Analysis

Epi-info version 7 was used for data entry. Data cleaning and coding were performed before any statistical analysis. Data analysis was performed with SPSS version 20 software. To describe the participant characteristics and study objectives, key variables are summarized in terms of descriptive frequencies and mean (SD).

Model fit was checked using the Hosmer and Lemeshow goodness-of-fit test. Bivariate analysis was then used to examine the relationship between individual independent variables and the dependent variable. Variables with $P \leq .20$ in the bivariate analysis were then entered into a multivariable logistic regression model to examine the relationship between selected independent variables and the outcome variable. In the multivariate analysis, variables with $P < .05$ according to the odds ratio were determined to be statistically significant independent predictors.

Ethical Statement

On behalf of the UOG College of Medicine and Health Sciences, ethical approval was secured from the Institute of Public Health (number IPH/840/02/2020). Written consent was obtained from

each study participant. The data were collected anonymously and participants' privacy was respected. Privacy of all of the information gathered was maintained and the data were solely utilized for research purposes.

Results

Characteristics of the Study Participants

With a response rate of 90.5%, 383 of the total disseminated questionnaires were returned. The majority of the 383 study participants were male (Table 1). The mean age of the participants was 28.3 (SD 3.37) years. Nurses accounted for the highest proportion of professionals, followed by medical doctors and midwives. Less than 30% of the participants had received information-retrieval training on eHealth information sources (Table 1).

Table 1. Sociodemographic characteristics of the study participants (N=383).

Variables	Participants, n (%)
Sex	
Male	239 (62.4)
Female	144 (37.6)
Age group (years)	
20-24	18 (4.7)
25-29	254 (66.3)
≥30	111 (29.0)
Professional background	
Nurse	158 (41.3)
Medical doctor	94 (24.5)
Pharmacist	30 (7.8)
Midwife	54 (14.1)
Laboratory	24 (6.3)
Other	23 (6.0)
Work experience (years)	
1-3	226 (59.0)
4-6	119 (31.1)
≥7	38 (9.9)
Computer access	
No	166 (43.3)
Yes	217 (56.7)
Received training on information retrieval	
No	269 (70.2)
Yes	114 (29.8)

eHealth Literacy

The overall mean score for eHealth literacy was 29.21 (SD 7.08), which is considered to be moderate; among the 383 health professionals surveyed, 225 (58.7%) had high eHealth literacy and the other 41.3% (n=158) had low eHealth literacy, defined as those scoring above and below the mean on the eHEALS, respectively. As shown in Table 2, 240 (62.7%) of the total survey participants reported knowing what health resources are available on the internet and 235 (61.3%) of the 383 participants agreed with the statement "I know where to find helpful health resources on the internet."

Only 221 (57.7%) of the 383 participants claimed that they could identify high-quality health resources from low-quality resources, indicating that almost half of the health professionals have trouble distinguishing quality health resources on the internet; 231 (60.3%) of the participants reported that they have the skills to evaluate health resources they found on the internet (Table 2).

Based on the participants' background characteristics, the majority of both the male and female participants had a high level of eHealth literacy. By contrast, nearly half of the nurses and one-third of medical doctors had a lower level of eHealth literacy (Table 3).

Table 2. Distribution of eHealth Literacy Scale (eHEALS) responses (N=383).

eHEALS statements	Rating scale, n (%)					Mean score
	Strongly disagree	Disagree	Unsure	Agree	Strongly agree	
I know what health resources are available on the internet	7 (1.8)	47 (12.3)	89 (23.2)	196 (51.2)	44 (11.5)	3.58
I know where to find helpful health resources on the internet	11 (2.9)	50 (13.1)	87 (22.7)	161 (42.0)	74 (19.3)	3.62
I know how to find helpful health resources on the internet	8 (2.1)	30 (7.8)	111 (29.0)	171 (44.6)	63 (16.4)	3.66
I know how to use the health information I find on the internet to help me	13 (3.4)	31 (8.1)	94 (24.5)	163 (42.6)	82 (21.4)	3.70
I know how to use the internet to answer my questions about health	14 (3.7)	30 (7.8)	87 (22.7)	191 (49.9)	61 (15.9)	3.67
I have the skills I need to evaluate the health resources I find on the internet	12 (3.1)	49 (12.8)	91 (23.8)	165 (43.1)	66(17.2)	3.58
I can tell high-quality health resources from low-quality resources	17 (4.4)	36 (9.4)	109 (28.5)	157 (41.0)	64 (16.7)	3.56
I feel confident in using information from the internet to make health decisions	7 (1.8)	34 (8.9)	64 (16.7)	187 (48.8)	91 (23.8)	3.84

Table 3. eHealth literacy level by background characteristics.

Variables	Low eHealth literacy, n (%)	High eHealth literacy, n (%)
Sex		
Male (n=239)	99 (41.4)	140 (58.6)
Female (n=144)	59 (41.0)	85 (59.0)
Age group (years)		
20-24 (n=18)	8(44.4)	10 (55.6)
25-29 (n=254)	106 (41.7)	148 (58.3)
≥30 (n=111)	44 (39.6)	67 (60.4)
Professional background		
Nurse (n=158)	75(47.5)	83(52.5)
Medical doctor (n=94)	31 (33.0)	63 (67.0)
Pharmacist (n=30)	11 (36.7)	19 (63.3)
Midwife (n=54)	23 (41.8)	32 (58.2)
Laboratory (n=24)	11 (45.8)	13 (54.2)
Other (n=23)	7 (31.8)	15 (68.2)
Work experience (years)		
1-3 (n=226)	94 (41.6)	132(58.4)
4-6 (n=119)	48 (40.3)	71 (59.7)
≥7 (n=38)	16 (42.1)	22 (57.9)

Computer Knowledge, Information Retrieval Training Need, and Perception Toward eHealth Information Resources

The majority of the study participants (314/383, 82.0%) stated that they require training in retrieving information from eHealth information sources. Approximately one-third of the participants (131/383, 34.2%) had poor computer knowledge. The majority of the participants (268/383, 70.0%) agreed that eHealth information resources are a useful tool in supporting them in making health-related decisions, and 57.7% (221/383) of the participants also perceived that retrieving information from these sources is easy.

Factors Associated With eHealth Literacy

Bivariable and multivariable logistic regression analyses were used to discover potential predictor variables linked with eHealth literacy. Variables with $P<.20$ in the bivariable analysis were further considered in the multivariable logistic regression analysis. Finally, the multivariable logistic regression analysis revealed that the variables perceived usefulness, computer access, perceived ease of use of eHealth information resources,

and computer knowledge were significantly associated with health professionals' eHealth literacy; thus, all three hypotheses were supported.

Perceived usefulness was significantly associated with eHealth literacy. Health professionals who perceived using health information resources located on the internet as useful were approximately 2-times more eHealth literate than their counterparts who perceived using eHealth information resources as not useful. Respondents who had computer access were also more than 2-times more likely to be eHealth literate than those who did not have computer access (Table 4).

The computer knowledge of health professionals was also significantly associated with eHealth literacy. Study participants who had good computer knowledge were 2.3-times more eHealth literate than those who had poor computer knowledge (Table 4).

Table 4. Multivariable logistic regression analysis of factors associated with eHealth literacy.

Variables	Adjusted odds ratio (95% CI)	*P* value
Sex		
Male (reference)	1 (1-1)	—[a]
Female	1.17 (0.718-2.479)	.52
Age group (years)		
20-24 (reference)	1 (1-1)	—
25-29	0.774 (0.242-2.479)	.68
≥30	0.63 (0.162-2.467)	.51
Computer access		
No (reference)	1 (1-1)	—
Yes	2.32 (1.389-3.861)	<.001
Work experience (years)		
1-3 (reference)	1 (1-1)	—
4-6	0.99 (0.529-1.85)	.97
≥7	0.97 (0.336-2.79)	.95
Computer knowledge		
Poor (reference)	1 (1-1)	—
Good	2.34 (1.442-3.787)	.001
Perceived usefulness		
Not useful (reference)	1 (1-1)	—
Useful	1.82 (1.075-3.091)	.03
Perceived ease of use		
Not easy (reference)	1 (1-1)	—
Easy	4.53 (2.768-7.401)	<.001

[a]Not applicable.

Discussion

Principal Findings

Higher levels of eHealth literacy may help people make better health-related decisions, resulting in better health outcomes [28]. This study found that the higher the eHealth literacy, the more it will promote social media use for health information, health information–seeking behaviors, and self-care agency. This will lead to better health-promoting behavior by raising the motivation and intention for health promotion [29]. The main purpose of this study was to estimate the level of eHealth literacy and identify its potential predictors among health care providers working in Amhara regional state teaching hospitals in Ethiopia during the COVID-19 pandemic. Our results showed that health professionals in Amhara regional state teaching hospitals have a moderate level of eHealth literacy (mean 29.21). The findings also revealed that computer access, perceived usefulness of eHealth information resources, perceived ease of use of eHealth information resources, and computer knowledge were significantly associated with health professionals' eHealth literacy level.

Comparison With Prior Work

In a study on Iranian medical and health science university students, the mean eHEALS score was 28.21 [30]. However, health professionals in Germany have higher eHealth literacy levels [31]. The possible reason for this discrepancy might be

due to the variation of internet penetration between these countries, which is 19% in Ethiopia [32] and 93% in Germany [33]. Additionally, due to the limited availability of health-related information in languages other than English, geographical location, cultural, and language barriers may have an impact on eHEALS scores [34,35].

The majority of participants agreed that the internet assisted them in making health-related decisions, and while they believe they know where to find helpful health resources on the internet and how to use them, nearly two-fifth of participants were not confident in their ability to evaluate the information they have retrieved. In particular, the participants' inability to distinguish between high- and low-quality health resources on the internet suggests a potential weakness in their ability to recognize crucial characteristics that would aid in determining which website may be reliable. Despite the increasing availability of eHealth information and increased acceptance of this mode of communication, all populations, including health professionals, may lack the skills to keep up with this dynamic and changing medium [36].

In terms of professional background, we found that nurses and medical doctors had greater eHealth literacy levels than midwives, pharmacists, and other health professionals. This finding is consistent with a previous study conducted among health care workers in Vietnam [16]. This might be due to the fact that doctors have more professional training than other health professionals, and they have been identified as the group with the greatest capacity for finding, analyzing, justifying, and using health care–related information [37]. Furthermore, because nurses and doctors are the primary caregivers responsible for educating and directing patients, health literacy has been identified as a strategic approach for improving patient–health care worker communication [38].

This study also investigated the factors that can influence eHealth literacy. Computer access, perceived usefulness of eHealth information resources, perceived ease of use, and computer knowledge were identified as significant predictors. Participants who had computer access in their working area were more likely to be eHealth literate than those who did not. Low access to technologies could be the main reason behind the low access to eHealth services [39].

The participants' eHealth literacy score was associated with their computer knowledge. This finding is consistent with studies that showed a positive association between technology literacy and eHealth literacy [26,40]. A Bangladeshi study also revealed a significant association between eHealth literacy and computer knowledge among university students [25]. Participants who had computer expertise had higher eHealth literacy levels than those who had poor computer knowledge. Because of Ethiopia's status as a developing country, computer access is limited, literacy skills are insufficient, and health professionals consequently do not have equal access to eHealth resources.

In contrast to other studies [16], demographic variables such as the participants' sex and age were not found to be significantly associated with eHealth literacy in this study. In a study conducted in Jordan, a nonsignificant association was found between sex, age, and eHealth literacy [22]. Similarly, no significant association was found between gender and eHealth literacy in an Italian study, whereas there was a significant association between age and eHealth literacy [23]. Further studies are needed to investigate the associations between gender, age, and eHealth literacy.

eHealth literacy was also significantly associated with the perceived usefulness of eHealth information sources. Participants who perceived eHealth information resources to be useful for making decisions were more eHealth literate than those who perceived these resources as not being as useful. These findings are supported by studies conducted among nursing students in Jordan [22] and Nepal [24].

It has been suggested that eHealth should be integrated into the health care system, as it can provide certain benefits for improving the quality of health care received [41]. Health professionals should be informed on the latest information and skills to acquire competency in using eHealth resources for patient care and clinical decision-making. According to a preliminary situation assessment, eHealth initiatives in Ethiopia are characterized as being of small scale and unable to effectively communicate with each other (ie, low interoperability). Accordingly, the Ethiopian government developed and formulated a national eHealth strategy for coordinating and streamlining the eHealth initiatives underway in the country as well as for establishing a foundation for sustainable eHealth implementation [42]. When promoting eHealth literacy among health professionals, perceived usefulness and ease of use of eHealth information sources, along with training on information retrieval, computer knowledge, and access should be considered.

Conclusions

This study provides an overview of health professionals' eHealth literacy levels in the Amhara regional state, Ethiopia, revealing that more than half of these professionals have a high degree of eHealth literacy. Additionally, the factors associated with eHealth literacy were explored, with the results suggesting significant associations of perceived usefulness, perceived ease of use, computer access, and computer knowledge. To improve health professionals' eHealth literacy, which could help them assist in decision-making, multidisciplinary approaches are needed. This would help to minimize the risk of infectious diseases such as COVID-19 and further mitigate its impacts. Health professionals also require eHealth literacy to assist their patients in obtaining more up-to-date, reliable, and high-quality information. It is crucial to provide training and support to health care workers on how to find, interpret, and, most importantly, evaluate the quality of health information found on the internet to improve their eHealth literacy level. Further research is needed to explore the role of eHealth literacy for mitigating pandemics in developing countries.

Acknowledgments

We would like to express our gratitude to the study participants for their willingness to participate. We also express gratitude to the data collectors and supervisors for their time and effort.

Authors' Contributions

All authors made a substantial contribution to the study's conception, design, methodology, and data analysis. GAT drafted the manuscript. All authors read the manuscript, revised it critically for important intellectual content, and approved the final version.

Conflicts of Interest

None declared.

References

1. Norman C. eHealth literacy 2.0: problems and opportunities with an evolving concept. J Med Internet Res 2011 Dec 23;13(4):e125 [FREE Full text] [doi: 10.2196/jmir.2035] [Medline: 22193243]
2. Lagoe C, Atkin D. Health anxiety in the digital age: An exploration of psychological determinants of online health information seeking. Comput Human Behav 2015 Nov;52:484-491. [doi: 10.1016/j.chb.2015.06.003]
3. Powell JA, Darvell M, Gray JAM. The doctor, the patient and the world-wide web: how the internet is changing healthcare. J R Soc Med 2017 Nov 23;96(2):74-76. [doi: 10.1177/014107680309600206]
4. National Research Council (US) Committee on Enhancing the Internet for Health Applications: Technical Requirements Implementation Strategies. Networking health: prescriptions for the internet. Washington, DC: National Academies Press; 2000.
5. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
6. Diviani N, van den Putte B, Giani S, van Weert JC. Low health literacy and evaluation of online health information: a systematic review of the literature. J Med Internet Res 2015 May 07;17(5):e112 [FREE Full text] [doi: 10.2196/jmir.4018] [Medline: 25953147]
7. Bassetti M, Vena A, Giacobbe DR. The novel Chinese coronavirus (2019-nCoV) infections: challenges for fighting the storm. Eur J Clin Invest 2020 Mar 05;50(3):e13209 [FREE Full text] [doi: 10.1111/eci.13209] [Medline: 32003000]
8. Phelan AL, Katz R, Gostin LO. The novel coronavirus originating in Wuhan, China: challenges for global health governance. JAMA 2020 Feb 25;323(8):709-710 [FREE Full text] [doi: 10.1001/jama.2020.1097] [Medline: 31999307]
9. Chong YY, Cheng HY, Chan HYL, Chien WT, Wong SYS. COVID-19 pandemic, infodemic and the role of eHealth literacy. Int J Nurs Stud 2020 Aug;108:103644 [FREE Full text] [doi: 10.1016/j.ijnurstu.2020.103644] [Medline: 32447127]
10. Kim S, Son YJ. Relationships between eHealth literacy and health behaviors in Korean adults. Comput Inform Nurs 2017 Feb;35(2):84-90. [doi: 10.1097/CIN.0000000000000255] [Medline: 27258808]
11. Negro-Calduch E, Azzopardi-Muscat N, Nitzan D, Pebody R, Jorgensen P, Novillo-Ortiz D. Health information systems in the COVID-19 pandemic: a short survey of experiences and lessons learned from the European region. Front Public Health 2021 Sep 28;9:676838. [doi: 10.3389/fpubh.2021.676838] [Medline: 34650946]
12. Schmidt AE, Abboud LA, Bogaert P. Making the case for strong health information systems during a pandemic and beyond. Arch Public Health 2021 Jan 29;79(1):13 [FREE Full text] [doi: 10.1186/s13690-021-00531-5] [Medline: 33514433]
13. Gong K, Xu Z, Cai Z, Chen Y, Wang Z. Internet hospitals help prevent and control the epidemic of COVID-19 in China: multicenter user profiling study. J Med Internet Res 2020 Apr 14;22(4):e18908 [FREE Full text] [doi: 10.2196/18908] [Medline: 32250962]
14. Paakkari L, Okan O. COVID-19: health literacy is an underestimated problem. Lancet Public Health 2020 May;5(5):e249-e250 [FREE Full text] [doi: 10.1016/S2468-2667(20)30086-4] [Medline: 32302535]
15. Hwang AR, Kang H. Influence of eHealth literacy on health promoting behaviors among university students. J Korean Soc School Health 2019;32(3):165-174. [doi: 10.15434/kssh.2019.32.3.165]
16. Do BN, Tran TV, Phan DT, Nguyen HC, Nguyen TTP, Nguyen HC, et al. Health literacy, eHealth literacy, adherence to infection prevention and control procedures, lifestyle changes, and suspected COVID-19 symptoms among health care workers during lockdown: online survey. J Med Internet Res 2020 Nov 12;22(11):e22894 [FREE Full text] [doi: 10.2196/22894] [Medline: 33122164]
17. Bastick Z. Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation. Comput Human Behav 2021 Mar;116:106633. [doi: 10.1016/j.chb.2020.106633]
18. Barua Z, Barua S, Aktar S, Kabir N, Li M. Effects of misinformation on COVID-19 individual responses and recommendations for resilience of disastrous consequences of misinformation. Prog Disaster Sci 2020 Dec;8:100119 [FREE Full text] [doi: 10.1016/j.pdisas.2020.100119] [Medline: 34173443]
19. Metzger MJ, Flanagin AJ. Using Web 2.0 technologies to enhance evidence-based medical information. J Health Commun 2011 Jul 29;16(Suppl 1):45-58. [doi: 10.1080/10810730.2011.589881] [Medline: 21843095]

20. Shiferaw KB, Tilahun BC, Endehabtu BF, Gullslett MK, Mengiste SA. E-health literacy and associated factors among chronic patients in a low-income country: a cross-sectional survey. BMC Med Inform Decis Mak 2020 Aug 06;20(1):181 [FREE Full text] [doi: 10.1186/s12911-020-01202-1] [Medline: 32762745]
21. Stellefson M, Hanik B, Chaney B, Chaney D, Tennant B, Chavarria EA. eHealth literacy among college students: a systematic review with implications for eHealth education. J Med Internet Res 2011 Dec 01;13(4):e102 [FREE Full text] [doi: 10.2196/jmir.1703] [Medline: 22155629]
22. Tubaishat A, Habiballah L. eHealth literacy among undergraduate nursing students. Nurse Educ Today 2016 Jul;42:47-52. [doi: 10.1016/j.nedt.2016.04.003] [Medline: 27237352]
23. Del Giudice P, Bravo G, Poletto M, De Odorico A, Conte A, Brunelli L, et al. Correlation between eHealth literacy and health literacy using the eHealth Literacy Scale and real-life experiences in the health sector as a proxy measure of functional health literacy: cross-sectional web-based survey. J Med Internet Res 2018 Oct 31;20(10):e281. [doi: 10.2196/jmir.9401] [Medline: 30381283]
24. Sharma S, Oli N, Thapa B. Electronic health-literacy skills among nursing students. Adv Med Educ Pract 2019;10:527-532. [doi: 10.2147/AMEP.S207353] [Medline: 31410077]
25. Islam MM, Touray M, Yang HC, Poly TN, Nguyen PA, Li YCJ, et al. E-Health literacy and health information seeking behavior among university students in Bangladesh. Stud Health Technol Inform 2017;245:122-125. [Medline: 29295065]
26. Xesfingi S, Vozikis A. eHealth literacy: in the quest of the contributing factors. Interact J Med Res 2016 May 25;5(2):e16 [FREE Full text] [doi: 10.2196/ijmr.4749] [Medline: 27226146]
27. Geography of Ethiopia. Wikipedia. URL: https://en.wikipedia.org/wiki/Geography_of_Ethiopia [accessed 2022-06-28]
28. Werts N, Hutton-Rogers L. Barriers To achieving e-Health literacy. Am J Health Sci 2013 Aug 14;4(3):115-120. [doi: 10.19030/ajhs.v4i3.8007]
29. Kim S, Oh J. The relationship between e-health literacy and health-promoting behaviors in nursing students: a multiple mediation model. Int J Environ Res Public Health 2021 May 28;18(11):5804 [FREE Full text] [doi: 10.3390/ijerph18115804] [Medline: 34071469]
30. Dashti S, Peyman N, Tajfard M, Esmaeeli H. E-Health literacy of medical and health sciences university students in Mashhad, Iran in 2016: a pilot study. Electron Physician 2017 Mar 25;9(3):3966-3973 [FREE Full text] [doi: 10.19082/3966] [Medline: 28461871]
31. Hennemann S, Beutel ME, Zwerenz R. Ready for eHealth? Health professionals' acceptance and adoption of eHealth interventions in inpatient routine care. J Health Commun 2017 Mar 19;22(3):274-284. [doi: 10.1080/10810730.2017.1284286] [Medline: 28248626]
32. Kemp S. Digital 2020 Ethiopia. DataReportal. 2020 Feb 17. URL: https://datareportal.com/reports/digital-2020-ethiopia [accessed 2022-06-28]
33. Kemp S. Digital 2020: Germany. DataReportal. 2020 Feb 12. URL: https://datareportal.com/reports/digital-2020-germany [accessed 2022-06-28]
34. Mitsutake S, Shibata A, Ishii K, Oka K. Associations of eHealth literacy with health behavior among adult internet users. J Med Internet Res 2016 Jul 18;18(7):e192 [FREE Full text] [doi: 10.2196/jmir.5413] [Medline: 27432783]
35. Tariq A, Khan SR, Basharat A. Internet use, eHealth literacy, and dietary supplement use among young adults in Pakistan: cross-sectional study. J Med Internet Res 2020 Jun 10;22(6):e17014 [FREE Full text] [doi: 10.2196/17014] [Medline: 32519974]
36. van Deursen AJAM, van Dijk JAGM. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011 Apr 29;13(2):e35 [FREE Full text] [doi: 10.2196/jmir.1581] [Medline: 21531690]
37. Palesy D, Jakimowicz S. Health literacy training for Australian home care workers: Enablers and barriers. Home Health Care Serv Q 2019 Apr 20;38(2):80-95. [doi: 10.1080/01621424.2019.1604458] [Medline: 31007141]
38. Tavakoly Sany SB, Behzhad F, Ferns G, Peyman N. Communication skills training for physicians improves health literacy and medical outcomes among patients with hypertension: a randomized controlled trial. BMC Health Serv Res 2020 Jan 23;20(1):60 [FREE Full text] [doi: 10.1186/s12913-020-4901-8] [Medline: 31973765]
39. Newman L, Patel K, Falls Prevention Project. The Role and Impact of Digital and Traditional Information and Communication Pathways in Health Service Access and Equity. Australian Policy Online. 2012 Sep. URL: https://apo.org.au/sites/default/files/resource-files/2012-08/apo-nid64561.pdf [accessed 2022-06-28]
40. Amante DJ, Hogan TP, Pagoto SL, English TM, Lapane KL. Access to care and use of the internet to search for health information: results from the US National Health Interview Survey. J Med Internet Res 2015 Apr 29;17(4):e106 [FREE Full text] [doi: 10.2196/jmir.4126] [Medline: 25925943]
41. Vanagas G, Engelbrecht R, Damaševičius R, Suomi R, Solanas A. eHealth solutions for the integrated healthcare. J Healthc Eng 2018 Jul 10;2018:3846892. [doi: 10.1155/2018/3846892] [Medline: 30123441]
42. Federal Ministry of Health Ethiopia. Ethiopian National eHealth Strategic - World Health Organization. SlideLegend. URL: https://slidelegend.com/ethiopian-national-ehealth-strategic-world-health-organization_59ed930f1723ddffdd4ad7a2.html [accessed 2022-06-28]

Abbreviations

eHEALS: eHealth Literacy Scale
UOG: University of Gondar

Edited by A Mavragani; submitted 05.01.22; peer-reviewed by A Garcia, H Liu; comments to author 15.05.22; revised version received 26.05.22; accepted 08.06.22; published 13.07.22.

Please cite as:
Tesfa GA, Yehualashet DE, Ewune HA, Zemeskel AG, Kalayou MH, Seboka BT
eHealth Literacy and its Associated Factors Among Health Professionals During the COVID-19 Pandemic in Resource-Limited Settings: Cross-sectional Study
JMIR Form Res 2022;6(7):e36206
URL: https://formative.jmir.org/2022/7/e36206
doi:10.2196/36206
PMID:35737897

Original Paper

The Impact of SARS-CoV-2 (COVID-19) on the Acuity of Mental Health–Related Diagnosis at Admission for Young Adults in New York City and Washington, DC: Observational Study

Amanda Fialk[1], PhD; Alexa Connors[1], MS; Brianna Cerrito[1], BA; Karee Jones[1], BA; Frank Buono[2], PhD

[1]The Dorm, New York City, NY, United States

[2]Department of Psychiatry, Yale School of Medicine, Yale University, New Haven, CT, United States

Corresponding Author:
Frank Buono, PhD
Department of Psychiatry
Yale School of Medicine
Yale University
300 George Street
New Haven, CT, 06510
United States
Phone: 1 8478496007
Email: frank.buono@yale.edu

Abstract

Background: The COVID-19 pandemic has required restrictive measures to mitigate transmission of the virus. Evidence has demonstrated increased generalized anxiety and depression among young adults due to the COVID-19 pandemic. However, minimal research has examined the longitudinal effect of COVID-19 over the course of time and its impact on anxiety and depression. Additionally, age and gender have been found to play a significant role on individuals' mental health, with young adults and women particularly at risk.

Objective: The aim of this study was to examine the impact of the COVID-19 pandemic on anxiety and depression upon admissions to treatment.

Methods: This was an observational study that was completed longitudinally in which the grouping variable split the time interval into five equal groups for assessments over each period of time. A total of 112 young adults (aged 18-25 years) were recruited for the study. Participants completed assessments online through a Qualtrics link.

Results: Psychometric properties of the admission assessments were uniformly highly statistically significant. There was a significant difference in generalized anxiety between the group-1 and group-3 time intervals. No significant difference was found across the time intervals for depression. Differences in predicting the impact of the psychometrics scores were found with respect to gender. Only the ability to participate and the quality-of-life subfactor of the Functional Assessment of Chronic Illness Therapy (FACIT) assessment were significant.

Conclusions: This study sought to understand the impact that COVID-19 has had on young adults seeking mental health services during the pandemic. Gender emerged as a clear significant factor contributing to increased anxiety in young adults seeking mental health services during the pandemic. These findings have critical importance to ensuring the potential treatment success rate of clients, while providing an overarching understanding of the impact of the pandemic and establishing clinical recommendations for the treatment of individuals who are seeking out treatment.

(JMIR Form Res 2022;6(7):e39217) doi:10.2196/39217

KEYWORDS

COVID-19; young adults; mental health treatment; anxiety; depression; youth; mental health; Qualtrics; psychometric properties; gender; mental health service; quality of life

Introduction

On March 11, 2020, the World Health Organization declared the coronavirus SARS-CoV-2 (COVID-19) outbreak as a pandemic. To mitigate transmission, restrictive measures were imposed across the United States, including the closure and modification of businesses and mandates for the use of face masks and vaccinations, which have had an immediate and unprecedented impact on psychological health [1-9]. Incidences of depression and anxiety, among other struggles such as substance use, posttraumatic stress disorder (PTSD), eating disorders, and suicidality, have increased as a result of pandemic-related stressors such as lockdowns, exposure to infected people, loss of loved ones, and economic hardship [10,11]. Age and gender have been found to play significant predictive roles on the impact that COVID-19 stressors have had on individuals' mental health, with young adults and women identified to be particularly at risk [5,6].

Several studies have demonstrated an increased incidence of depression and anxiety during the COVID-19 pandemic among adults [6,12-17]. One US study found that in states that had more than 50 COVID-19 cases since March 10, 2020, each additional day resulted in an 11% increase in the probability of ascending into a higher incidence of mental health crises [18]. Quantitative studies conducted in other countries echo these findings. Ammerman et al [19] investigated the association between COVID-19 and the prevalence of suicidal thoughts and behaviors, showing that 45% of those who endorsed past-month suicidal ideation reported that their thoughts were directly related to COVID-19. Moreover, 9% of the sample reported intentionally exposing themselves to COVID-19, with 50% of these individuals intending to kill themselves through willfully contracting the COVID-19 virus [19]. Taken together, these findings suggest a significant predictive relationship between COVID-19 stressors and the prevalence and severity of individuals' mental health, including anxiety and depression. However, there is a lack of research examining the impact of COVID-19 over the prolonged course of the pandemic.

Predictive factors related to the impact of COVID-19 stressors have also begun to emerge in the current literature. Age has been found to play a significant predictive role in the impact of COVID-19 on mental health. Evidence from several recent studies suggests that young adults are particularly at risk for experiencing mental health problems during the pandemic [12,14,20-24]. Studies have reported high levels of depression and anxiety symptoms in college students and young adults as a result of COVID-19 pandemic–related stressors [3,5,6,14,16,23,25]. More specifically, researchers evaluating the impacts of the COVID-19 quarantine in France among 69,054 university students found a high prevalence of suicidal ideation (11.4%), severe distress (22.4%), heightened perceived stress (24.7%), severe depression (16.1%), and high levels of anxiety (27.5%) [23]. Researchers also examined the prevalence of PTSD, depression, and psychological risk factors in 2485 home-quarantined college students, finding that PTSD had a prevalence of 2.7% and depression had a prevalence of 9.0% [25]. García-Portilla et al [26] examined the psychological effects of the pandemic across age groups and found that those under 60 years old were at greatest risk. One potential related factor contributing to this finding is lack of resilience, as younger individuals have less years of experience in coping with hardship, which is a developmentally acquired skill [26,27]. These findings suggest that age has the potential to be a strong predictive factor of interest in the relative impact of COVID-19 stressors on key facets of mental health.

Gender has also been found to play a significant predictive role in the relative impact of COVID-19 on mental health. Turna et al [28] examined the effects of the COVID-19 pandemic on multiple facets of mental well-being. Anxiety and depression were assessed using the General Anxiety Disorder (GAD-7), Patient Health Questionnaire (PHQ-9), and Perceived Stress Scale among 632 participants, 82% of whom identified as female. The results showed that nearly one-third (31%) of participants met the criteria for generalized anxiety disorder and 29% met the criteria for major depressive disorder. Female gender was significantly predictive of psychiatric symptoms at a 99% CI. These findings are also supported by evidence found in a study by Emery et al [29] that looked at how COVID-19 impacted the mental health and behaviors of young adults in the United States: female gender, prepandemic mental illness symptoms, and self-reported stress were found to be emergent risk factors for mental illness. Kecojevic et al [30] obtained similar findings when examining the impact of COVID-19 on the mental health of college students, in which female students reported higher anxiety than male students. Collectively, these results suggest that female gender is a risk factor for COVID-19–related anxiety [28-30].

Emerging adulthood is a time of uncertainty and exploration, and the COVID-19 pandemic has had a compounding effect on this process, resulting in a significant decline in life satisfaction and a significant increase in mental health problems [31]. The aim of this study was therefore to examine the longitudinal impact of the COVID-19 pandemic on anxiety and depression upon admissions to treatment at a long-term and intensive outpatient mental health program for young adults (The Dorm). Specifically, this study examined the prevalence and acuity levels of generalized anxiety and depression diagnoses upon admissions in the first 15 months of the COVID-19 epidemic. Particular attention was paid to how gender identity was associated with acuity levels of anxiety and/or depression. We anticipated that as the pandemic continued over time, so would depression and generalized anxiety among young adults.

Methods

Participants

We attempted to recruit 133 individuals for this study, 112 of whom ultimately volunteered. The 21 individuals who did not complete the study cited lack of interest for not participating. Among the 112 young adults that agreed to participate (mean age 22.3, SD 3.2 years), 53 (47.3%) self-identified as female. Inclusion criteria were: (1) admitted to The Dorm, (2) provided informed consent, and (3) at least 18 years of age. Exclusion criteria were: (1) unfit to complete the survey due to medical

or psychological constraint and (2) not fluent in the English language.

Setting

The Dorm is an intensive outpatient program for young adults aged 18-35 years operating in New York City and Washington, DC. Empirically supported behavioral psychosocial methodologies are implemented to serve a variety of mental health illnesses and co-occurring disorders over the course of a phased treatment, which is typically a 1-year admittance (on average). In addition, holistic approaches such as exercise, yoga, Reiki, horticulture, community service, meditation, and mindfulness are an integral part of the treatment model. All clients participate in family programming, including weekly parent coaching, parent groups, and family groups. Clients work with both a therapist and a clinical coach, and participate in 3-30 hours a week of group therapy depending on the treatment phase.

Assessments

Generalized Anxiety Disorder

The GAD-7 is a 7-item, self-rated scale developed in correspondence with the Diagnostic and Statistical Manual and updated for the 5th Edition, which is used as a screening tool and severity indicator for generalized anxiety disorder. Reliability of this assessment demonstrated a Cronbach α of .92 [32].

Patient Health Questionnaire

The PHQ-9 is a 9-item self-report scale that is used for screening, diagnosing, monitoring, and measuring the severity of depression (mild, scores of 5-9; moderate, scores of 10-14; moderately severe, scores of 15-19; severe, scores of 20-21) [33]. Reliability of this assessment demonstrated a Cronbach α of .89.

Barratt Impulsivity Scale

The Barratt Impulsivity Scale (BIS-11) is a 4-point Likert scale from “rarely/never” (score=1) to “almost always/always” (score=4), where 4 indicates the most impulsive response [34,35]. However, there is no standard BIS-11 score that is consistently used to designate an individual as “highly impulsive.”

Functional Assessment of Chronic Illness Therapy Measurement System

The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System is a 27-item questionnaire that focuses on the domains of physical well-being, social/family well-being, emotional well-being, and functional well-being. Items are scored on a 5-point Likert scale, where participants endorse “not at all” to “very much” as the statement applies over the past 7 days [36].

Patient-Reported Outcomes Measurement Information System Scales

Ability to Participate in Social Roles and Activities

The Patient-Reported Outcomes Measurement Information System (PROMIS) Ability to Participate in Social Roles and Activities is an 8-item self-report questionnaire that assesses the perceived ability to perform one’s usual social roles and activities. Items are worded negatively in terms of perceived limitations and responses are reverse-coded, where 5 indicates “never” and 1 indicates “always”; thus, higher scores represent fewer limitations (better abilities). The item bank does not use a time frame (eg, over the past 7 days) when assessing the ability to participate in social roles and activities.

General Self-Efficacy

PROMIS General Self-Efficacy is a 10-item measure used to assess a person’s belief in their capacity to manage daily stressors and have control over meaningful events [33]. This measure was derived from the National Institutes of Health Toolbox Self-Efficacy Item Bank by creating new “confidence” response options that mirrored the same response options as in the PROMIS measures of the Self-Efficacy for Managing Chronic Conditions scale: “I am not at all confident,” “I am a little confident,” “I am somewhat confident,” “I am quite confident,” “I am very confident” [37].

Friendship (Ages 18+) Fixed Form

PROMIS Friendship (Ages 18+) Fixed Form is an 8-item scale evaluating the perceptions of the respondent on their availability of friends or companions with whom to interact or affiliate with in the past month on a 5-point Likert scale (1=never, 5=always).

Perceived Rejection (Ages 18+)

PROMIS Perceived Rejection (Ages 18+) is an 8-item scale evaluating how often people perceive others to be arguing/yelling at them and the perceived insensitivity over the past month on a 5-point Likert scale (1=never, 5=always).

Sleep-Related Impairment Questionnaire

The PROMIS Sleep-Related Impairment Questionnaire is an 8-item questionnaire that measures self-reported alertness, sleepiness, tiredness, and functional impairments associated with sleep problems during waking hours within the past 7 days [38]. This measure uses a 5-point Likert scale (1=not at all, 5=very much).

Procedure

All potential participants were consented either through a private meeting using Zoom virtual conference software or in person under standardized COVID-19 protocols. Consent forms were read orally, and any questions were answered by the investigator or research assistant before asking the participant to electronically sign their name on the tablet. By electronically signing their name, the participant was assigned a random 6-digit number generated by Qualtrics survey management tools. All consented individuals were tracked and followed through a password-protected Microsoft Access file. Each participant was asked to complete the assessments online through the Qualtrics link. The research assistant stayed with the participants and provided assistance as needed until the completion of the survey. The total length of time required to complete the survey did not exceed 40 minutes.

Ethical Considerations

The study was approved by Yale School of Medicine's institutional review board prior to participant recruitment (IRES number 2000026514). This research was performed in alignment with the World Medical Association Declaration of Helsinki (2018).

Data Analysis

Data were analyzed using SPSS for Windows, version 26.0. The grouping variable split the time interval into five equal phases (1=March 2020 through May 2020; 2=June 2020 through August 2020; 3=September 2020 through November 2020; 4=December 2020 through February 2021; 5=March 2021 through May 2021).

Mean group and overall differences were analyzed using one-way analysis of variance, whereas dichotomous variables (eg, age, gender, location) were analyzed using Fisher-exact or χ^2 tests. Strengths of association between variables were examined by the Pearson correlation coefficient and by linear regression analyses with generation of odds ratios and accompanying 95% CIs. All statistical tests were two-tailed with an α level of .05. Additionally, linear regression was used with gender (male, female, transgender, and other) as the dependent variable to evaluate its impact on psychometric assessments.

Results

The demographic characteristics of the participants are summarized in Table 1. Over 98% of participants were single at admission and the majority either had some college or a high school/equivalent education when combined. The majority of participants were white, with equal representation of self-identified males and females at admission, along with individuals identifying as transgender and other represented. The highest percentage of admissions were unemployed 45%, and nearly 60% of those admitted had experienced at least a singular traumatic event in their lifetime.

Table 2 presents a correlation matrix of the psychometric properties of the admission assessments, which were uniformly highly statistically significant. Figure 1 presents the participants' average acuity of primary diagnosis of generalized anxiety and/or depression across the time intervals, along with the range of scores within each grouping. There was a significant difference in generalized anxiety levels between the group-1 and group-3 time intervals. No significant difference was noted across the time intervals for the depression assessment.

Additionally, the findings from the linear regression demonstrated differences across gender and the admission psychometrics. Significant findings were noted (F_{101}=2.025, P=.04) to predict the impact of the psychometrics scores on gender. However, as noted in the coefficients subtable, only ability to participate (P=.002) and the quality-of-life subfactor of the FACIT (P=.02) assessments were significant with other values being not significant. In contrast to initial indications, depression and anxiety were not significant factors across gender at admission to this program.

Table 1. Demographics of admission intakes (N=112).

Characteristic	Value	*P* value
Age (years), mean (SD)	22.2 (3.2)	N/A[a]
Gender, n (%)		.44
Male	53 (47.3)	
Female	53 (47.3)	
Transgender	1 (0.8)	
Other	5 (4.5)	
Race, n (%)		.11
White	84 (75.0)	
Black	9 (8.0)	
Hispanic	5 (4.5)	
Asian	14 (12.5)	
Marital status, n (%)		.10
Single	110 (98.2)	
Married	2 (1.8)	
Highest education, n (%)		.74
Some high school	4 (3.6)	
High school graduation/GED[b]	44 (39.3)	
Some college	46 (41.1)	
Associates degree	2 (1.8)	
Bachelor's degree	18 (16.1)	
Employment, n (%)		.47
Full time work	14 (12.5)	
Student	47 (42.0)	
Unemployed	51 (45.5)	
Experienced trauma in life, n (%)		.46
Yes	67	
No	45	

[a]N/A: not applicable.

[b]GED: General Educational Development.

Table 2. Correlational table of psychometric measures at admission.

Variable	GAD-7[a] total	PHQ-9[b]	FACIT[c] physical	FACIT well-being	Ability to participate	Impulsivity	Friendship	Rejection	Self-efficacy	Sleep
GAD-7 total										
r	1	0.740	0.643	–0.421	–0.594	0.077	–0.131	0.381	–0.306	–0.110
P value	—[d]	<.001	<.001	<.001	<.001	.42	.17	<.001	<.001	.25
PHQ-9										
r	0.740	1	0.749	–0.579	0.650	0.036	–0.290	0.395	–0.395	–0.080
P value	<.001	—	<.001	<.001	<.001	.71	.002	<.001	<.001	.40
FACIT physical										
r	0.643	0.749	1	–0.398	0.575	0.129	–0.084	0.356	–0.251	–0.080
P value	<.001	<.001	—	<.001	<.001	.17	.38	<.001	.007	.40
FACIT well-being										
r	–0.421	–0.579	–0.398	1	–0.523	0.525	0.551	–0.196	0.618	0.237
P value	<.001	<.001	<.001	—	<.001	<.001	<.001	.04	<.001	.01
Ability to participate										
r	0.594	0.650	0.575	–0.523	1	–0.029	–0.320	0.417	–0.356	–0.109
P value	<.001	<.001	<.001	<.001	—	.76	<.001	<.001	<.001	.25
Impulsivity										
r	0.077	0.036	0.129	0.525	–0.029	1	0.345	0.189	0.454	0.149
P value	.42	.71	.17	<.001	.76	—	<.001	.04	<.001	.11
Friendship										
r	–0.131	–0.290	–0.084	0.551	–0.320	0.345	1	–0.228	0.544	0.193
P value	.165	.002	.38	<.001	<.001	<.001	—	.02	<.001	.04
Rejection										
r	0.381	0.395	0.356	–0.196	0.417	0.189	–0.228	1	–0.121	0.000
P value	<.001	<.001	<.001	.04	<.001	.04	.02	—	.20	>.99
Self-efficacy										
r	–0.306	–0.395	–0.251	0.618	–0.356	0.454	0.544	–0.121	1	0.316
P value	<.001	<.001	.007	<.001	<.001	<.001	<.001	.20	—	<.001
Sleep										
r	–0.110	–0.080	–0.080	0.237	–0.109	0.149	0.193	0.000	0.316	1
P value	.25	.40	.40	.01	.25	.11	.04	>.99	<.001	—

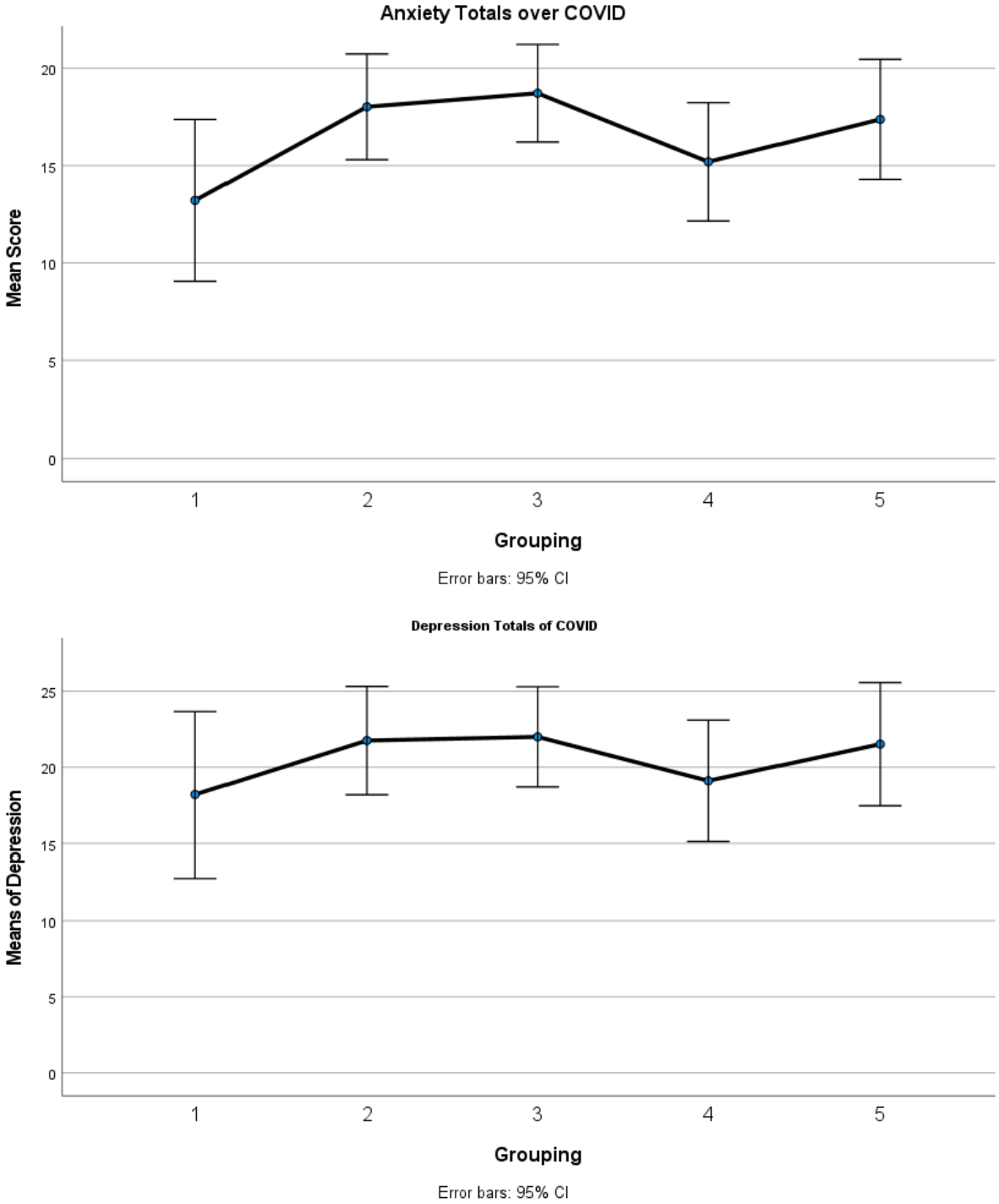
[a]GAD-7: 7-item General Anxiety Disorder scale.

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

[c]FACIT: Functional Assessment of Chronic Illness Therapy.

[d]Not applicable.

Figure 1. Aggregate differences of anxiety and depression across the past five consecutive quarters.



Discussion

Principal Findings

Restrictive measures put in place to mitigate transmission of the coronavirus SARS-CoV-2 (COVID-19) have had an immediate and unprecedented impact on people's psychological health and well-being, especially as they relate to the incidences and severity of depression and anxiety. This study sought to understand the impact that COVID-19 has had on young adults seeking mental health services at an outpatient program during the pandemic. The results indicated a significant increase in anxiety in young adults across the measured time periods of the pandemic, in particular between March and November of 2020, or phases 1-3 as defined in this study. This increase in anxiety coincides with a rising number of cases and COVID-19 becoming the third leading cause of death in the world, as death tolls were hitting their highest peak since 2015 (15.9% increase) [39]. The American Psychiatric Association issued a public opinion poll in 2020, which found that 80% of participants were either somewhat or extremely anxious about keeping themselves and their families safe in 2020, while 75% were worried about COVID-19 specifically.

The results also demonstrated that self-identified gender was a significant predictor of mental health. Female clients scored

lower than their male counterparts on measures of "quality of life" and "ability to participate in daily activities," both of which are subscores that are reliably indicative of broader mental well-being. This is consistent with the literature, which suggests that female-identified individuals are more significantly impacted by psychological distress than their male-identified counterparts [40-42]. In particular, female-identified individuals between the ages of 26-35 years are most at risk in terms of vulnerability to stress [40]. Future research should consider the biological, social, and cognitive mechanisms underlying these gender differences.

Comparison With Prior Work

Mental health conditions account for 16% of the global burden of disease and injury in people aged 10-19 years. Young adulthood is a critical time of interpersonal, intrapersonal, social, educational, and vocational development [43]. Research shows that the brain does not fully develop until the age of 25, and risk factors for mental health may be of increased concern for the developing brain [8]. Half of all mental health disorders in adulthood start by age 14 [44]. Furthermore, COVID-19 has only exacerbated these statistics, heightening the public health urgency [12,14,20-23]. While young adults were at lower risk for the direct health effects of hospitalization or death due to COVID-19, research is needed to understand any long-lasting mental health effects that may be detrimental to development. The findings of this study echo the need to understand the long-term public health impact on young adult mental well-being. Future research should continue to evaluate the public health impact of COVID-19 on young adults in a systematic manner.

The findings emphasize the importance of secondary mental health concerns among young adults, especially female clients in this case. While this study demonstrated that young adults experienced a significant increase in anxiety during the pandemic and that gender was a significant factor in the compromised mental well-being of young adults, it will be necessary to continue to study the lasting and secondary impact of anxiety to prepare for and provide successful treatment to young adults. Anxiety is correlated with and can exacerbate secondary mental health concerns and symptoms such as substance use, PTSD, and suicidality. For instance, the prevalence of PTSD in 2485 home-quarantined college students was reported to be 2.7% [25]. In a study that looked at the association between COVID-19 and the prevalence of suicidal thoughts and behaviors, researchers found that 45% of those who demonstrated past-month suicidal ideation reported that their thoughts were directly related to COVID-19 [19]. Finally, substance use may have increased significantly during COVID-19, especially among those who already had preexisting mental health conditions such as anxiety and depression [45,46]. These secondary mental health concerns have the potential to compromise young adults beyond their psychological and emotional well-being. Secondary mental health concerns can lead to reduced productivity at home, school, and in the labor market. Special attention to the secondary effects of heightened anxiety in young adults because of COVID-19–associated stressors is vital to ensure the critical need for mental health preparedness from a global perspective.

Limitations

There are several limitations of this study. The longitudinal research design limits the generalizability of the findings due to not having a repeated measurement. Future research should evaluate the long-term impact of similar findings in a more systematic manner that includes a more rigorous research design methodology. Additionally, the sample size is relatively small, thus providing an initial understanding of the acuity of anxiety and depression in young adults seeking intensive outpatient programming during the pandemic. Another limitation is the lack of racial and socioeconomic diversity within the sample. A majority of the sample were white individuals. Additionally, research has demonstrated that marginalized populations such as non-white Hispanic and Asian populations experienced higher levels of the subjective perception of distress, worry, and fear [47]. Further research is needed to explore the impact of COVID-19 on racially and ethnically diverse populations as well as those with different socioeconomic statuses. Additionally, we did not collect information on the socioeconomic status of the participants. Future research is needed to understand if this is a covariate for increased anxiety and depression. While we recognize this as a contributing limitation within the manuscript, it should be noted that longitudinal data discussing the impact of COVID-19 on mental health are currently lacking. Lastly, due to restrictions put in place by COVID-19 and the corresponding state regulations, treatment for individuals within the first quarter was primarily conducted virtually instead of in person. While this deviates from the current protocol in place, all attempts were made to make the virtual sessions as equally representative as the in-person interactions.

Conclusion

This study attempted to understand the impact of COVID-19 on the admissions of an outpatient program for young adults. It was apparent that COVID-19 had a larger impact on the quality of life for women than for men. Additionally, it was apparent that there was a significant difference in generalized anxiety between the first and third intervals during the COVID-19 pandemic. These findings highlight the need to better understand the mental health impact of the COVID-19 pandemic.

Authors' Contributions

AF wrote the majority of the manuscript and provided editing of the manuscript. AC conducted the literature review, wrote a portion of the manuscript, and edited the manuscript. BC edited the manuscript, formatted the manuscript, and wrote portions of the manuscript. KJ edited the manuscript and wrote portions of the manuscript. FB obtained the consent, designed the research study, conducted the analyses, wrote a portion of the manuscript, edited the manuscript, and provided mentorship.

Conflicts of Interest

AF, BC, AC, and KJ are employed by The Dorm; however, they were not responsible for the data management or analysis. FB has no potential conflicts of interest to disclose.

References

1. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920 [FREE Full text] [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
2. Cao W, Fang Z, Hou G, Han M, Xu X, Dong J, et al. The psychological impact of the COVID-19 epidemic on college students in China. Psychiatry Res 2020 May;287:112934 [FREE Full text] [doi: 10.1016/j.psychres.2020.112934] [Medline: 32229390]
3. Glowacz F, Schmits E. Psychological distress during the COVID-19 lockdown: the young adults most at risk. Psychiatry Res 2020 Nov;293:113486 [FREE Full text] [doi: 10.1016/j.psychres.2020.113486] [Medline: 33007682]
4. Hossain MM, Sultana A, Purohit N. Mental health outcomes of quarantine and isolation for infection prevention: a systematic umbrella review of the global evidence. Epidemiol Health 2020 Jun 02;42:e2020038. [doi: 10.4178/epih.e2020038] [Medline: 32512661]
5. Huang Y, Zhao N. Generalized anxiety disorder, depressive symptoms and sleep quality during COVID-19 outbreak in China: a web-based cross-sectional survey. Psychiatry Res 2020 Jun;288:112954 [FREE Full text] [doi: 10.1016/j.psychres.2020.112954] [Medline: 32325383]
6. Liu C, Stevens C, Conrad R, Hahm H. Evidence for elevated psychiatric distress, poor sleep, and quality of life concerns during the COVID-19 pandemic among U.S. young adults with suspected and reported psychiatric diagnoses. Psychiatry Res 2020 Oct;292:113345 [FREE Full text] [doi: 10.1016/j.psychres.2020.113345] [Medline: 32745794]
7. 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]
8. Sharma A, Pillai D, Lu M, Doolan C, Leal J, Kim J, et al. Impact of isolation precautions on quality of life: a meta-analysis. J Hosp Infect 2020 May;105(1):35-42. [doi: 10.1016/j.jhin.2020.02.004] [Medline: 32059996]
9. Zhang SX, Wang Y, Rauch A, Wei F. Unprecedented disruption of lives and work: health, distress and life satisfaction of working adults in China one month into the COVID-19 outbreak. Psychiatry Res 2020 Jun;288:112958 [FREE Full text] [doi: 10.1016/j.psychres.2020.112958] [Medline: 32283450]
10. Chan ASW, Ho JMC, Li JSF, Tam HL, Tang PMK. Impacts of COVID-19 pandemic on psychological well-being of older chronic kidney disease patients. Front Med 2021;8:666973. [doi: 10.3389/fmed.2021.666973] [Medline: 34124096]
11. Fofana NK, Latif F, Sarfraz S, Bilal, Bashir MF, Komal B. Fear and agony of the pandemic leading to stress and mental illness: an emerging crisis in the novel coronavirus (COVID-19) outbreak. Psychiatry Res 2020 Sep;291:113230 [FREE Full text] [doi: 10.1016/j.psychres.2020.113230] [Medline: 32593067]
12. González-Sanguino C, Ausín B, Castellanos M, Saiz J, López-Gómez A, Ugidos C, et al. Mental health consequences during the initial stage of the 2020 Coronavirus pandemic (COVID-19) in Spain. Brain Behav Immun 2020 Jul;87:172-176 [FREE Full text] [doi: 10.1016/j.bbi.2020.05.040] [Medline: 32405150]
13. Kibbey MM, Fedorenko EJ, Farris SG. Anxiety, depression, and health anxiety in undergraduate students living in initial US outbreak "hotspot" during COVID-19 pandemic. Cogn Behav Ther 2021 Sep 12;50(5):409-421. [doi: 10.1080/16506073.2020.1853805] [Medline: 33433271]
14. Kujawa A, Green H, Compas BE, Dickey L, Pegg S. Exposure to COVID-19 pandemic stress: associations with depression and anxiety in emerging adults in the United States. Depress Anxiety 2020 Dec 10;37(12):1280-1288. [doi: 10.1002/da.23109] [Medline: 33169481]
15. Ma Z, Zhao J, Li Y, Chen D, Wang T, Zhang Z, et al. Mental health problems and correlates among 746 217 college students during the coronavirus disease 2019 outbreak in China. Epidemiol Psychiatr Sci 2020 Nov 13;29:e181. [doi: 10.1017/s2045796020000931]
16. Ochnik D, Rogowska AM, Kuśnierz C, Jakubiak M, Schütz A, Held MJ, et al. Mental health prevalence and predictors among university students in nine countries during the COVID-19 pandemic: a cross-national study. Sci Rep 2021 Sep 20;11(1):18644. [doi: 10.1038/s41598-021-97697-3] [Medline: 34545120]
17. Reading Turchioe M, Grossman LV, Myers AC, Pathak J, Creber RM. Correlates of mental health symptoms among US adults during COVID-19, March-April 2020. Public Health Rep 2021 Nov 19;136(1):97-106 [FREE Full text] [doi: 10.1177/0033354920970179] [Medline: 33211985]
18. Holingue C, Kalb LG, Riehm KE, Bennett D, Kapteyn A, Veldhuis CB, et al. Mental distress in the United States at the beginning of the COVID-19 pandemic. Am J Public Health 2020 Nov;110(11):1628-1634. [doi: 10.2105/ajph.2020.305857]
19. Ammerman BA, Burke TA, Jacobucci R, McClure K. Preliminary investigation of the association between COVID-19 and suicidal thoughts and behaviors in the U.S. J Psychiatr Res 2021 Feb;134:32-38. [doi: 10.1016/j.jpsychires.2020.12.037] [Medline: 33360222]

20. Breslau J, Finucane ML, Locker AR, Baird MD, Roth EA, Collins RL. A longitudinal study of psychological distress in the United States before and during the COVID-19 pandemic. Prev Med 2021 Feb;143:106362. [doi: 10.1016/j.ypmed.2020.106362] [Medline: 33388325]
21. French MT, Mortensen K, Timming AR. Psychological distress and coronavirus fears during the initial phase of the COVID-19 pandemic in the United States. J Ment Health Policy Econ 2020 Sep 01;23(3):93-100. [Medline: 32853158]
22. Riedel-Heller S, Richter D. COVID-19 pandemic and mental health of the general public: is there a tsunami of mental disorders? Psychiatr Prax 2020 Nov 02;47(8):452-456. [doi: 10.1055/a-1290-3469] [Medline: 33137827]
23. Wathelet M, Duhem S, Vaiva G, Baubet T, Habran E, Veerapa E, et al. Factors associated with mental health disorders among university students in France confined during the COVID-19 pandemic. JAMA Netw Open 2020 Oct 01;3(10):e2025591 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.25591] [Medline: 33095252]
24. Zheng J, Morstead T, Sin N, Klaiber P, Umberson D, Kamble S, et al. Psychological distress in North America during COVID-19: the role of pandemic-related stressors. Soc Sci Med 2021 Feb;270:113687. [doi: 10.1016/j.socscimed.2021.113687] [Medline: 33465600]
25. Tang W, Hu T, Hu B, Jin C, Wang G, Xie C, et al. Prevalence and correlates of PTSD and depressive symptoms one month after the outbreak of the COVID-19 epidemic in a sample of home-quarantined Chinese university students. J Affect Disord 2020 Sep 01;274:1-7 [FREE Full text] [doi: 10.1016/j.jad.2020.05.009] [Medline: 32405111]
26. García-Portilla P, de la Fuente Tomás L, Bobes-Bascarán T, Jiménez Treviño L, Zurrón Madera P, Suárez Álvarez M, et al. Are older adults also at higher psychological risk from COVID-19? Aging Ment Health 2021 Jul 01;25(7):1297-1304. [doi: 10.1080/13607863.2020.1805723] [Medline: 32870024]
27. Power E, Hughes S, Cotter D, Cannon M. Youth mental health in the time of COVID-19. Ir J Psychol Med 2020 Dec;37(4):301-305 [FREE Full text] [doi: 10.1017/ipm.2020.84] [Medline: 32611470]
28. Turna J, Zhang J, Lamberti N, Patterson B, Simpson W, Francisco AP, et al. Anxiety, depression and stress during the COVID-19 pandemic: results from a cross-sectional survey. J Psychiatr Res 2021 May;137:96-103 [FREE Full text] [doi: 10.1016/j.jpsychires.2021.02.059] [Medline: 33667763]
29. Emery RL, Johnson ST, Simone M, Loth KA, Berge JM, Neumark-Sztainer D. Understanding the impact of the COVID-19 pandemic on stress, mood, and substance use among young adults in the greater Minneapolis-St. Paul area: Findings from project EAT. Soc Sci Med 2021 May;276:113826 [FREE Full text] [doi: 10.1016/j.socscimed.2021.113826] [Medline: 33743209]
30. Kecojevic A, Basch CH, Sullivan M, Davi NK. The impact of the COVID-19 epidemic on mental health of undergraduate students in New Jersey, cross-sectional study. PLoS One 2020;15(9):e0239696 [FREE Full text] [doi: 10.1371/journal.pone.0239696] [Medline: 32997683]
31. Preetz R, Filser A, Brömmelhaus A, Baalmann T, Feldhaus M. Longitudinal changes in life satisfaction and mental health in emerging adulthood during the COVID-19 pandemic. Risk and protective factors. Emerg Adulthood 2021 Oct 05;9(5):602-617. [doi: 10.1177/21676968211042109]
32. 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]
33. 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]
34. Patton JH, Stanford MS, Barratt ES. Factor structure of the Barratt impulsiveness scale. J Clin Psychol 1995 Nov;51(6):768-774. [doi: 10.1002/1097-4679(199511)51:6<768::aid-jclp2270510607>3.0.co;2-1] [Medline: 8778124]
35. Stanford M, Mathias C, Dougherty D, Lake S, Anderson N, Patton J. Fifty years of the Barratt Impulsiveness Scale: an update and review. Pers Individ Diff 2009 Oct;47(5):385-395. [doi: 10.1016/j.paid.2009.04.008]
36. Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation. Health Qual Life Outcomes 2003 Dec 16;1:79 [FREE Full text] [doi: 10.1186/1477-7525-1-79] [Medline: 14678568]
37. Kupst MJ, Butt Z, Stoney CM, Griffith JW, Salsman JM, Folkman S, et al. Assessment of stress and self-efficacy for the NIH Toolbox for Neurological and Behavioral Function. Anxiety Stress Coping 2015 Feb 10;28(5):531-544 [FREE Full text] [doi: 10.1080/10615806.2014.994204] [Medline: 25577948]
38. Hanish AE, Lin-Dyken DC, Han JC. PROMIS sleep disturbance and sleep-related impairment in adolescents: examining psychometrics using self-report and actigraphy. Nurs Res 2017;66(3):246-251 [FREE Full text] [doi: 10.1097/NNR.0000000000000217] [Medline: 28448375]
39. Ahmad FB, Cisewski JA, Miniño A, Anderson RN. Provisional mortality data - United States, 2020. MMWR Morb Mortal Wkly Rep 2021 Apr 09;70(14):519-522. [doi: 10.15585/mmwr.mm7014e1] [Medline: 33830988]
40. Al Dhaheri AS, Bataineh MF, Mohamad MN, Ajab A, Al Marzouqi A, Jarrar AH, et al. Impact of COVID-19 on mental health and quality of life: Is there any effect? A cross-sectional study of the MENA region. PLoS One 2021 Mar 25;16(3):e0249107 [FREE Full text] [doi: 10.1371/journal.pone.0249107] [Medline: 33765015]
41. Mazza C, Ricci E, Biondi S, Colasanti M, Ferracuti S, Napoli C, et al. A nationwide survey of psychological distress among Italian people during the COVID-19 pandemic: immediate psychological responses and associated factors. Int J Environ Res Public Health 2020 May 02;17(9):3165 [FREE Full text] [doi: 10.3390/ijerph17093165] [Medline: 32370116]

42. 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]
43. Wood D, Crapnell T, Lau L, Bennett A, Lotstein D, Ferris M. Emerging adulthood as a critical stage in the life course. In: Halfon N, Forrest CB, Lerner RM, Faustman EM, editors. Handbook of life course health development. Cham: Springer International Publishing; 2018:123-143.
44. Adolescent and young adult health. World Health Organization. 2021 Jan 18. URL: https://www.who.int/news-room/fact-sheets/detail/adolescents-health-risks-and-solutions [accessed 2022-07-01]
45. Capasso A, Jones A, Ali S, Foreman J, Tozan Y, DiClemente R. Increased alcohol use during the COVID-19 pandemic: the effect of mental health and age in a cross-sectional sample of social media users in the U.S. Prev Med 2021 Apr;145:106422 [FREE Full text] [doi: 10.1016/j.ypmed.2021.106422] [Medline: 33422577]
46. Lechner WV, Laurene KR, Patel S, Anderson M, Grega C, Kenne DR. Changes in alcohol use as a function of psychological distress and social support following COVID-19 related University closings. Addict Behav 2020 Nov;110:106527 [FREE Full text] [doi: 10.1016/j.addbeh.2020.106527] [Medline: 32679435]
47. Fitzpatrick KM, Drawve G, Harris C. Facing new fears during the COVID-19 pandemic: the state of America's mental health. J Anxiety Disord 2020 Oct;75:102291 [FREE Full text] [doi: 10.1016/j.janxdis.2020.102291] [Medline: 32827869]

Abbreviations

BIS-11: Barratt Impulsivity Scale
FACIT: Functional Assessment of Chronic Illness Therapy
GAD-7: Generalized Anxiety Disorder scale
PHQ-9: Patient Health Questionnaire
PROMIS: Patient Reported Outcomes Measurement Information System
PTSD: posttraumatic stress disorder

Edited by A Mavragani; submitted 02.05.22; peer-reviewed by J Sung, ASW Chan ; comments to author 03.06.22; revised version received 21.06.22; accepted 27.06.22; published 14.07.22.

Please cite as:
Fialk A, Connors A, Cerrito B, Jones K, Buono F
The Impact of SARS-CoV-2 (COVID-19) on the Acuity of Mental Health–Related Diagnosis at Admission for Young Adults in New York City and Washington, DC: Observational Study
JMIR Form Res 2022;6(7):e39217
URL: https://formative.jmir.org/2022/7/e39217
doi:10.2196/39217
PMID:35767688

Original Paper

Remote Analysis of Respiratory Sounds in Patients With COVID-19: Development of Fast Fourier Transform–Based Computer-Assisted Diagnostic Methods

Gregory Furman[1*], Prof Dr; Evgeny Furman[2*], Prof Dr; Artem Charushin[3*], Prof Dr; Ekaterina Eirikh[3], Prof Dr; Sergey Malinin[4], BSc; Valery Sheludko[5*], MD; Vladimir Sokolovsky[1*], Prof Dr; David Shtivelman[1*], MA

[1]Physics Department, Ben-Gurion University of the Negev, Beer-Sheva, Israel

[2]Department of Pediatric, EA Vagner Perm State Medical University, Perm, Russian Federation

[3]Department of Ear, Nose and Throat, EA Vagner Perm State Medical University, Perm, Russian Federation

[4]Central Research Laboratory, EA Vagner Perm State Medical University, Perm, Russian Federation

[5]Perm Regional Clinical Infectious Diseases Hospital, Perm, Russian Federation

[*]these authors contributed equally

Corresponding Author:
Gregory Furman, Prof Dr
Physics Department
Ben-Gurion University of the Negev
POB 653
Beer-Sheva, 8410501
Israel
Phone: 972 +972547684245
Email: gregoryf@bgu.ac.il

Abstract

Background: Respiratory sounds have been recognized as a possible indicator of behavior and health. Computer analysis of these sounds can indicate characteristic sound changes caused by COVID-19 and can be used for diagnostics of this illness.

Objective: The aim of the study is to develop 2 fast, remote computer-assisted diagnostic methods for specific acoustic phenomena associated with COVID-19 based on analysis of respiratory sounds.

Methods: Fast Fourier transform (FFT) was applied for computer analysis of respiratory sound recordings produced by hospital doctors near the mouths of 14 patients with COVID-19 (aged 18-80 years) and 17 healthy volunteers (aged 5-48 years). Recordings for 30 patients and 26 healthy persons (aged 11-67 years, 34, 60%, women), who agreed to be tested at home, were made by the individuals themselves using a mobile telephone; the records were passed for analysis using WhatsApp. For hospitalized patients, the illness was diagnosed using a set of medical methods; for outpatients, polymerase chain reaction (PCR) was used. The sampling rate of the recordings was from 44 to 96 kHz. Unlike usual computer-assisted diagnostic methods for illnesses based on respiratory sound analysis, we proposed to test the high-frequency part of the FFT spectrum (2000-6000 Hz).

Results: Comparing the FFT spectra of the respiratory sounds of patients and volunteers, we developed 2 computer-assisted methods of COVID-19 diagnostics and determined numerical healthy-ill criteria. These criteria were independent of gender and age of the tested person.

Conclusions: The 2 proposed computer-assisted diagnostic methods, based on the analysis of the respiratory sound FFT spectra of patients and volunteers, allow one to automatically diagnose specific acoustic phenomena associated with COVID-19 with sufficiently high diagnostic values. These methods can be applied to develop noninvasive screening self-testing kits for COVID-19.

(JMIR Form Res 2022;6(7):e31200) doi:10.2196/31200

KEYWORDS

COVID-19; audio analysis; remote computer diagnosis; respiratory sounds; respiratory analysis; modeling; computer-assisted methods; diagnostics

Introduction

The World Health Organization (WHO) reported that till October 16, 2021, about 241 million people were infected with the novel coronavirus (COVID-19) worldwide and about 18 million people currently have the disease [1]. COVID-19 is a public health problem in all countries regardless of their level of development.

It is known that SARS-CoV-2 causes a severe lower respiratory disease with high mortality and evidence of systemic spread [2]. The virus is able to actively multiply in the epithelium of the airways. Intense cough is 1 of the main symptoms of COVID-19. It is known that the highest density of cough receptors is in the larynx [3]. Anatomically, a dry cough can be associated with the effect of the virus on the cough receptors of the larynx due to infection with COVID-19. SARS-CoV-2 can penetrate into the smallest airways, where it infects cells and causes bilateral pneumonia, often with respiratory failure [4-7]. The damage of various airways, caused by SARS-CoV-2, alters sound formation in the patient and changes the characteristics of respiratory sounds. Detection of characteristic respiratory sounds (cough, wheezes, asthma wheezing, shortness of breath, etc) is a widely used way of diagnostics of pulmonary diseases, which is applied to develop new computer-assisted diagnostic methods (eg, [8-12] and references therein).

At present, diagnostics of COVID-19 is based on clinical symptoms, chest X-ray/computed tomography (CT), and coronavirus tests (polymerase chain reaction [PCR]), molecular tests, antigen tests, and specific antibodies to SARS-CoV-2) [6,7,13]. New and more contagious COVID-19 virus strains are appearing in the United Kingdom, the Republic of South Africa, Vietnam, and some other countries; in India, the daily number of new infected persons was up to 300,000 in May 2021. Therefore, it is desirable to develop vast, cheap, and widely available remote methods of COVID-19 diagnostics. One of these methods can be based on computer-assisted analysis of respiratory sounds of the patient and on comparison of the sound characteristics between a patient and a healthy volunteer.

The objectivity of auscultatory diagnostics can be significantly enhanced by using digitized audio signals and computer processing of these signals. Automated adventitious sound detection or classification is a promising solution to overcome the limitations of conventional auscultation and to assist in the monitoring of relevant diseases, such as asthma, chronic obstructive pulmonary disease (COPD), and pneumonia [14]. Olvera-Montes et al [15] used the detection of respiratory crackle sounds through an Android smartphone–based system for the diagnostics of pneumonia and monitoring of the patient's state.

Reyes et al [16] used a smartphone-based system for automated bedside detection of crackle sounds in patients with diffuse interstitial pneumonia. The performance of automated detection was analyzed using (1) synthetic fine and coarse crackle sounds randomly inserted into basal respiratory sounds acquired from healthy subjects with different signal-to-noise ratios and (2) real bedside-acquired respiratory sounds from patients with interstitial diffuse pneumonia. In simulated scenarios, for fine crackles, an accuracy ranging from 84.86% to 89.16%, a sensitivity ranging from 93.45% to 97.65%, and a specificity ranging from 99.82% to 99.84% were found. The detection of coarse crackles was found to be a more challenging task in the simulated scenarios. In the case of real data, the results show the feasibility of using the developed mobile health system in a clinical noncontrolled environment to help the expert in evaluating the pulmonary state of a subject.

The overview concerns the potential for computer audition (CA), that is, the use of speech and sound analysis by artificial intelligence to help in COVID-19 diagnostics [17]. Automatic recognition and monitoring of breathing, dry and wet coughing or sneezing sounds, speech under cold, eating behavior, sleepiness, or pain are used. Schuller et al [17] concluded that CA appears to be ready for implementation of (pre-) diagnosis and monitoring tools.

It was considered that the acquired breathing sounds can be analyzed using advanced signal processing and in tandem with new deep machine learning and pattern recognition techniques to separate the breathing phases, estimate the lung volume, estimate oxygenation, and further classify the breathing data input into healthy or unhealthy cases [18]. Computer analysis of breath sounds can be important for identification of specific changes in these sounds, caused by COVID-19.

Brown et al [19] used the exploring automatic diagnostics of COVID-19 from crowdsourced respiratory sound data. The results of early works [16-19] and references therein allow one to suggest that respiratory sounds can be useful in COVID-19 diagnostics.

The purpose of this study is to 2 develop fast, remote methods of diagnostics of specific acoustic phenomena associated with COVID-19 based on computer-assisted analysis of respiratory sounds. The developed methods are based on analysis of fast Fourier transform (FFT) spectra of respiratory sounds recorded near the mouth. We proposed to use a personal computer, a modern mobile telephone, or a smartphone for registration, recording of respiratory sounds, and their analysis. The developed methods can be applied as additional screening methods of COVID-19 diagnostics and as personalized screening self-testing kits for COVID-19. Such self-tests would serve as an early step before further procedures ordered by a doctor (PCR test, lung CT, X-ray, etc). We restricted our work to consideration of breathing sounds (and not cough and voice samples). Lung diseases, such as asthma, COPD, and pneumonia, cause specific changes in the FFT spectra of respiratory sounds in the frequency range from 100 to 2500 Hz, and this range is usually considered during development of computer-assisted diagnostic methods (eg, [8-12,20-22] and references therein). We considered the larger frequency range up to 6000 Hz, and it was shown that for diagnostics of COVID-19, the frequency range from 2000 to 6000 Hz is significant. This allows us to diagnose specific acoustic phenomena associated with COVID-19 for patients with other lung diseases as well.

Methods

Patients

In this study, 14 patients with COVID-19 and 17 healthy volunteers participated. COVID-19 in the patients was diagnosed using medical methods, such as analysis of clinical symptoms, chest X-ray and CT, coronavirus tests (PCR test for SARS-CoV-2 RNA [the main method], specific antibody test for SARS-CoV-2). The clinical examination of the patients and the recording of their respiratory sounds were carried out at Perm Infectious Hospital, and the volunteers' respiratory sounds were recorded at Perm State Medical University (Perm, Russia).

Ethical Considerations

This study complied with the Declaration of Helsinki (adopted in June 1964, Helsinki, Finland), revised in October 2000 (Edinburg, Scotland) and was overseen by the independent ethics committee of Perm State Medical University (approval code: 5/21). Written agreements from the patients and volunteers were obtained.

Methods of Recording Respiratory Sounds

Doctors performed all the recordings, and patients and volunteers were instructed to remain calm and to breathe easily. No special measures to reduce ambient noise were applied. The respiratory sounds of patients were recorded in m4a format using a Honor dua-1 22 smartphone at a distance of 2 cm from the mouth for about 20 s; the sampling rate was 48 kHz. The respiratory sounds of volunteers were recorded near the mouth using a mobile telephone (the sampling rate was 44.1 kHz, mp3 format) and a computer-based recording system (the sampling rate was 96 kHz, wav format) [20-25]. Our comparison of the records carried out with the help of various devices showed that although the mp3 and m4a formats compress a signal, the FFT spectra of a sound recorded in the wav, mp3, and m4a formats are practically the same. To analyze respiratory sounds, we proposed to consider the normalized FFT spectra that largely decrease the influence of the recording format.

Simultaneous recordings of respiratory sounds in the same point of a patient in different formats were made. The normalized integral characteristics of the FFT spectra used for diagnostics differed by less than 3% for different formats.

The beginnings and ends of the recordings made using the smartphone and the mobile telephone contained temporal parts, in which respiratory sounds are not recorded. In these parts, there are short impulses, the amplitudes of which can be much larger than the respiratory sound maximum. The impulses do not reflect the processes in the airways. The applied preliminary processing removed these parts.

Diagnostic Methods

The 2 proposed diagnostic methods were based on the fact that lung diseases cause changes in the airways and these changes are reflected in the spectra of respiratory sounds. This approach has been applied in the development of computer-assisted methods of diagnostics of various lung diseases [8-12,20-26]. These developed methods were based on the analysis of the FFT spectra of respiratory sounds in the frequency range from 100 to 2500 Hz and the comparison of the spectra of patients and healthy volunteers.

We proposed to compare different parts of the FFT spectrum of a patient with COVID-19 in a higher frequency range from 2000 to 6000 Hz. Examples of the amplitude-frequency dependences (spectra) of FFT for a patient with COVID-19 and a healthy volunteer are presented in Figure 1. In many cases, the FFT spectra of volunteers do not possess well-defined maximums and minimums, as shown in Figure 1a. One can see several differences (Table 1) in the spectra for a volunteer and a patient: the maxima and minima of their spectra are located in various frequency ranges; these differences could be used to formulate potential healthy-ill criteria, which are presented in Table 2. Similar locations of the maxima and minima were observed, for example, in the spectra for the first, third, and ninth patients. The clearly seen amplitude increase in the frequency range from 1000 to 1500 Hz in the spectrum for the first patient (Figure 1b) can be a result of concomitant disease (upper respiratory tract infection).

In Table 2:

where f_a is the frequency of the extremum, A(f) is the harmonic amplitude at frequency f, and Δf is the half of the frequency range, chosen equal to 300 Hz. In the program, the integral is replaced by a sum of harmonic amplitudes with a frequency in the range from $(f_a - \Delta f)$ to $(f_a + \Delta f)$.

The criteria under test were formulated as ratios of the integrals of the harmonic amplitudes over various frequency ranges (Table 2 and Equation 1). So, the criteria values were independent of the breathing intensity.

Adventitious sounds caused by an illness change not only the amplitude-frequency dependence of the respiratory sound FFT spectrum but also the frequency-amplitude dependence. The second proposed method is based on analysis of differences between the frequency-amplitude dependences of FFT spectra for patients and volunteers.

The moments of the frequency (MF) distribution can be considered as potential healthy-ill criteria:

and

where i_{min} and i_{max} are the harmonic numbers corresponding to the minimal (f_{min}=2000 Hz) and maximal (f_{max}=5900 Hz) frequencies, respectively, and f_i is the frequency of the i-th harmonic.

The MF distribution was also independent of the breathing intensity.

Figure 1. Amplitudes of FFT harmonics for the first healthy volunteer (a) and the first patient with COVID-19 (b). The amplitudes are given in arbitrary units. FFT: fast Fourier transform.

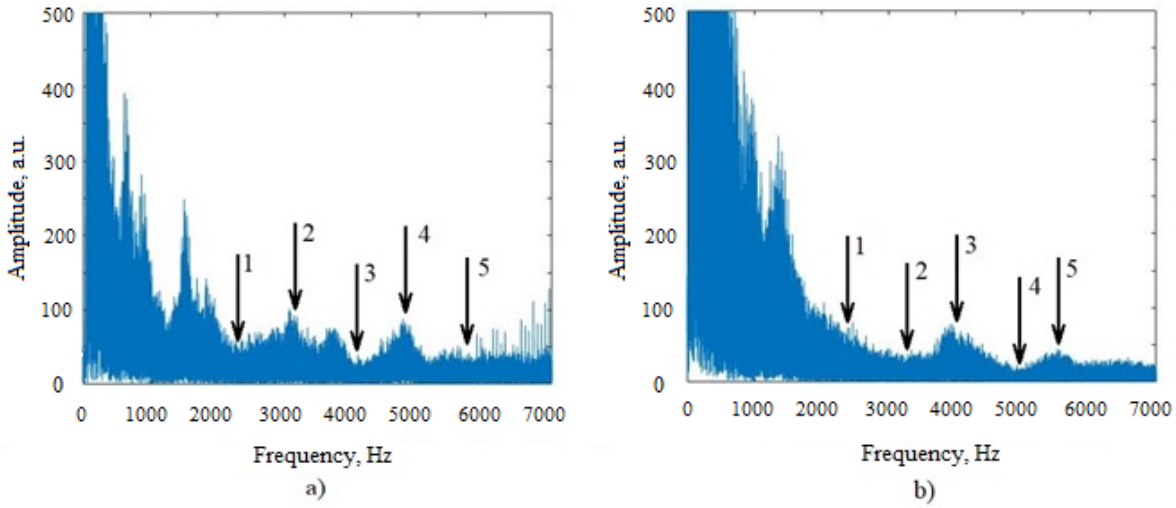


Table 1. Comparison of the FFT[a] spectra of a volunteer and a patient with COVID-19.

Volunteer	Patient
Minimum at about 2300 Hz	No extremum in the frequency range
Maximum at 3100 Hz	Minimum at 3300 Hz
Minimum at 4100 Hz	Maximum at 3900 Hz
Maximum at 4900 Hz	Minimum at 5000 Hz
No extremum at a frequency above 5300 Hz	Maximum at 5600 Hz

[a]FFT: fast Fourier transform.

Table 2. Healthy-ill criteria.

Criteria	Healthy	Ill
k_1 = I(2300)/I(3200)	$k_1<1$	$k_1>1$
k_2 = I(3200)/I(4000)	$k_2>1$	$k_2<1$
k_3 = I(4000)/I(5000)	$k_3<1$	$k_3>1$
k_4 = I(5000)/I(5600)	$k_4>1$	$k_4<1$

Results

The First Method

The results of the calculation of potential healthy-ill criteria k_1, k_2, k_3, and k_4 are presented in Tables 3 and 4.

For the reader's convenience, Figure 2 presents the distribution of the criteria through patients with COVID-19 and volunteers. Comparing the results presented in Tables 3 and 4 and Figure 2, one can see that the most reliable result is given by the high-frequency criterion k_4.

For healthy volunteers, the criterion should be >1; this was correct in T_N=15 (88.2%) of 17 cases and incorrect in F_N=2 (11.8%) cases. For patients with COVID-19, the criterion should be <1; this was observed in T_P=11 (78.6%) of 14 cases. COVID-19 was not diagnosed in F_P=3 (21.4%) cases.

The criterion for the 12th patient was close to the boundary value of 1. The second proposed method also had the healthy-ill criterion close to the boundary value (Figure 3). Analysis of the recording for the patient showed that the outside noise was high, and this can cause incorrect diagnostics.

Additionally, we applied the proposed method for diagnostics of COVID-19 using the respiratory sound recordings of 30 patients and 26 healthy persons (ages 11-67 years, 34, 60%, women). None of them reported that they had other lung diseases. These recordings were made by the persons themselves, who agreed to be tested at home, using a mobile telephone. The sound recordings were passed to us through WhatsApp. The cases of COVID-19 illness were confirmed by PCR tests, and in all cases, the illness was asymptomatic or mild. The proposed method gave the correct diagnosis for 26 (86.7%) of 30 patients and 22 (84.6%) of 26 healthy ones. The

PCR tests for persons who were incorrectly diagnosed as ill gave negative results.

Several patients made a few recordings during the process of the disease. Development of a spectrum of a patient is presented in Figure 3. Though the spectrum varied noticeably, the k_4 criterion was less than 1 during the illness. The pathological process was characterized by a clearly seen increase in harmonic amplitudes in a wide frequency range. The third-day spectrum corresponded to the most severe condition of the person (according to their message).

Table 3. Criteria k_1, k_2, k_3, and k_4 for patients with COVID-19.

Patient number	Age (years), gender (F=female, M=male)	Diagnosis	k_1	k_2	k_3	k_4
1	18, F	COVID-19, upper respiratory tract infection	>1	<1	>1	<1
2	80, F	COVID-19, unilateral pneumonia	>1	>1	>1	<1
3	47, F	COVID-19, bilateral pneumonia	>1	<1	>1	<1
4	58, F	COVID-19, bilateral pneumonia	>1	>1	>1	<1
5	62, M	COVID-19, pneumonia	>1	>1	>1	<1
6	65, F	COVID-19, pneumonia	<1	<1	<1	>1
7	28, F	COVID-19, unilateral pneumonia with hydrothorax, HIV infection	<1	>1	>1	<1
8	75, M	COVID-19, upper respiratory tract infection	<1	>1	>1	<1
9	38, F	COVID-19, upper respiratory tract infection, exacerbation of COPD[a]	>1	<1	>1	<1
10	36, F	COVID-19, pneumonia	>1	>1	>1	>1
11	20, M	COVID-19, pneumonia	>1	>1	>1	<1
12	56, M	COVID-19, pneumonia	>1	>1	>1	~1
13	20, M	COVID-19, pneumonia	>1	>1	>1	<1
14	32, M	COVID-19, pneumonia	>1	>1	>1	<1

[a]COPD: chronic obstructive pulmonary disease.

Table 4. Criteria k_1, k_2, k_3, and k_4 for volunteers.

Volunteer number	Age (years), gender (F=female, M=male)	k_1	k_2	k_3	k_4
1	22, F	<1	>1	<1	>1
2	47, M	>1	>1	>1	>1
3	48, F	>1	<1	<1	>1
4	17, M	>1	>1	<1	<1
5	8, M	>1	<1	>1	>1
6	5, F	>1	<1	>1	>1
7	5, F	>1	>1	>1	>1
8	11, M	>1	>1	>1	>1
9	5, M	>1	>1	>1	>1
10	14, M	>1	<1	>1	>1
11	5, F	>1	>1	>1	>1
12	12, F	>1	>1	>1	>1
13	9, M	<1	>1	>1	>1
14	10, F	<1	<1	>1	>1
15	10, F	<1	<1	>1	>1
16	8, M	>1	<1	>1	>1
17	14, M	>1	<1	>1	<1

Figure 2. Healthy-ill criteria according to Tables 3 and 4: blue circles for volunteers and red squares for patients.

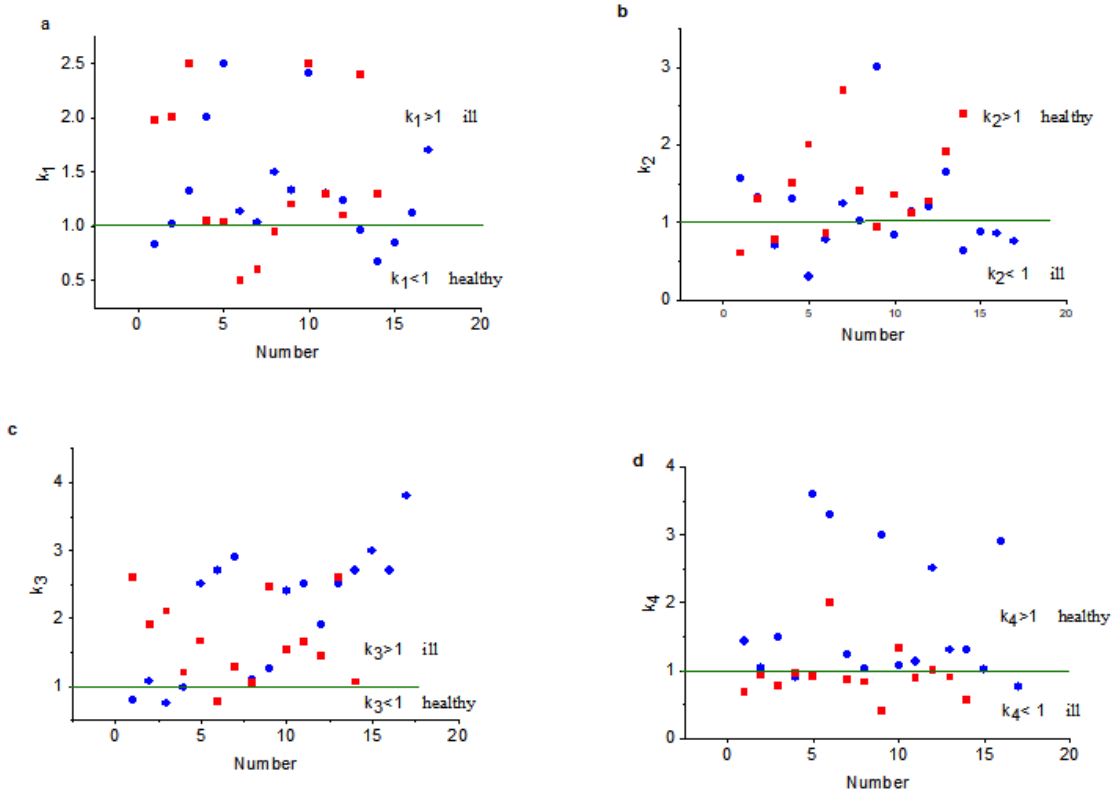
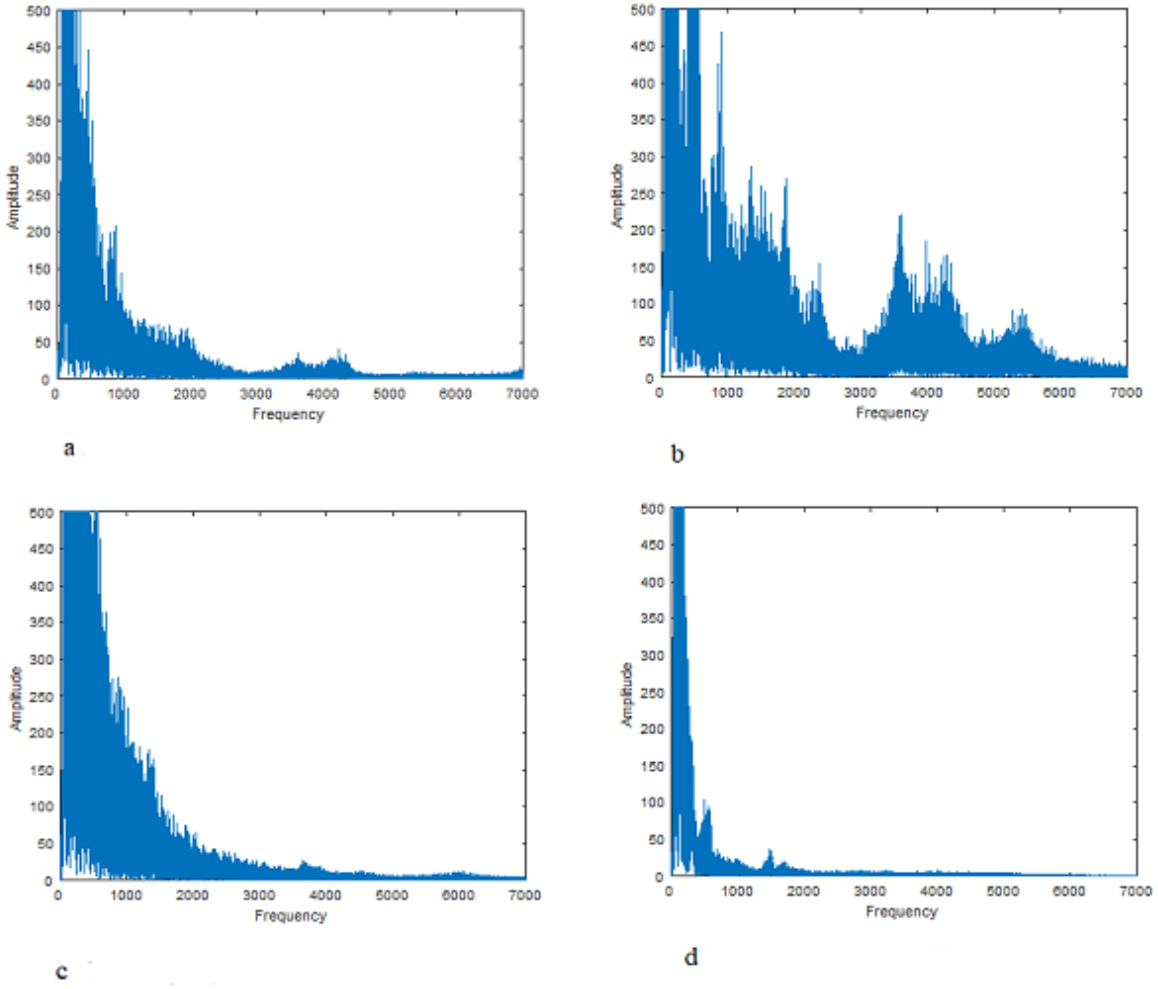


Figure 3. Change in the respiratory sound FFT spectrum of a patient with COVID-19 in the disease process. On the first day, the illness was diagnosed using the PCR test; on the last (11th) day, the person was diagnosed as healthy. The amplitudes are given in arbitrary units. FFT: fast Fourier transform; PCR: polymerase chain reaction.



The Second Method

We compared the values of MF and MF_n (n=2, 3, and 4) as well as MF_n/MF and MF_n/MF_2. The best result was obtained for MF_4/MF (Figure 4). In this figure, the blue line corresponds to the boundary value, which was selected by us as equal to 0.8: If $(MF_4/MF)\times10^{-9}$ is >0.8, the examined person is ill; if it is <0.8, the person is healthy. The second method, like the first one, gave incorrect diagnostics for the sixth patient and overdiagnosis for the fourth volunteer. The respiratory sound recording for this volunteer was characterized by a relatively low level of a signal and high noise. The overdiagnosis for the fourth volunteer could be the result of low quality of the recorded signal.

The second method correctly diagnosed patients as sick in T_P=13 (92.8%) of 14 cases and volunteers as healthy in T_N=14 (82.3%) of 17 cases. The method misdiagnosed patients as healthy only once, F_P=1 (7.2%), and healthy volunteers as sick in F_N=3 (17.7%) cases.

The second method applied to the respiratory sounds recorded by persons at home demonstrated a diagnostics effectivity that was close to that of the first method.

The results allow us to assess the main characteristics of the 2 proposed methods. Here, for characterization of the proposed methods, we used sensitivity, specificity, and the Youden index. The sensitivity of a method is determined by the formula [27]

and its specificity as

The sensitivity and specificity of the first method from Tables 1 and 2 were estimated as S_e=0.786 and S_p=0.882, respectively, and for the second method as (Figure 2) as S_e=0.93 and S_p=0.824. Here, we considered the results obtained only for people who were tested at the hospital.

The Youden index is calculated using the following formula:

The Youden index for the first method was about 0.67 and for the second was 0.754.

For the reader's convenience, these results and the sensitivity, specificity, and Youden index for both methods are presented in Table 5.

Figure 4. Healthy-ill criterion ($MF_4/MF)\times10^{-9}$: blue circles for volunteers and red squares for patients. Each number on the horizontal axis indicates the number of a patient or a volunteer according to Tables 3 and 4, respectively. The blue line corresponds to the boundary value of 0.8. MF: moments of the frequency.

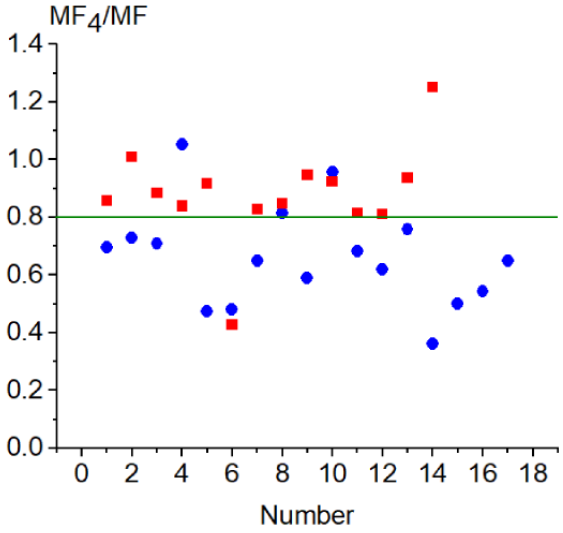


Table 5. Comparison of the 2 proposed methods.

Method	T_P	F_P	T_N	F_N	S_e	S_p	J
First	11	3	15	2	0.786	0.882	0.67
Second	13	1	14	3	0.93	0.824	0.754

Discussion

Principal Findings

The proposed computer-assisted diagnostic methods for COVID-19, which are based on analysis of respiratory sounds recorded near the mouth, demonstrated high diagnostic accuracy. For people tested at the hospital, the second method demonstrated better characteristics (sensitivity of 0.93, specificity of 0.824, and Youden index of 0.754) than the first one.

Both methods demonstrated close diagnostic characteristics when analyzing respiratory sound recordings made by persons themselves using a mobile telephone at home and submitted to us through WhatsApp. The proposed methods correctly diagnosed 86.7% of patients and 84.6% of healthy ones. These results demonstrate the possible application of the proposed methods for remote diagnostics.

Although a relatively low number (due to pandemic limitations) of the examined patients with COVID-19 and healthy volunteers did not allow estimating the method characteristics with high accuracy, the proposed methods correctly diagnose patients with COVID-19 in a wide age range, and the proposed criteria of healthy/ill are independent of the patient's age, sex, etc, as well as concomitant diseases, such as upper respiratory tract infection, pneumonia, exacerbation of COPD, and bilateral pneumonia (Table 3).

The patient and volunteer groups contained members of various genders and ages (from 5 to 80 years).

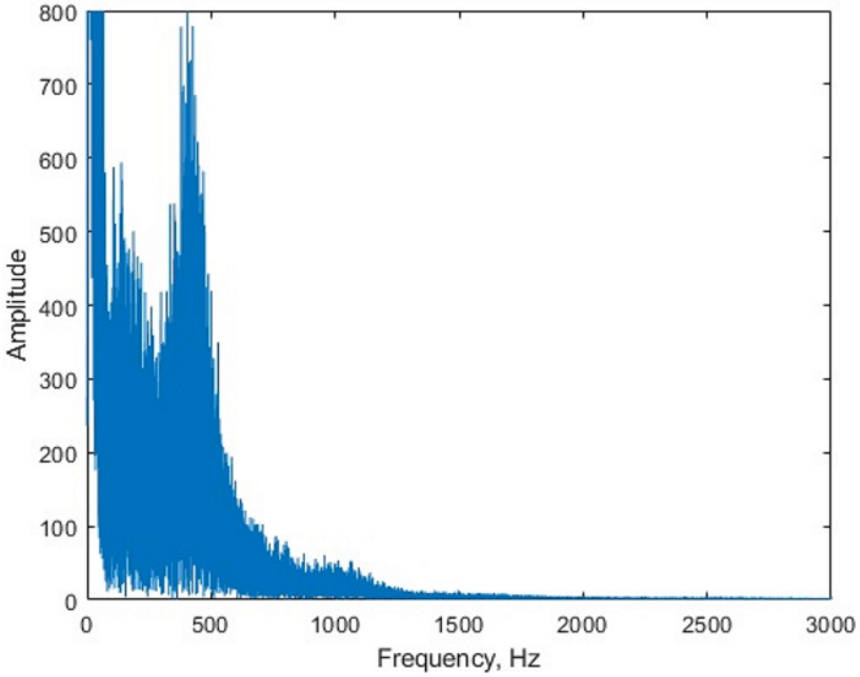
High diagnostic characteristics of the proposed methods independent of age were achieved due to comparison of various parts of the respiratory sound FFT spectrum. For example, in the first method, the ratio k_4 of integrals over various frequency ranges was determined and compared with the boundary value: If $k_4>1$, the examined person is healthy, and if $k_4<1$, the person is ill. The ratio is independent of the recording devices and sampling rate, and as our results showed, the boundary value does not depend on the individual patient characteristics and even on concomitant diseases (Table 3).

Comparison With Prior Work

Each lung disease is characterized by specific changes in the airways or lungs. These changes cause abnormal (adventitious) sounds, which can be separated into several types (wheezing, stridor, crackles, etc). These abnormal sounds are characterized by their durations and specific frequency ranges being below 2500 Hz [8,10,12]. For example, bronchial asthma is characterized by airway obstruction and inflammatory process, which covers all airways, from the central to the peripheral parts of the tracheobronchial tree (small bronchi) [21]. Asthmatic changes in the lungs cause typical respiratory sounds, with the main frequency in low-frequency ranges between 100 and 1000 Hz [12,28,29] and between 400 and 1600 Hz [12,30] (Figure 5). For analysis of respiratory sounds of patients with pulmonary diseases (asthma, pneumonia, HIV infection, etc) and

development of computer-assisted diagnostic methods, the frequency range from 100 to 2500 Hz is usually considered.

Figure 5. FFT spectra of sound signals for an asthmatic patient. The amplitudes are given in arbitrary units. FFT: fast Fourier transform.



For diagnostics of specific acoustic phenomena associated with COVID-19, we proposed to consider the high-frequency range of respiratory sound FFT spectra. It can be assumed that an upper respiratory tract injury leads to the appearance of changes in a higher-frequency part of the spectrum; it can be associated with COVID-19 [31]. Severe forms of tracheobronchitis were consistently present in 88% of COVID-19 cases [32-34]. This assumption is confirmed by our results: the most reliable criterion is the k_4 criterion, which is determined for the high-frequency range of the FFT spectrum, from 4700 to 5900 Hz (the first proposed method).

A decrease in the diagnostic value of the k_1, k_2, and k_3 criteria can be the result of concomitant diseases, which cause changes in the respiratory sounds in the lower-frequency range [8,10,12]. Another reason of the low accuracy of diagnostics based on the k_1, k_2, and k_3 criteria can be a higher sensitivity of the parameters f_a and Δf at low frequencies to individual characteristics of patients (age, sex, weight, etc) and also to fatigue and anxiety.

The second proposed method of computer-assisted diagnostics of COVID-19 is also based on the consideration of the high-frequency range of the FFT spectrum, from 2000 to 6000 Hz.

The high diagnostic accuracy is achieved in both methods due to our offer to compare various parts of the FFT spectrum of a patient (volunteer). This allows us to minimize the influence of the breathing intensity as well as the gender and age dependences of the FFT spectrum.

One of the ways to increase the diagnostic values of the proposed computer-assisted methods is to create a big database and determine the parameters (f_a and Δf for the first method and f_{min} and f_{max} for the second one) using machine learning.

The proposed methods for COVID-19 diagnostics are based on the consideration of the high-frequency ranges of FFT spectra. The most reliable result is given by the high-frequency criterion k_4 for the frequency range above 4700 Hz. Other lung illnesses do not cause abnormal respiratory sounds (adventitious sounds) in the considered frequency range; changes caused by them are between 50 and 2500 Hz (see Figure 5 and [8,10,12,28,35]). This fact and the independence of the proposed criteria of the concomitant diseases allow us to assume that the criteria can be used for diagnostics of COVID-19. We analyzed FFT spectra for several patients with other lung diseases (without COVID-19), such as asthma (Figure 5), bilateral pneumonia, pneumonia, and upper respiratory tract infection, and did not find these specific changes in the high-frequency range above 4700 Hz.

Limitations

The proposed screening self-tests would serve as a preliminary step before further procedures are ordered by a doctor. The results of the screening self-tests should be confirmed by other diagnostic methods (chest X-ray/CT and coronavirus tests, such as PCR test, antigen test, and specific SARS-CoV-2 antibody test).

Conclusion

The high-frequency range of the respiratory sound FFT spectrum contains information about the health state of the examined person. The proposed computer-assisted methods based on analysis of this spectrum part can be applied as fast, remote additional screening methods (telemedicine) for specific acoustic phenomena associated with COVID-19. The methods demonstrate sufficiently high diagnostic values. The methods can be a basis for the development of noninvasive screening self-testing kits for COVID-19. To increase the accuracy and

reliability of the methods, a big database of respiratory sounds of patients with COVID-19 and volunteers should be created.

Acknowledgments

This study was supported by a grant from the Ministry of Science & Technology (MOST; 3-16500), Israel & Russian Foundation (RFBR; joint research project 19-515-06001).

Conflicts of Interest

None declared.

References

1. Worldometer. Reported Cases and Deaths by Country or Territory. URL: https://www.worldometers.info/coronavirus/#countries [accessed 2022-05-26]
2. Plotkin S. The new coronavirus, the current king of China. J Pediatric Infect Dis Soc 2020 Feb 28;9(1):1-2 [FREE Full text] [doi: 10.1093/jpids/piaa018] [Medline: 32083284]
3. Sant'Ambrogio G, Sant'Ambrogio FB. Role of laryngeal afferents in cough. Pulm Pharmacol 1996 Oct;9(5-6):309-314. [doi: 10.1006/pulp.1996.0040] [Medline: 9232668]
4. Raj VS, Mou H, Smits SL, Dekkers DHW, Müller MA, Dijkman R, et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature 2013 Mar 14;495(7440):251-254 [FREE Full text] [doi: 10.1038/nature12005] [Medline: 23486063]
5. Wang K, Chen W, Zhou YS, Lian JQ, Zhang Z, Du P, et al. SARS-CoV-2 invades host cells via a novel route: CD147-spike protein. bioRxiv Preprint posted online December 4, 2020. [doi: 10.1101/2020.03.14.988345] [Medline: 33277466]
6. Yang S, Shi Y, Lu H, Xu J, Li F, Qian Z, et al. Clinical and CT features of early stage patients with COVID-19: a retrospective analysis of imported cases in Shanghai, China. Eur Respir J 2020 Apr 26;55(4):2000407 [FREE Full text] [doi: 10.1183/13993003.00407-2020] [Medline: 32217649]
7. Zhou P, Yang XL, Wang X, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 2020 Feb 03;579(7798):270-273. [doi: 10.1038/s41586-020-2012-7] [Medline: 33199918]
8. Ellington LE, Gilman RH, Tielsch JM, Steinhoff M, Figueroa D, Rodriguez S, et al. Computerised lung sound analysis to improve the specificity of paediatric pneumonia diagnosis in resource-poor settings: protocol and methods for an observational study. BMJ Open 2012 Feb 03;2(1):e000506 [FREE Full text] [doi: 10.1136/bmjopen-2011-000506] [Medline: 22307098]
9. Marques A, Oliveira A, Jácome C. Computerized adventitious respiratory sounds as outcome measures for respiratory therapy: a systematic review. Respir Care 2014 May 17;59(5):765-776 [FREE Full text] [doi: 10.4187/respcare.02765] [Medline: 24046460]
10. Mhetre R, Bagal PU. Respiratory sound analysis for diagnostic information. IOSRJEEE 2014;9(5):42-46. [doi: 10.9790/1676-09544246]
11. Fenton TR, Pasterkamp H, Tal A, Chernick V. Automated spectral characterization of wheezing in asthmatic children. IEEE Trans Biomed Eng 1985 Jan;BME-32(1):50-55. [doi: 10.1109/tbme.1985.325616]
12. Reichert S, Gass R, Hajjam A, Brandt C, Nguyen E, Baldassari K, et al. The ASAP project: a first step to an auscultation's school creation. Respir Med CME 2009;2(1):7-14. [doi: 10.1016/j.rmedc.2009.01.001]
13. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA 2020 Apr 07;323(13):1239-1242. [doi: 10.1001/jama.2020.2648] [Medline: 32091533]
14. Pramono RXA, Bowyer S, Rodriguez-Villegas E. Automatic adventitious respiratory sound analysis: a systematic review. PLoS One 2017 May 26;12(5):e0177926 [FREE Full text] [doi: 10.1371/journal.pone.0177926] [Medline: 28552969]
15. Olvera-Montes N, Reyes B, Charleston-Villalobos S, Gonzalez-Camarena R, MejiaAvila M, Dorantes-Mendez G, et al. Detection of respiratory crackle sounds via an Android smartphone-based system. Annu Int Conf IEEE Eng Med Biol Soc 2018;2018:1620-1623. [doi: 10.1109/embc.2018.8512672]
16. Reyes B, Olvera-Montes N, Charleston-Villalobos S, González-Camarena R, Mejía-Ávila M, Aljama-Corrales T. A smartphone-based system for automated bedside detection of crackle sounds in diffuse interstitial pneumonia patients. Sensors (Basel) 2018 Nov 07;18(11):3813 [FREE Full text] [doi: 10.3390/s18113813] [Medline: 30405036]
17. Schuller BW, Schuller DM, Qian K, Liu J, Zheng H, Li X. Front Digit Health 2021 Mar 29;3:564906 [FREE Full text] [doi: 10.3389/fdgth.2021.564906] [Medline: 34713079]
18. Faezipour M, Abuzneid A. Smartphone-based self-testing of COVID-19 using breathing sounds. Telemed J E Health 2020 Oct 01;26(10):1202-1205. [doi: 10.1089/tmj.2020.0114] [Medline: 32487005]
19. Brown C, Chauhan J, Grammenos A. Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data. arXiv Preprint posted online August 23, 2020.. [doi: 10.1145/3394486.3412865]

20. Furman E, Yakovleva E, Malinin S, Furman G, Sokolovsky V. Computer-assisted assay of respiratory sound of children suffering from bronchial asthma. Clin Med 2014;6(1):83-87.
21. Furman E, Yakovleva E, Malinin S, Furman G, Meerovich V, Sokolovsky V. P50: a new modality using breath sound analysis in pediatric asthma. Clin Transl Allergy 2014 Feb 28;4(S1):105. [doi: 10.1186/2045-7022-4-s1-p105]
22. Furman E, Rocheva E, Malinin S, Furman G, Sokolovsky V. Comparative effectiveness of computer analysis of the energy characteristics of the respiratory noise spectrum at three points for the diagnosis of bronchial obstructive syndrome in children with bronchial asthma. Permskii meditsinskii zhurnal 2015:32-88. [doi: 10.1183/13993003.congress-2019.pa739]
23. Furman E, Sokolovsky V, Furman G, Meerovich VM, Malinin SV, Rocheva EV. Mathematical model of breath sound propagation in respiratory tract. Russ J Biomech 2018;22(2):142-152. [doi: 10.15593/RJBiomech/2018.2.03]
24. Malinin S, Furman E, Rocheva E, Sokolovsky V, Furman G. The home remote diagnostics of bronchial asthma in children with the using of telemedical system. Eur Respir J 2019;54:PA739. [doi: 10.1183/13993003.congress-2019.PA739]
25. Furman E, Malinin S, Furman G, Meerovich V, Sokolovsky V, Rocheva E. Respiratory sound analysis for bronchial asthma diagnostics. IOSRJEN 2020;10(1):53-59. [doi: 10.1183/13993003.congress-2019.pa739]
26. Reddel H, FitzGerald J, Bateman E, Bacharier LB, Becker A, Brusselle G, et al. GINA 2019: a fundamental change in asthma management; treatment of asthma with short-acting bronchodilators alone is no longer recommended for adults and adolescents. Eur Respir J 2019 Jun;53(6):1901046 [FREE Full text] [doi: 10.1183/13993003.01046-2019] [Medline: 31249014]
27. Baratloo A, Hosseini M, Negida A, El Ashal G. Part 1: simple definition and calculation of accuracy, sensitivity and specificity. Emergency (Tehran) 2015;3(2):48-49 [FREE Full text] [Medline: 26495380]
28. Sovijärvi A, Dalmasso F, Vanderschoot J, Malmberg L, Righini G, Stoneman SA. Definition of terms for applications of respiratory sounds. Eur Respir Rev 2000;10:577-610.
29. Kosasih K, Abeyratne U, Swarnkar V. High frequency analysis of cough sounds in pediatric patients with respiratory diseases. 2012 Presented at: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 28-Aug-2012 to 1-Sep-2012; San Diego, CA. [doi: 10.1109/embc.2012.6347277]
30. Taplidou S, Hadjileontiadis L, Kitsas I. On applying continuous wavelet transform in wheeze analysis. 2004 Presented at: 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2004; San Francisco, CA. [doi: 10.1109/iembs.2004.1404073]
31. Moussavi Z. Fundamentals of respiratory sounds and analysis. Synth Lect Biomed Eng 2006 Jan;1(1):1-68. [doi: 10.2200/s00054ed1v01y200609bme008]
32. Lai C, Ko W, Lee P, Jean S, Hsueh P. Extra-respiratory manifestations of COVID-19. Int J Antimicrob Agents 2020 Aug;56(2):106024 [FREE Full text] [doi: 10.1016/j.ijantimicag.2020.106024] [Medline: 32450197]
33. Brann D, Tsukahara T, Weinreb C, Lipovsek M, Van den Berge K, Gong B, et al. Non-neuronal expression of SARS-CoV-2 entry genes in the olfactory system suggests mechanisms underlying COVID-19-associated anosmia. Sci Adv 2020 Jul 31;6(31):eabc5801 [FREE Full text] [doi: 10.1126/sciadv.abc5801] [Medline: 32937591]
34. Borczuk AC, Salvatore SP, Seshan SV, Patel SS, Bussel JB, Mostyka M, et al. COVID-19 pulmonary pathology: a multi-institutional autopsy cohort from Italy and New York City. Mod Pathol 2020 Nov 02;33(11):2156-2168 [FREE Full text] [doi: 10.1038/s41379-020-00661-1] [Medline: 32879413]
35. Sovijarvi A, Malmberg L, Charbonneau G, Vandershoot J, Dalmasso F, Sacco C, et al. Characteristics of breath sounds and adventitious respiratory sounds. Eur Respir Rev 2000 Jan;10(77):591-596.

Abbreviations

CA: computer audition
COPD: chronic obstructive pulmonary disease
CT: computed tomography
FFT: fast Fourier transform
MF: moments of the frequency
PCR: polymerase chain reaction

Edited by A Mavragani; submitted 13.06.21; peer-reviewed by Y Man, J op den Buijs; comments to author 13.10.21; revised version received 18.10.21; accepted 11.05.22; published 19.07.22.

Please cite as:
Furman G, Furman E, Charushin A, Eirikh E, Malinin S, Sheludko V, Sokolovsky V, Shtivelman D
Remote Analysis of Respiratory Sounds in Patients With COVID-19: Development of Fast Fourier Transform–Based Computer-Assisted Diagnostic Methods
JMIR Form Res 2022;6(7):e31200
URL: https://formative.jmir.org/2022/7/e31200
doi:10.2196/31200
PMID:35584091

Original Paper

Public Attitudes Regarding Trade-offs Between the Functional Aspects of a Contact-Confirming App for COVID-19 Infection Control and the Benefits to Individuals and Public Health: Cross-sectional Survey

Seiji Bito[1], MD, MSHS; Yachie Hayashi[1], BSc; Takanori Fujita[2], MD, PhD; Shigeto Yonemura[3], MD, MA

[1]Division of Clinical Epidemiology, National Hospital Organization Tokyo Medical Center, Tokyo, Japan

[2]Department of Health Policy Management, Keio University School of Medicine, Tokyo, Japan

[3]The Graduate Schools for Law and Politics, University of Tokyo, Tokyo, Japan

Corresponding Author:
Seiji Bito, MD, MSHS
Division of Clinical Epidemiology
National Hospital Organization Tokyo Medical Center
2-5-1 Higashigaoka
Meguro-ku
Tokyo, 1528602
Japan
Phone: 81 334113130
Email: seiji.bito@kankakuki.jp

Abstract

Background: It is expected that personal health information collected through mobile information terminals will be used to develop health strategies that benefit the public. Against this background, several countries have actively attempted to use mobile phones to control infectious diseases. These collected data, such as activity logs and contact history, are countermeasures against diseases such as COVID-19. In Japan, the Ministry of Health, Labor, and Welfare has developed and disseminated a contact-confirming app (COVID-19 Contact-Confirming Application [COCOA]) to the public, which detects and notifies individuals whether they have been near someone who had subsequently tested positive for COVID-19. However, there are concerns about leakage and misuse of the personal information collected by such information terminals.

Objective: This study aimed to investigate the possible trade-off between effectiveness in preventing infectious diseases and infringement of personal privacy in COCOA. In addition, we analyzed whether resistance to COCOA would reduce if the app contributed to public health or if a discount was provided on mobile phone charges.

Methods: A cross-sectional, quantitative survey of Japanese citizens was conducted using *Survey Monkey*, a general-purpose web-based survey platform. When developing the questions for the questionnaire, we included the installation status of COCOA and recorded the anxiety stemming from the potential leakage or misuse of personal information collected for COVID-19 infection control. The respondents were asked to rate various factors to determine their perceptions on a 5-point scale.

Results: In total, 1058 participants were included in the final analysis. In response to the question of whether the spread of the disease was being controlled by the infection control measures taken by the government, 25.71% (272/1058) of the respondents answered that they strongly agreed or agreed. One-quarter of the respondents indicated that they had already installed COCOA. This study found that the sense of resistance to government intervention was not alleviated by the benefits provided to individuals when using the app. The only factors that were positively associated with the response *absolutely opposed to use* of the app, even with a discount on mobile phone use charges, were those regarding leaks and misuse of personal information, which was true for all functions (function A: odds ratio [OR] 1.8, 95% CI 1.3-2.4; function B: OR 1.9, 95% CI 1.5-2.6; function C: OR 1.8, 95% CI 1.4-2.4).

Conclusions: Public organizations need to emphasize the general benefits of allowing them to manage personal information and assure users that this information is being managed safely rather than offering incentives to individuals to provide such personal information. When collecting and using citizens' health information, it is essential that governments and other entities focus on contributing to the public good and ensuring safety rather than returning benefits to individual citizens.

(JMIR Form Res 2022;6(7):e37720) doi:10.2196/37720

KEYWORDS

internet questionnaire survey; contact-confirming app; COVID-19; privacy; ethics in public health; health application; application development; health service; mobile phone; survey platform; public health; digital information; privacy; health information

Introduction

Background

With the development of information technology, the use of big data has become more common. Indeed, there are many examples where big data have been successfully used to promote general public health by referring to the personal information collected via the data terminals of smartphones, as smartphones are carried around by people everywhere and form a part of their lives [1,2].

Against this background, several countries have actively attempted to use mobile phones for infectious disease control by launching, as well as promoting the download of, apps to collect data, such as activity logs and contact history, as a countermeasure against COVID-19. Each country has unique methods for doing so [3,4]. For example, in South Korea and Singapore, the government has taken the lead in developing and disseminating contact-tracing apps that track the history of individuals who are infected and established the necessary system to rapidly share data with municipalities and medical institutions [5-7].

Although the situation in early 2022—in which unlimited personal information could be uploaded to a specific information cloud or similar means for secondary use—might have been convenient, it involved various ethical, legal, and social problems [8-10]. For example, many contact-tracing apps developed to prevent the spread of COVID-19 have GPS-tracking capabilities, and the information collected by these apps includes detailed movement records of individuals. If this information were to be leaked or used for purposes other than as intended, an individual's privacy would be severely compromised. Moreover, because of the nature of the information, the purpose for which it is to be used must be communicated, along with details about the agencies managing the information and the scope of information use [11]. Wacksman [12] warned that the unprecedented demands stemming from the response to the spread of COVID-19 pose new risks, such as the collection of vast amounts of data and the risk of increasing demands for data sharing that go beyond the intended purpose.

In Japan, the Ministry of Health, Labor, and Welfare has developed and disseminated a contact-confirming app (COVID-19 Contact-Confirming Application [COCOA]) to the public, which notifies individuals whether they have been near an individual who has tested positive for COVID-19. The Ministry of Health, Labor, and Welfare explains the purpose of COCOA on its website as follows:

> *With knowledge of the possibility that an individual was in contact with a "positive" individual, users can receive early support from public health centers, such as by undergoing a test. As the number of users increases, it is hoped that this initiative will help prevent the spread of infection.*

As of April 2022, there have been >32.5 million downloads of COCOA, and it is thought that approximately 1 in 5 people living in Japan have installed the app on their mobile device or devices. In Japan, COCOA is the most widely used contact verification app, and citizens have only 2 choices: installing or not installing COCOA.

COCOA was developed as a Bluetooth-based exposure notification app rather than a contact-tracing app. Thus, current COCOA functions do not have the ability to track moving individuals through GPS or collect personal movement logs. If users who have installed this app on their mobile device come into close contact (within 1 m for ≥15 minutes) with an individual who is COVID-19 positive (positive registration is confirmed by self-report), the Bluetooth communication function detects this and informs the person that they have been in high-risk contact with an individual who is COVID-19 positive.

The information delivered through the notification does not include the location information of the app user who is infected—only a contact code is exchanged when a user contacts another app user's smartphone. Owing to the nature of COCOA's functions, there is very little risk of personal information being shared with government authorities, and the possibility of private information being externally leaked or used for other purposes is also considered to be extremely low [13].

However, whether COCOA functions effectively as an infection control measure against COVID-19 warrants further investigation, as there appear to be 3 reasons for its likely dysfunction. First, the app involves a double opt-in by users in terms of positive result registration through COCOA, which requires both the installation of COCOA on mobile devices and the proactive registration of a positive result (as sharing test results is not mandatory). Consequently, it is assumed that there are many undetected cases for which no alerts are issued, even when a person has actually been in contact with an individual who is COVID-19 positive [13,14]. Second, the app cannot collect data on people's movements, which makes it almost impossible for authorities to intervene directly or implement specific strategies, such as infection control, using the log data. Third, even for an alert indicating such contact, there is no function to suggest specific actions to be taken afterward; thus, the app cannot facilitate effective behavioral changes at the individual level to prevent further infections [13,14]. Several empirical investigations into the attitudes of actual users (citizens) of contact-tracing apps and exposure notifications toward the future use of such apps (regarding the balance between risks related to the leakage of private information, the use of such apps for other purposes, and the benefits to public health and individuals) have been conducted since the start of

the COVID-19 pandemic [15-19]. In a survey aimed at examining why some Australian citizens had not downloaded the app launched by their government, Thomas et al [18] reported that citizens were concerned about the app's technical functionalities and privacy and were distrusting of the government. According to a citizen survey on COCOA conducted by Machida et al [15], the main concerns that inhibited app use were identified as *lack of knowledge on how to use it*, *privacy concerns*, *doubts about the effectiveness of the app*, and *concerns about battery consumption and communication costs*. By contrast, there is a lack of international evidence on citizens' acceptance of apps when there is a trade-off between the app's features and the associated personal and public health benefits.

Objectives

In this study, we conducted a quantitative survey of Japanese citizens to identify the extent to which they were resistant to features that could be implemented in contact-confirming apps. This study investigates the possible trade-off between the effective prevention of infectious diseases and infringements of personal privacy. In addition, we analyzed whether this resistance would decrease if users perceived a contribution in terms of public health or if they received a discount on their mobile phone charges in exchange for app use.

Methods

Study Design and Participants

In this study, we conducted a cross-sectional web-based survey using *Survey Monkey*, a general-purpose web-based survey platform [20]. The target population was set as the general population, aged ≥18 years, living in Japan, and having their own mobile devices. We recruited participants with a target of 1000 respondents. For this, we invited a sample of residents who fit the general target population requirements from among the survey monitors of the contracted research company, Asmarq.co.jp [21]. A sample of 1300 respondents was initially recruited, and if the final number of respondents did not reach 1000, additional samples were added. This process ended when the number of respondents exceeded 1000. The survey was anonymous as the researcher could track neither the respondents nor the nonrespondents. The researcher did not pay any incentives directly to the respondents; however, a sampling fee was paid to the research company. Participants were required to check the details and overview of the survey presented on the website. The actual survey commenced once they provided informed consent. The survey was conducted in June 2021.

Instruments

A set of multiple-choice questions was developed to accomplish the objectives of this study through a cross-sectional survey on the internet. The specific survey items are listed in Multimedia Appendix 1.

As independent variables were assumed to be associated with dependent variables, we asked about the respondents' characteristics, frequency of use of social networking sites (SNSs), COVID-19 infection history, fears related to COVID-19 infection, level of adherence to infection prevention behaviors (behavioral restrictions and mask wearing), and their assessment of government agencies and hospital responsibility for infectious disease prevention.

As dependent variables, we tried to include the installation status of COCOA, as well as record respondents' anxiety regarding potential leakage or misuse of personal information collected for COVID-19 infection control. Respondents were asked about their anxiety on a 5-point scale from *strongly agree* to *strongly disagree*, with *neutral* as their neutral option. In addition, respondents were asked about their degree of resistance to the use of the app if the following 3 functions (that COCOA does not currently have) were added:

- Function A: If a respondent is found to have COVID-19, information on the respondent's infection would be provided to the authorities via a medical institution and reflected in COCOA notifications.
- Function B: In addition to function A, this function would use the location information function of smartphones to track and prompt individuals to take specific actions, such as requesting a medical examination.
- Function C: All movement data of people who have installed the app would be recorded through their mobile devices and aggregated by the government; these data would be used by the government to plan and implement specific infection prevention measures.

The respondents were asked to indicate their level of resistance to the use of the app if the various functions were added on a 5-point scale: *absolutely opposed to use*, *significant resistance to use*, *some resistance to use*, *not much resistance to use*, and *no resistance to use*. Finally, we asked about the resistance to each function when it was understood that the addition of functions A, B, or C would bring about a trade-off in benefits to individuals versus the public. Specifically, participants were asked about their resistance to each feature in the following cases:

- Condition 1: if 50% of the population uses the app with each feature added, the degree of spread of the infection would be halved compared with nonuse of the app.
- Condition 2: if installing the app with each feature resulted in a discount on their monthly mobile phone bill; the options were set as approximately ¥200 (US $1.48) per month, ¥500 (US $3.70) per month, ¥1000 (US $7.29) per month, ¥2000 (US $14.78) per month, ¥3000 (US $22.17) per month, ¥5000 (US $36.95) per month, and would not want to use it irrespective of the discount.

Data Collection

Participants sampled by the survey company visited Survey Monkey's website to take the survey we created. Thereafter, respondents were informed via the web about the purpose of the survey and the advantages and disadvantages of participating; only those who agreed to participate became respondents. The survey ended when the number of respondents exceeded 1000. Responses to the web-based survey were collected in the cloud, and the individual forms were downloaded by the principal investigator.

Statistical Analysis

SPSS (version 27; IBM Corp) was used for statistical analysis. All data entered for the survey, including age, were categorical. Descriptive statistics included descriptions of all independent and dependent variables, including patient characteristics. Logistic regression analysis was performed to analyze the factors that could explain the dependent variables. In performing the logistic regression analysis, we attempted to dichotomize the dependent variables into *0* and *1*. As a criterion for dichotomization, we categorized anxiety regarding leakage or misuse of personal information collected for COVID-19 countermeasures between *agree* and *neither*. In addition, resistance to the proposed newly added functions was dichotomized between *absolutely opposed to use* and *significantly resistant to use*. The former option was dichotomized to determine whether a person had a small amount of anxiety. For the latter, we considered the presence of a clear and strong refusal to use the service as an important threshold for attitude. All statistical tests were 2-tailed, and $P \leq .05$ was set as the threshold for statistical significance.

Ethics Approval

This study was reviewed and approved by the Biomedical Research Ethics Committee of the National Hospital Organization Tokyo Medical Center in May 2021 (approval number R21-030).

Results

Participants' Sociodemographic Information

We recruited 1300 people through a research company, of whom 1191 (91.62%) accessed the survey website. The eligibility criteria were being aged ≥18 years, living in Japan, and having their own mobile phones. Of the 1191 individuals, 1096 (92.02%) met the eligibility criteria, and after excluding those who did not give consent or dropped out, 1058 (88.83%) participants were included in the analysis.

The basic characteristics and frequency of SNS use by the respondents are presented in Table 1. Regarding respondents' age, 8.82% (93/1058) were aged ≤34 years, 71.73% (759/1058) were aged 35 to 64 years, and 19.48% (206/1058) were aged ≥65 years. Regarding gender, 43.37% (459/1058) were female; regarding marital status, 66.6% (705/1058) were married; and regarding the number of people living together, 16.41% (174/1058) lived alone. The proportion of respondents who had personally contracted COVID-19 was 1.12% (12/1058), whereas the proportion of respondents whose family or close friends had been infected was 3.2% (34/1058). The frequency of SNS use (ie, Twitter, Facebook, and Instagram) was as follows (in the order *not used*, *used several times per month or less*, and *used several times per week or more*): 46.34% (487/1058), 20.6% (218/1058), 33.41% (353/1058) for Twitter; 52.32% (553/1058), 26.59% (281/1058), 21.1% (223/1058) for Facebook; and 58.78% (622/1058), 16.91% (179/1058), and 24.3% (257/1058) for Instagram.

Table 1. Respondent demographics (N=1058).

Characteristics and response options	Number of responses, n (%)
Basic characteristics	
Age (years)	
≤34	93 (8.8)
35-64	759 (71.7)
≥65	206 (19.5)
Gender	
Female	459 (43.4)
Male	599 (56.6)
Marital status	
Married	705 (66.6)
Unmarried	353 (33.4)
Households	
Living alone	174 (16.4)
2 people	365 (34.5)
≥3people	519 (49.1)
COVID-19 infections	
Respondent	
No	1028 (98.9)
Yes	11 (1.1)
Family and close friends	
No	995 (96.8)
Yes	33 (3.2)
Use of SNS[a] (frequency)	
Twitter	
Does not use	486 (46)
Several times per month or less	218 (20.6)
Several times per week or more	353 (33.4)
Facebook	
Does not use	552 (52.3)
Several times per month or less	281 (26.6)
Several times per week or more	223 (21.1)
Instagram	
Does not use	620 (58.8)
Several times per month or less	178 (16.9)
Several times per week or more	256 (24.3)
Degree of compliance with preventative behaviors	
Wearing masks when out	
I strictly adhere to this	716 (69.1)
I adhere to this with some exceptions	250 (24.1)
I am neutral to this	51 (4.9)
I do not really adhere to this	9 (0.9)
I do not ever adhere to this	10 (1)

Characteristics and response options	Number of responses, n (%)
Restrictions on activities such as meetings	
I strictly adhere to this	598 (56.9)
I adhere to this with some exceptions	333 (31.7)
I am neutral to this	85 (8.1)
I do not really adhere to this	17 (1.6)
I do not ever adhere to this	18 (1.7)
Assessment of authorities in controlling the spread of infections	
National and prefectural governments	
Strongly agree	55 (5.2)
Agree	217 (20.6)
Neutral	356 (33.8)
Disagree	237 (22.5)
Strongly disagree	188 (17.9)
Local health centers	
Strongly agree	57 (5.4)
Agree	267 (25.5)
Neutral	464 (44.4)
Disagree	162 (15.5)
Strongly disagree	96 (9.2)
Assessment of authorities in controlling the spread of infections	
Medical institutions such as hospitals	
Strongly agree	79 (7.5)
Agree	345 (32.8)
Neutral	456 (43.3)
Disagree	103 (9.8)
Strongly disagree	69 (6.6)
Assessment of measures against infection	
National and prefectural governments	
Strongly agree	61 (5.8)
Agree	269 (25.6)
Neutral	337 (32.1)
Disagree	192 (18.3)
Strongly disagree	191 (18.2)
Local health centers	
Strongly agree	90 (8.6)
Agree	372 (35.7)
Neutral	387 (37.2)
Disagree	121 (11.6)
Strongly disagree	72 (6.9)
Medical institutions like hospitals	
Strongly agree	128 (12.2)
Agree	413 (39.5)
Neutral	374 (35.8)

Characteristics and response options	Number of responses, n (%)
Disagree	75 (7.1)
Strongly disagree	56 (5.4)
Fear of respondent becoming infected	
In terms of health deterioration to oneself	
Strongly agree	378 (35.8)
Agree	437 (41.5)
Neutral	144 (13.6)
Disagree	66 (13.6)
Strongly disagree	30 (6.3)
In terms of transmission to one's family or friends	
Strongly agree	422 (2.8)
Agree	414 (40.1)
Neutral	147 (39.3)
Disagree	39 (14)
Strongly disagree	31 (2.9)
In terms of transmission to other people	
Strongly agree	367 (34.8)
Agree	429 (40.7)
Neutral	176 (16.7)
Disagree	52 (4.9)
Strongly disagree	31 (2.9)
Anxiety regarding the handling of one's personal information	
In terms of data leaks and misuse	
Strongly agree	135 (12.9)
Agree	296 (28.3)
Neutral	386 (36.9)
Disagree	178 (17)
Strongly disagree	51 (4.9)
Use of COCOA[b]	
Have you installed COCOA?	
Yes	259 (24.9)
No	782 (75.1)

[a]SNS: social networking site.

[b]COCOA: COVID-19 Contact-Confirming Application.

Awareness of COVID-19 and Infectious Disease Control

The respondents' attitudes toward COVID-19 and infection control measures are shown in Table 1. Respondents were asked how strictly they adhered to the government's requests to wear a mask when going out and restrict their activities, such as attending meetings in person. Approximately 93.21% (986/1058) and 88.59% (937/1058) of respondents said that they strictly adhered to or adhered to the requests with some exceptions for wearing masks and for restricting activities such as meetings, respectively.

In response to the question of whether the spread of the disease was being controlled by the infection control measures taken by the government and various prefectures, 25.8% (273/1058) of respondents answered that they *strongly agreed* or *agreed*. When asked if the government and prefectures were conducting infection control measures with a strong sense of responsibility, 31.41% (332/1058) of respondents answered that they strongly agreed or agreed. When asked how fearful they were of harm

to their own health, the transmission of the disease to their family and friends, or transmission to others if they were to contract COVID-19, the respondents answered that they either *strongly agree*" or *agree* with these 3 scenarios at 77.28% (818/1058), 79.43% (840/1058), and 75.5% (799/1058), respectively. When asked whether they were concerned that personal information and other data collected as part of the countermeasures against COVID-19 by the national and prefectural governments would be leaked or misused for other purposes, 41.17% (436/1058) of respondents answered that they *strongly agree* or *agree*.

One-quarter of the respondents indicated that they had already installed COCOA. They were given an overview of the functions of the current COCOA. Once each of the 3 aforementioned functions (A, B, and C) were added, the percentage of respondents who said that they would be absolutely opposed to its use or had significant resistance to its use was 27.31% (289/1058) for function A and 31.4% (332/1058) for function B. The difference compared with when only function A was added was not significant. With the addition of function C, this percentage was 33.63% (355/1058).

Respondents' Resistance to the Additional Functions, Even With Benefits and Trade-offs

The percentage of respondents who said they were *absolutely opposed to use* of the app even if half of the population used it with functions A, B, and C (enabling the degree of spread to be reduced to half of what it would be if the app were not used) was 11.6% (123/1058) for function A, 13.33% (141/1058) for function B, and 13.71% (145/1058) for function C. The percentage of respondents who said that they were *absolutely opposed to use* of the app, even if the effect of using it was to halve the spread of infections, was 10.36% (110/1058) for function A, 12.21% (129/1058) for function B, and 12.43% (131/1058) for function C. In response to the question "Would you accept the use of the app with each of the functions added, if the following reductions in mobile phone usage charges were applied?" the respondents who selected "I would not want to use the app irrespective of the discount" was 34.62% (366/1058) for function A, 36.77% (389/1058) for function B, and 37.3% (395/1058) for function C (Table 2).

Table 2. Resistance to using the COVID-19 Contact-Confirming Application under various conditions when functions are added to the current version (N=1058).

Condition and response option	Function A (responses), n (%)	Function B (responses), n (%)	Function C (responses), n (%)
Feature addition only			
No resistance to use	141 (13.5)	123 (11.9)	124 (12)
Not much resistance to use	282 (27.1)	251 (24.2)	243 (23.4)
Some resistance to use	334 (32.1)	337 (32.5)	321 (31)
Significant resistance to use	164 (15.7)	187 (18.1)	206 (19.9)
Absolutely opposed to use	121 (11.6)	138 (13.3)	142 (13.7)
A function is added, thereby halving the spread of infections			
No resistance to use	140 (13.4)	118 (11.3)	120 (11.5)
Not much resistance to use	304 (29.1)	259 (24.9)	256 (24.7)
Some resistance to use	339 (32.5)	369 (35.5)	354 (34.1)
Significant resistance to use	152 (14.6)	168 (16.1)	180 (17.3)
Absolutely opposed to use	108 (10.4)	127 (12.2)	129 (12.4)
Feature additions mean reduced charges for device use			
Would accept at a discount of about ¥200 (US $1.48) a month	112 (10.7)	83 (8.0)	88 (8.5)
Would accept at a discount of about ¥500 (US $3.70) per month	147 (14.0)	112 (10.7)	97 (9.3)
Would accept at a discount of about ¥1000 (US $7.29) per month	154 (14.7)	150 (14.4)	147 (14.1)
Would accept at a discount of about ¥2000 (US $14.78) per month	87 (8.3)	123 (11.8)	110 (10.6)
Would accept at a discount of about ¥3000 (US $22.17) per month	52 (5)	53 (5.1)	65 (6.3)
Would accept at a discount of about ¥5000 (US $36.95) per month	133 (12.7)	138 (13.2)	145 (13.9)
I would not want to use it regardless of the discount	362 (34.6)	384 (36.8)	388 (337.3)

Factors Associated With the Dependent Variables

The results of the logistic regression of the relationship between those who installed COCOA and the basic characteristics of the respondents are shown in Table 3. There was no significant difference in terms of gender or age groups; regarding the frequency of SNS use, those who used Facebook at least several times per week were more likely to have COCOA installed (odds ratio [OR] 1.5, 95% CI 1.0-2.2). Although there was no significant difference in terms of their own or family members' history of COVID-19 infection, those who answered *strongly agree* or *agree* to the question of whether they feared they might infect others if they contracted COVID-19 were more likely to have installed COCOA (OR 2.2, 95% CI 1.1-4.7). In addition, those who reported adhering to government-mandated behavioral guidelines for infection control were more likely to have installed the app (OR 2.0, 95% CI 1.0-4.0).

The association between the response *absolutely opposed to use* of the contact-confirming app with added functions A, B, and C and the independent variables of the respondents are provided in Table 4. The respondents' basic characteristics, frequency of SNS use, history of COVID-19 infection, fear of the consequences of being infected themselves, and the degree of adherence to infection control behaviors were not significantly associated with the dependent variables for any of the functions A, B, and C. However, when all functions were added, those who were concerned about leaks or misuse of their personal information were more likely to respond that they were *absolutely opposed to use* of a contact app with more functions added than those who were not concerned (function A: OR 2.1, 95% CI 1.4-3.3; function B: OR 2.3, 95% CI 1.5-3.5; function C: OR 2.5, 95% CI 1.7-3.7). For function B, there was a negative correlation between the answers of *strongly agree* and *agree* in response to the question of whether the national and prefectural governments were controlling the spread of infections with a strong sense of responsibility and the response of *absolutely opposed to use* with the new functions (OR 0.5, 95% CI 0.2-0.9).

Table 3. Factors related to COVID-19 Contact-Confirming Application installation.

Independent variables	Odds ratio (95% CI)
Female (reference: male)	0.7 (0.5-1.0)
Age (years) group (reference: aged <34 years)	
35-64	1.0(0.6-1.8)
≥65	0.8 (0.4-1.6)
Living alone (reference: living with others)	1.1 (0.7-1.6)
SNS[a] accessed occasionally per week (reference: accessed below this amount)	
Twitter	1.1 (0.8-1.6)
Facebook	1.5 (1.0-2.2)
Instagram	1.3 (0.9-1.9)
COVID-19 infection history (reference: no history)	
Respondent has a history of COVID-19	1.3 (0.6-2.6)
Respondent's family has a history of COVID-19	1.2 (0.9-1.6)
Respondent fears COVID-19 infection (reference: disagree)	
Fear of health deterioration to oneself	1.0 (0.6-1.7)
Fear of transmission to one's family and friends	0.6 (0.3-1.3)
Fear of transmission to people other than family members or friends	2.2 (1.1-4.7)
Compliance with preventative behaviors (reference: I do not do this)	
Wearing masks when going out	0.7 (0.3-1.6)
Restrictions on activities such as meetings	2.0 (1.0-4.0)
Assessment of authorities (reference: disagree)	
Infection is being controlled through infection countermeasures conducted by authorities	1.3 (0.8-2.0)
Authorities are conducting infection control with a sense of responsibility	0.9 (0.6-1.5)
I am concerned that my personal information collected by authorities might be leaked or misused for other purposes	0.9 (0.7-1.2)

[a]SNS: social networking site.

Table 4. Factors associated with the response "I am absolutely opposed to using the app" if functions A to C are added.

Independent variables	Function A, odds ratio (95% CI)	Function B, odds ratio (95% CI)	Function C, odds ratio (95% CI)
Female (reference: male)	0.7 (0.5-1.2)	0.8 (0.5-1.2)	0.7 (0.5-1.1)
Age (years) group (reference: aged <34 years)			
35-64	0.9 (0.4-1.8)	1.0 (0.5-2.0)	1.0 (0.5-2.0)
≥65	0.7 (0.3-1.7)	0.8 (0.3-1.9)	1.0 (0.4-2.2)
Living alone (reference: living with others)	0.9 (0.5-1.6)	0.8 (0.5-1.4)	0.9 (0.5-1.5)
SNS[a] accessed occasionally per week (reference: accessed below this amount)			
Twitter	1.0 (0.6-1.6)	1.0 (0.6-1.7)	1.0 (0.6-1.6)
Facebook	0.8 (0.5-1.5)	0.7 (0.4-1.2)	0.7 (0.4-1.2)
Instagram	1.2 (0.7-2.2)	1.3 (0.7-2.2)	1.1 (0.7-1.9)
COVID-19 infection history (reference: no history)			
Respondent has a history of COVID-19	1.5 (0.7-3.4)	1.5 (0.7-3.4)	1.2 (0.5-3.0)
Respondent's family has a history of COVID-19	0.9 (0.6-1.6)	0.9 (0.5-1.5)	0.8 (0.5-1.4)
Respondent fears COVID-19 infection (reference: disagree)			
Fear of health deterioration to oneself	0.9 (0.5-1.8)	0.9 (0.5-1.7)	0.8 (0.4-1.6)
Fear of transmission to one's family and friends	0.7 (0.3-1.7)	0.8 (0.3-2.0)	0.7 (0.3-1.6)
Fear of transmission to people other than family members or friends	0.9 (0.4-2.2)	0.7 (0.3-1.6)	1.0 (0.4-2.3)
Compliance with preventative behaviors (reference: I do not do this)			
Wearing masks when going out	0.8 (0.3-1.8)	0.8 (0.3-1.8)	0.9 (0.4-2.0)
Restrictions on activities such as meetings	0.5 (0.3-1.0)	0.5 (0.3-1.0)	0.5 (0.3-1.0)
Assessment of authorities (reference: disagree)			
Infection is being controlled through countermeasures conducted by authorities	1.2 (0.6-2.4)	1.3 (0.6-2.6)	1.0 (0.5-2.0)
Authorities are controlling the infection with a sense of responsibility	0.6 (0.3-1.1)	0.5 (0.2-0.9)	0.6 (0.3-1.2)
I am concerned that my personal information collected by the authorities might be leaked or misused for other purposes	2.1 (1.4-3.3)	2.3 (1.5-3.5)	2.5 (1.7-3.7)

[a]SNS: social networking site.

We also identified the factors associated with the response *absolutely opposed to use*, even when users were presented with trade-offs between privacy on the one hand and public health benefits from each feature's addition and the personal benefits of the feature additions in the form of discounted mobile phone use charges on the other hand (Tables 5 and 6, respectively). Even if the addition of each function reduced the spread of infection by half, the factors that were significantly associated with the response *absolutely opposed to use* of the app were the same for all functions (A, B, and C). One of the factors was positively correlated, and two were negatively correlated. The factor that was positively correlated was the respondents who were concerned about data leaks and misuse of personal information (function A: OR 2.2, 95% CI 1.4-3.5; function B: OR 2.2, 95% CI 1.4-3.3; function C: OR 2.5, 95% CI 1.6-3.8). Conversely, the factors that presented negative correlations were adherence to the government-imposed restrictions on activities (function A: OR 0.4, 95% CI 0.2-0.8; function B: OR 0.4, 95% CI 0.2-0.7; function C: OR 0.4, 95% CI 0.2-0.8) and agreement that the national and local governments were conducting infection control with a sense of responsibility (function A: OR 0.4, 95% CI 0.2-0.8; function B: OR 0.5, 95% CI 0.2-1.0; function C: OR 0.5, 95% CI 0.2-0.9).

The only factor that was positively associated with the response *absolutely opposed to use*, even if there was a discount on mobile phone use charges, was concern about leaks and misuse of the collected personal information, which was true for all functions (function A: OR 1.8, 95% CI 1.3-2.4; function B: OR 1.9, 95% CI 1.5-2.6; function C: OR 1.8, 95% CI 1.4-2.4). In contrast, adhering to government restrictions on activities had a negative correlation with the dependent variables for all functional additions (function A: OR 0.5, 95% CI 0.3-0.8; function B: OR 0.5, 95% CI 0.3-0.9; function C OR 0.6, 95% CI 0.3-0.9). Fear of transmission to family members if respondents became infected was negatively correlated with functions A and C (function A: OR 0.5, 95% CI 0.2-0.9; function C: OR 0.5, 95% CI 0.3-0.9).

Table 5. Factors associated with the response "I am absolutely opposed to the use of the app" if it is assumed that the spread of infections will be halved by the addition of functions A to C.

Independent variables	Function A, odds ratio (95% CI)	Function B, odds ratio (95% CI)	Function C, odds ratio (95% CI)
Female (reference: male)	0.6 (0.4-1.0)	0.7 (0.5-1.1)	0.7 (0.4-1.0)
Age (years) group (reference: aged <34 years)			
35-64	1.0 (0.4-2.1)	0.8 (0.4-1.5)	0.8 (0.4-1.5)
≥65	1.0 (0.4-2.6)	0.7 (0.3-1.6)	0.8 (0.4-1.9)
Living alone (reference: living with others)	1.0 (0.6-1.8)	0.9 (0.5-1.6)	1.0 (0.6-1.7)
SNS[a] accessed occasionally per week (reference: accessed below this amount)			
Twitter	0.9 (0.5-1.6)	0.8 (0.5-1.4)	0.9 (0.5-1.5)
Facebook	1.0 (0.5-1.7)	0.8 (0.5-1.5)	0.8 (0.5-1.4)
Instagram	1.4 (0.8-2.6)	1.5 (0.8-2.6)	1.5 (0.9-2.7)
COVID-19 infection history (reference: no history)			
Respondent has a history of COVID-19	1.6 (0.6-4.4)	1.5 (0.6-4.1)	1.6 (0.7-3.6)
Respondent's family has a history of COVID-19	0.6 (0.2-1.4)	0.6 (0.2-1.3)	0.9 (0.5-1.5)
Respondent fears COVID-19 infection (reference: disagree)			
Fear of health deterioration to oneself	0.8 (0.4-1.7)	1.0 (0.5-1.9)	1.0 (0.5-1.9)
Fear of transmission to one's family and friends	0.4 (0.1-1.1)	0.5 (0.2-1.3)	0.5 (0.2-1.3)
Fear of transmission to people other than family members or friends	1.3 (0.5-3.8)	1.0 (0.4-2.5)	1.0 (0.4-2.6)
Compliance with preventative behaviors (reference: I do not do this)			
Wearing masks when going out	1.1 (0.5-2.7)	1.2 (0.5-2.8)	1.0 (0.4-2.4)
Restrictions on activities such as meetings	0.4 (0.2-0.8)	0.4 (0.2-0.7)	0.4 (0.2-0.8)
Assessment of authorities (reference: disagree)			
Infection is being controlled through infection countermeasures conducted by authorities	1.6 (0.7-3.5)	1.2 (0.6-2.5)	1.2 (0.6-2.5)
Authorities control infections with a sense of responsibility	0.4 (0.2-0.8)	0.5 (0.2-1.0)	0.5 (0.2-0.9)
I am concerned that my personal information collected by authorities might be leaked or misused for other purposes	2.2 (1.4-3.5)	2.2 (1.4-3.3)	2.5 (1.6-3.8)

[a]SNS: social networking site.

Table 6. Factors associated with the response "I would not want to use the app" when a fee reduction is presented as a trade-off for additional functions A to C.

Independent variables	Function A, odds ratio (95% CI)	Function B, odds ratio (95% CI)	Function C, odds ratio (95% CI)
Female (reference: male)	1.2 (1.0-1.7)	1.2 (0.9-1.6)	1.2 (0.9-1.6)
Age (years) group (reference: aged <34 years)			
35-64 years	0.9 (0.5-1.4)	0.8 (0.5-1.3)	0.9 (0.5-1.4)
≥65 years	0.7 (0.4-1.4)	0.7 (0.4-1.2)	0.8 (0.4-1.5)
Living alone (reference: living with others)	1.1 (0.7-1.5)	1.1 (0.8-1.6)	1.1 (0.7-1.6)
SNS[a] accessed occasionally per week (reference: accessed below this amount)			
Twitter	0.9 (0.6-1.2)	0.9 (0.6-1.3)	1.0 (0.7-1.4)
Facebook	0.9 (0.6-1.4)	0.9 (0.6-1.3)	0.8 (0.5-1.1)
Instagram	0.7 (0.4-1.0)	0.8 (0.5-1.1)	0.8 (0.5-1.1)
COVID-19 infection history (reference: no history)			
Respondent has a history of COVID-19	0.9 (0.4-1.9)	0.8 (0.4-1.8)	0.9 (0.4-2.0)
Respondent's family has a history of COVID-19	0.9 (0.6-1.3)	0.9 (0.6-1.3)	0.9 (0.6-1.2)
Respondent fears COVID-19 infection (reference: disagree)			
Fear of health deterioration to oneself	1.1 (0.7-1.8)	1.0 (0.6-1.6)	1.1 (0.7-1.8)
Fear of transmission to one's family and friends	0.5 (0.2-0.9)	0.7 (0.3-1.2)	0.5 (0.3-0.9)
Fear of transmission to people other than family members or friends	0.9 (0.5-1.7)	0.8 (0.4-1.4)	1.0 (0.5-1.7)
Compliance with preventative behaviors (reference: I do not do this)			
Wearing masks when going out	0.9 (0.5-1.9)	0.9 (0.5-1.8)	1.0 (0.5-1.9)
Restrictions on activities such as meetings	0.5 (0.3-0.8)	0.5 (0.3-0.9)	0.6 (0.3-0.9)
Assessment of authorities (reference: disagree)			
Infection is being controlled through infection countermeasures conducted by authorities	0.9 (0.6-1.5)	0.9 (0.6-1.4)	0.8 (0.5-1.3)
Authorities control infections with a sense of responsibility	0.7 (0.5-1.1)	0.7 (0.5-1.1)	0.7 (0.5-1.1)
I am concerned that my personal information collected by authorities might be leaked or misused for other purposes	1.8 (1.3-2.4)	1.9 (1.5-2.6)	1.8 (1.4-2.4)

[a]SNS: social networking site.

Discussion

Principal Findings

In formulating the hypotheses for this study, we were interested in the public's attitude toward privacy risks as an inadvertent side effect of the convenience offered by information technology. As Yuan et al [22] suggested, privacy risk strategies are socially implemented after taking into account 2 types of trade-offs that affect individuals' disclosure behavior: those between the expected benefits of privacy disclosure and the effectiveness of risk-management methods. The contact-tracing app or contact notification app that was developed and used in many countries to prevent the spread of the COVID-19 pandemic epitomizes this theory. The contact notification app (COCOA) in operation in Japan was developed to minimize privacy risks to the public; however, its effectiveness in preventing infection was lower than expected. Therefore, the addition of functions that cannot be performed by the current COCOA, such as the collection of mobile logs and secondary use of the collected data, raises concerns about the increased risk to privacy and the functions themselves. The core objective of our research was to identify how the general public perceives the trade-offs between additional personal and public interest. We also wanted to investigate the extent to which people trade personal privacy risks for economic incentives, as noted by Hann et al [23]. This study was conducted on the assumption that when the administrative body of a country plans to develop and disseminate a contact-confirming app that is more effective in preventing the spread of infectious diseases, citizens who use the app may have some resistance to the government's collection of activity logs and provision of individual interventions for people with confirmed infections. We also hypothesized that this resistance might be mitigated by emphasizing the public health benefits of an effective contact-confirming app or through a discount on mobile phone use charges for individuals using the app [24,25].

The results of the logistic regression of factors associated with COCOA installation showed a positive correlation with fear of

transmission of infection to others in the event of oneself becoming infected; a reasonable interpretation of this result is possible. However, the positive correlation was not significant; rather, the proportion of citizens living in Japan who feel obligated to install COCOA, or who have a high assessment of the app's usefulness, is considered low.

The survey then introduced functions A, B, and C, which are expected to increase the risk of personal privacy violations concerning the current COCOA, and asked about the resistance to using the app when each of these functions was added. Approximately one-third of all respondents for functions A, B, and C indicated that they felt uncomfortable using the app. This result was more generous than expected at the beginning of the survey. Our initial assumption was that more than half of the survey respondents would react with at least moderate resistance to each of the additional functions. This result would suggest that the public is likely to be more receptive to the use of privacy-related information by the authorities for maintaining public health than we had expected.

Even when function B (ie, location tracking and specific encouragement of activities) were added to function A, there was no significant difference, suggesting that the additional fear and resistance to individual interventions are not necessarily significant. Similarly, there was no significant difference for function C, indicating that the resistance to the collection of personal information for infection prevention measures was not particularly high.

The percentage of respondents who answered *absolutely opposed to use* or *significant resistance to use* of the app decreased slightly when the addition of functions A, B, and C was known to reduce the degree of spread of infectious disease. This finding suggests that increased public health benefits can be traded against personal privacy risks. This result may be more characteristic of the trade-offs of interests in the somewhat special domain of health rather than the general market society principle [26,27]. Surprisingly, when presented with the condition that the addition of a feature would result in a discount on the cell phone bill, one-third of respondents stated that "I would not want to use it regardless of the discount" for all functions A, B, and C. Although it is difficult to explain this result via a literature review, we argue that this result was influenced by the health-related agenda and the fact that the project was a national government agency. When personal privacy is given up, the motivation behind it may not be to make money but to work together to help prevent the spread of infectious diseases. If so, we believe that bringing in financial incentives when promoting such behavior might inadvertently upset citizens' sentiments.

We also deem our interpretation of the results presented in Tables 4-6 to be relevant. Resistance to using apps was not associated with respondents' basic characteristics, an affinity for SNS, or history of COVID-19 but was negatively correlated with their evaluation of the executive branch regarding infection prevention activities and compliance concerning infection prevention behaviors. We interpreted this result as suggesting that the behavior of using exposure notification apps has an element of altruistic behavior that seeks the public good [28,29]. In contrast, as mentioned many times in previous literature, the sense of risk regarding information leakage had the highest contribution rate among the factors related to resistance to app use [16-18,30-32].

Generalizability

It may be acceptable for society to allow public organizations, such as national government agencies, to collect and use individuals' private information for specific purposes as part of nationwide efforts to control infectious diseases. However, such interventions must be balanced by the resistance felt by the public regarding the provision of private information [33]. In addition, it is necessary to have a continuous discussion between administrative agencies and the public regarding the conditions that would enable the formation of a social consensus on the collection and use of private information for administrative purposes. In this study, we presented a hypothetical basis for this discussion. The current COCOA in Japan is one of the weakest interventions for the collection and use of private information; however, the public response is that certain prerequisites must be in place, such as a clear purpose of use and a robust security environment. If these factors are satisfactorily explained to the public, the intervention will be stronger; however, even if this occurs, the possibility that the intervention will be accepted is still low [34]. For example, it may be possible to actively consider the collection of activity logs using the GPS function on people's phones [35,36]. However, these activities are tolerated by the public as they are aware of the greater public health risks posed by a global crisis such as the COVID-19 pandemic; a more cautious attitude is necessary when considering other general administration purposes [37,38].

An important interpretation to note is that there is no trade-off between the permissibility of government interventions and the provision of special incentives to individuals, such as discounts on mobile phone bills. This study revealed that such simplistic benefits may upset public sentiments and increase suspicion of the government. In a future society where personal health records are widely distributed, public organizations will increasingly require access to private information on the cloud. In such a case, it is more important for public organizations to increase the general benefits of managing personal information and provide assurance that this is being done safely rather than providing incentives to individuals who provide such personal information.

Study Limitations

This survey involved a representative sample of the Japanese population. The response rate was >10/13. The results could potentially be a decision resource for the government when considering whether to add more functions to the app.

However, 4 limitations must be noted. First, it is possible that the public did not understand the functions of the current COCOA. The questions asked in the survey assumed that additional functions would be added to the current COCOA, which may have been difficult for respondents to envision, along with the advantages and disadvantages that would result for themselves and the public. Second, as this was a quantitative

survey based on a specific hypothesis, it was not possible to set out the specific functions that the public would expect from COCOA. A survey with a qualitative approach would be necessary to explore the functions. Third, we were unable to examine the psychometric properties of the attitudes measured in this study. Usually, when attitudes are measured using questionnaires, their reliability and validity must be ensured. However, in this study, the lack of existing general-purpose scales and insufficient time for scale development created major limitations in interpreting the results. The existing literature identifies the main element of the public's resistance to the use of contact-tracing apps as a feeling of insecurity regarding the leakage of information or its use for purposes other than as intended [39,40]. As such an attitude is a very important concept for the future *Ethical, Legal, and Social Issues in Science and Technology* agenda in information technology, it may be necessary to develop a common rating scale with guaranteed psychometric validity [41,42]. Fourth, when interpreting the results of the survey, the impact of the *privacy paradox* along with clinical relevance must be considered. The privacy paradox is a situation in which people express concerns about privacy through their attitudes but do not hesitate to take actions that pose a significant risk in terms of revealing their private information in real life [43,44]. To avoid this paradox, we asked about respondents' attitudes regarding their actual behaviors rather than their conceptual concerns. However, when reflecting on the results of this study for policy recommendations, it must be stated that the evidence is influenced by this paradox.

Conclusions

In this study, we investigated the public's attitude toward the addition of several functions to COCOA, an app used to confirm COVID-19 contact distributed by the Japanese government. Although there was some resistance to its official and governmental release and use, there was little change in the public's resistance, as depicted in this study, regardless of the added functions' content. However, the sense of resistance to the government's intervention was not ameliorated by the incentives provided to individuals. On the basis of the results of this study, we believe that it is acceptable for citizens to implement functions that involve a certain degree of privacy risk when digitally processed personal privacy information is used primarily to promote health benefits, such as preventing the spread of infectious diseases, provided that the entities in charge of such projects, such as government agencies, are fully accountable for their actions. Furthermore, at least with respect to projects such as health policy, citizens will set aside their privacy concerns in view of broader public health interests rather than individual economic incentives. Thus, this study offers insights into potential strategies for governmental use of private information through mobile devices. Further social research and empirical evidence are needed on this topic to form an *ecosystem* of information and people, which is still expanding daily.

Acknowledgments

This research was part of this study, *Comprehensive Practical Research and Development Program to Address Ethical, Legal, and Social Issues in Science and Technology*, and it received funding from Research Institute of Science and Technology for Society, which is an initiative of the Japan Science and Technology Agency to promote research and development. The authors would like to thank Editage for their writing support in the *Abstract* section.

Authors' Contributions

SB was responsible for conceiving and designing the experiments, collecting and analyzing the data, and writing and revising the manuscript. SB, YH, and TF were responsible for data interpretation. All authors have read and approved the final version of the manuscript. All authors have agreed to its publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Specific questions on the survey.
[DOCX File , 29 KB - formative_v6i7e37720_app1.docx]

References

1. Paydar S, Emami H, Asadi F, Moghaddasi H, Hosseini A. Functions and outcomes of personal health records for patients with chronic diseases: a systematic review. Perspect Health Inf Manag 2021 Mar 15;18(Spring):1l [FREE Full text] [Medline: 34345228]
2. Kim JW, Ryu B, Cho S, Heo E, Kim Y, Lee J, et al. Impact of personal health records and wearables on health outcomes and patient response: three-arm randomized controlled trial. JMIR Mhealth Uhealth 2019 Jan 04;7(1):e12070 [FREE Full text] [doi: 10.2196/12070] [Medline: 30609978]
3. O'Connell J, O'Keeffe DT. Contact tracing for Covid-19 - a digital inoculation against future pandemics. N Engl J Med 2021 Aug 05;385(6):484-487. [doi: 10.1056/NEJMp2102256] [Medline: 34010529]

4. Blasimme A, Ferretti A, Vayena E. Digital contact tracing against COVID-19 in Europe: current features and ongoing developments. Front Digit Health 2021 Jun 17;3:660823 [FREE Full text] [doi: 10.3389/fdgth.2021.660823] [Medline: 34713135]
5. Park S, Choi GJ, Ko H. Information technology-based tracing strategy in response to COVID-19 in South Korea-privacy controversies. JAMA 2020 Jun 02;323(21):2129-2130. [doi: 10.1001/jama.2020.6602] [Medline: 32324202]
6. Park Y, Huh IS, Lee J, Kang CR, Cho SI, Ham HJ, Seoul Metropolitan Government COVID-19 Rapid Response (SCoRR) Team. Application of testing-tracing-treatment strategy in response to the COVID-19 outbreak in Seoul, Korea. J Korean Med Sci 2020 Nov 23;35(45):e396 [FREE Full text] [doi: 10.3346/jkms.2020.35.e396] [Medline: 33230987]
7. Lai SH, Tang CQ, Kurup A, Thevendran G. The experience of contact tracing in Singapore in the control of COVID-19: highlighting the use of digital technology. Int Orthop 2021 Jan;45(1):65-69 [FREE Full text] [doi: 10.1007/s00264-020-04646-2] [Medline: 33188602]
8. Cushman R, Froomkin AM, Cava A, Abril P, Goodman KW. Ethical, legal and social issues for personal health records and applications. J Biomed Inform 2010 Oct;43(5 Suppl):S51-S55 [FREE Full text] [doi: 10.1016/j.jbi.2010.05.003] [Medline: 20937485]
9. Gostin L. Health care information and the protection of personal privacy: ethical and legal considerations. Ann Intern Med 1997 Oct 15;127(8 Pt 2):683-690. [doi: 10.7326/0003-4819-127-8_part_2-199710151-00050] [Medline: 9382380]
10. Haynes CL, Cook GA, Jones MA. Legal and ethical considerations in processing patient-identifiable data without patient consent: lessons learnt from developing a disease register. J Med Ethics 2007 May;33(5):302-307 [FREE Full text] [doi: 10.1136/jme.2006.016907] [Medline: 17470509]
11. Carrión Señor I, Fernández-Alemán JL, Toval A. Are personal health records safe? A review of free Web-accessible personal health record privacy policies. J Med Internet Res 2012 Aug 23;14(4):e114 [FREE Full text] [doi: 10.2196/jmir.1904] [Medline: 22917868]
12. Wacksman J. Digitalization of contact tracing: balancing data privacy with public health benefit. Ethics Inf Technol 2021;23(4):855-861 [FREE Full text] [doi: 10.1007/s10676-021-09601-2] [Medline: 34131391]
13. Nakamoto I, Jiang M, Zhang J, Zhuang W, Guo Y, Jin MH, et al. Evaluation of the design and implementation of a peer-to-peer COVID-19 contact tracing mobile app (COCOA) in Japan. JMIR Mhealth Uhealth 2020 Dec 01;8(12):e22098 [FREE Full text] [doi: 10.2196/22098] [Medline: 33170801]
14. Kurita J, Sugawara T, Ohkusa Y. Effectiveness of COCOA, a COVID-19 contact notification application, in Japan. medRxiv 2020 Jul 14. [doi: 10.1101/2020.07.11.20151597]
15. Machida M, Nakamura I, Saito R, Nakaya T, Hanibuchi T, Takamiya T, et al. Survey on usage and concerns of a COVID-19 contact tracing application in Japan. Public Health Pract (Oxf) 2021 Nov;2:100125 [FREE Full text] [doi: 10.1016/j.puhip.2021.100125] [Medline: 34841372]
16. O'Callaghan ME, Buckley J, Fitzgerald B, Johnson K, Laffey J, McNicholas B, et al. A national survey of attitudes to COVID-19 digital contact tracing in the Republic of Ireland. Ir J Med Sci 2021 Aug;190(3):863-887 [FREE Full text] [doi: 10.1007/s11845-020-02389-y] [Medline: 33063226]
17. Smoll NR, Walker J, Khandaker G. The barriers and enablers to downloading the COVIDSafe app - a topic modelling analysis. Aust N Z J Public Health 2021 Aug;45(4):344-347 [FREE Full text] [doi: 10.1111/1753-6405.13119] [Medline: 33970555]
18. Thomas R, Michaleff ZA, Greenwood H, Abukmail E, Glasziou P. Concerns and misconceptions about the Australian government's COVIDSafe app: cross-sectional survey study. JMIR Public Health Surveill 2020 Nov 04;6(4):e23081 [FREE Full text] [doi: 10.2196/23081] [Medline: 33048826]
19. Lang R, Benham JL, Atabati O, Hollis A, Tombe T, Shaffer B, et al. Attitudes, behaviours and barriers to public health measures for COVID-19: a survey to inform public health messaging. BMC Public Health 2021 Apr 21;21(1):765 [FREE Full text] [doi: 10.1186/s12889-021-10790-0] [Medline: 33882896]
20. SurveyMonkey. URL: https://jp.surveymonkey.com/ [accessed 2022-04-22]
21. Asmarq. URL: https://www.asmarq.co.jp/ [accessed 2022-04-22]
22. Li Y. Theories in online information privacy research: a critical review and an integrated framework. Decis Support Syst 2012 Dec;54(1):471-481. [doi: 10.1016/j.dss.2012.06.010]
23. Hann IH, Hui KL, Lee SY, Png IP. Overcoming online information privacy concerns: an information-processing theory approach. J Manag Inf Syst 2007 Dec 08;24(2):13-42. [doi: 10.2753/mis0742-1222240202]
24. Mukherjee S, Manjaly JA, Nargundkar M. Money makes you reveal more: consequences of monetary cues on preferential disclosure of personal information. Front Psychol 2013 Nov 11;4:839 [FREE Full text] [doi: 10.3389/fpsyg.2013.00839] [Medline: 24273524]
25. Yuchao W, Ying Z, Liao Z. Health privacy information self-disclosure in online health community. Front Public Health 2020 Feb 4;8:602792 [FREE Full text] [doi: 10.3389/fpubh.2020.602792] [Medline: 33614566]
26. Dinev T, Hart P. Privacy concerns and internet use--a model of trade-off factors. Acad Manag Proc 2003 Aug;2003(1):D1-D6. [doi: 10.5465/ambpp.2003.13792464]
27. Caudill EM, Murphy PE. Consumer online privacy: legal and ethical issues. J Public Policy Mark 2000 Apr 1;19(1):7-19. [doi: 10.1509/jppm.19.1.7.16951]

28. Li T, Cobb C, Yang JJ, Baviskar S, Agarwal Y, Li B, et al. What makes people install a COVID-19 contact-tracing app? Understanding the influence of app design and individual difference on contact-tracing app adoption intention. Pervasive Mob Comput 2021 Aug;75:101439. [doi: 10.1016/j.pmcj.2021.101439]
29. Williams SN, Armitage CJ, Tampe T, Dienes K. Public attitudes towards COVID-19 contact tracing apps: a UK-based focus group study. Health Expect 2021 Apr;24(2):377-385 [FREE Full text] [doi: 10.1111/hex.13179] [Medline: 33434404]
30. Ernst CP. Risk hurts fun: the influence of perceived privacy risk on social network site usage. In: Ernst CP, editor. Factors Driving Social Network Site Usage. Wiesbaden, Germany: Springer; 2015:45-56.
31. Dyke SO, Dove ES, Knoppers BM. Sharing health-related data: a privacy test? NPJ Genom Med 2016 Aug 17;1(1):160241-160246 [FREE Full text] [doi: 10.1038/npjgenmed.2016.24] [Medline: 27990299]
32. Cho H, Ippolito D, Yu YW. Contact tracing mobile apps for COVID-19: privacy considerations and related trade-offs. arXiv 2020 Mar 30.
33. White L, van Basshuysen P. Privacy versus public health? A reassessment of centralised and decentralised digital contact tracing. Sci Eng Ethics 2021 Mar 29;27(2):23 [FREE Full text] [doi: 10.1007/s11948-021-00301-0] [Medline: 33779818]
34. Tran CD, Nguyen TT. Health vs. privacy? The risk-risk tradeoff in using COVID-19 contact-tracing apps. Technol Soc 2021 Nov;67:101755 [FREE Full text] [doi: 10.1016/j.techsoc.2021.101755] [Medline: 34566204]
35. Min-Allah N, Alahmed BA, Albreek EM, Alghamdi LS, Alawad DA, Alharbi AS, et al. A survey of COVID-19 contact-tracing apps. Comput Biol Med 2021 Oct;137:104787 [FREE Full text] [doi: 10.1016/j.compbiomed.2021.104787] [Medline: 34482197]
36. Wang S, Ding S, Xiong L. A new system for surveillance and digital contact tracing for COVID-19: spatiotemporal reporting over network and GPS. JMIR Mhealth Uhealth 2020 Jun 10;8(6):e19457 [FREE Full text] [doi: 10.2196/19457] [Medline: 32499212]
37. Kayaalp M. Patient privacy in the era of Big Data. Balkan Med J 2018 Jan 20;35(1):8-17 [FREE Full text] [doi: 10.4274/balkanmedj.2017.0966] [Medline: 28903886]
38. Yeh MJ. Participated without consent: mandatory authorization of government database for secondary use. Dev World Bioeth 2020 Dec;20(4):200-208. [doi: 10.1111/dewb.12259] [Medline: 32155680]
39. Gao Y, Li H, Luo Y. An empirical study of wearable technology acceptance in healthcare. Ind Manag Data Syst 2015;115(9):1704-1723. [doi: 10.1108/imds-03-2015-0087]
40. Kokolakis S. Privacy attitudes and privacy behaviour: a review of current research on the privacy paradox phenomenon. Comput Secur 2017 Jan;64:122-134. [doi: 10.1016/j.cose.2015.07.002]
41. Gerber N, Reinheimer B, Volkamer M. Investigating people's privacy risk perception. Proc Privacy Enhancing Technol 2019;2019(3):267-288. [doi: 10.2478/popets-2019-0047]
42. Swartz P, Da Veiga A, Martins N. Validating an information privacy governance questionnaire to measure the perception of employees. Inform Comput Secur 2021 May 18;29(5):761-786. [doi: 10.1108/ics-08-2020-0135]
43. Barth S, de Jong MD. The privacy paradox – investigating discrepancies between expressed privacy concerns and actual online behavior – a systematic literature review. Telemat Inform 2017 Nov;34(7):1038-1058. [doi: 10.1016/j.tele.2017.04.013]
44. Glasgow G, Butler S, Iyengar S. Survey response bias and the 'privacy paradox': evidence from a discrete choice experiment. Appl Econ Lett 2021;28(8):625-629. [doi: 10.1080/13504851.2020.1770183]

Abbreviations

COCOA: COVID-19 Contact-Confirming Application
OR: odds ratio
SNS: social networking site

Edited by A Mavragani; submitted 04.03.22; peer-reviewed by J Wacksman, H Moghaddasi, S Paydar; comments to author 18.04.22; revised version received 05.05.22; accepted 19.05.22; published 20.07.22.

Please cite as:
Bito S, Hayashi Y, Fujita T, Yonemura S
Public Attitudes Regarding Trade-offs Between the Functional Aspects of a Contact-Confirming App for COVID-19 Infection Control and the Benefits to Individuals and Public Health: Cross-sectional Survey
JMIR Form Res 2022;6(7):e37720
URL: https://formative.jmir.org/2022/7/e37720
doi:10.2196/37720
PMID:35610182

Original Paper

Preferences in the Willingness to Download a COVID-19 Contact Tracing App in the Netherlands and Turkey: Experimental Study

Frans Folkvord[1,2], PhD; Lutz Peschke[3], BSc, MSc, PhD; Yasemin Gümüş Ağca[3], PhD; Karlijn van Houten[2], MSc; Giacomo Stazi[2], MSc; Francisco Lupiáñez-Villanueva[2,4], PhD

[1]Tilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, Netherlands

[2]PredictBy, PredictBy Research, Barcelona, Spain

[3]Department of Communication and Design, Bilkent University, Ankara, Turkey

[4]Department of Information and Communication Science, Universitat Oberta de Catalunya, Barcelona, Spain

Corresponding Author:
Frans Folkvord, PhD
Tilburg School of Humanities and Digital Sciences
Tilburg University
Warrandelaan 2
Tilburg, 5037 AB
Netherlands
Phone: 31 682240480
Email: fransfolkvord@gmail.com

Abstract

Background: Despite the worldwide growth in using COVID-19 contact tracing apps (CTAs) and the potential benefits for citizens, governments, health care professionals, businesses, and other organizations, only a few studies have examined the factors affecting the levels of willingness to download a CTA.

Objective: This study aimed to investigate individuals' preferences in the willingness to download a health app.

Methods: We conducted an experimental study in 2 countries, the Netherlands (N=62) and Turkey (N=83), using 4 different vignettes (ie, data protection, manufacturer, reward, and gaming models) with different attributes. Participants were randomly assigned to 1 of the conditions within the vignettes.

Results: The results showed that data protection and gaming elements are factors that influence the willingness to download a COVID-19 CTA. More specifically, we see that data protection is an important factor explaining the willingness to download the app in Turkey, whereas including gaming elements significantly affects the willingness to download the app in the Netherlands.

Conclusions: COVID-19 CTAs are highly promising to reduce the spread of the virus and make it easier to open up society faster, especially because they can be used quickly and share information rapidly. COVID-19 CTA developers must ensure that their apps satisfactorily and sufficiently address ethical considerations, even in times of crisis. Furthermore, integrating gaming elements in the CTA could enhance the willingness to download the CTA.

(JMIR Form Res 2022;6(7):e37891) doi:10.2196/37891

KEYWORDS

COVID-19 tracing app; willingness to download; discrete choice task; pandemic; mitigation strategies; COVID-19; health application; mobile health; gaming; tracing application; digital health; data protection

Introduction

Background

The COVID-19 pandemic affected human health tremendously, as well as social and economic life, increasing the urge to find effective measures to start reviving public life as soon as possible while minimizing the risk of infection and hospitalization of patients [1]. Current developments across the world show that one strategy might be through technology, accompanied by mobile apps [2,3]. Considering that COVID-19 has proven to be highly contagious, especially newer variants such as Delta and Omicron [4]—sometimes without the carrier experiencing symptoms and infecting several others before actually testing positive—it is important to trace contacts and identify individuals who had close contact with the carrier.

Normally, this tracking is accomplished through a personal interview with the infected individual, establishing the people who have been in (close) contact with the carrier, and subsequently developing strategies to mitigate further spreading. However, it is difficult for people to accurately remember all the persons that they have been in close contact with, including those who cannot be identified because they are unknown to the carrier. Moreover, considering the rapidly increasing numbers of infected people in a majority of the countries worldwide at the moment, conducting the interviews manually would require a considerable workforce of trained individuals to hold the interviews effectively, which is a highly costly exercise and has shown to be ineffective in times of high infection rates [5].

Research has shown that COVID-19 contact tracings apps (CTAs) could mitigate the current pandemic by informing people instantly when they have been in close contact with an infected individual [5-9]. CTAs have been implemented in nearly all countries for the identification of pandemic hot spots, supporting collaborative information processes between the public and health authorities about critical contacts with infectious citizens. One of the main goals is to initiate knowledge circulation between the 5 systems of the Quintuple Helix [10], which includes the decision processes of policy makers. According to Oldeweme et al [6], the willingness to download processes of CTAs are fostered by the reduction of uncertainties—perceived privacy and performance risks—through trust in governments and public participation but do not reduce social risks and health-related COVID-19 concerns. Although these new possibilities promise several benefits for mitigating the pandemic, the actual willingness to download and the long-term use of mobile health (mHealth) apps is rather low and lags behind their potential [11]. As a result, despite the promising effects of the CTAs to mitigate the pandemic, there is insufficient evidence supporting the willingness to download preferences for these apps, although willingness to download is the first, and therefore important, step. This lack of evidence urges us to take a step back to better understand which factors influence the willingness to download CTAs [12,13].

A first step in this endeavor is to understand which factors affect people's willingness to download CTAs. Considering the limited understanding of the general cognitive motivators that trigger people's willingness to download health apps, especially in times of crisis, it is important to examine which attributes of health apps are preferred when downloading CTAs. An extensive systematic literature review [14] has shown that, in support of the Technology Acceptance Model, 8 key themes can be categorized for clinicians' adoption of mHealth tools: usefulness, ease of use, design, compatibility, technical issues, content, personalization, and convenience. In addition, from a patient's perspective [15], usefulness, ease of use, data-related factors, monetary factors, technical issues, and user experience were considered to be important factors for the adoption of an mHealth app.

Furthermore, intrinsic motivation to use such a health app is considered a strong predictor of actual willingness to download and use [16]. In addition, Salomoni et al [17] point out that rethinking how the users receive and interact with a health app is important to better understand the role of diegetic interfaces, which should also be taken into account when developing a CTA. Without the proper comprehension of the cognitive motivators that explain the willingness to download and use mHealth apps, it would be very difficult to establish the effectiveness of mHealth apps and fully understand individuals' use of such apps.

Attributes Explaining Preferences in the Willingness to Download CTAs

The rising use of mHealth apps threatens to change the way substantial amounts of health data will be managed, with a paradigm shift from mainframe systems located in the facilities of health care providers to apps on mobile phones and data stored in shared cloud services [18,19]. According to Klar and Lanzerath [20], besides the challenges of effectiveness, technological problems and the risks of privacy and equity have to be considered. Ryan [21] has, for example, evaluated South Korea's digital tracing app through the lens of 4 human rights principles to determine if this response was ethically justifiable (ie, necessary, proportional, scientifically valid, and time-bounded) and concluded that the Korean digital CTA was scientifically valid and proportionate—meeting the necessity requirement—but it was too vaguely defined to meet the time-boundedness requirement. More specifically, the prerequisite of an ethical deployment of CTAs is voluntariness, starting with deciding to carry a smartphone, choosing to download and install a CTA, leaving the CTA operating in the background all the times, and finally, sharing contact logs when tested positive. Peschke et al [22] discussed the voluntariness of CTA deployment in different countries. Based on an analysis of more than 35 CTAs worldwide [23], 2 main categories of CTA deployment could be identified. First, for the participation of the public life, such as entering university campuses and shopping malls in Turkey, an individual Hayat Eve Sığar (Turkish for "life fits into home") code has to be generated and presented to the gatekeepers. In these apps, the data of the users are collected in a central place and sometimes shared with other state institutions. If COVID-19 is detected within the passengers of the same vehicle or at the same location within the next 14 days, everyone at that place, family doctors, and the filiation team are informed [24]. Second, Google and Apple have developed a system called Exposure Notifications System to bring out anonymized identity information to the users' environment via Bluetooth. The importance of data protection and privacy issues regarded by potential users of the CTAs have not been sufficiently understood yet.

A recent study has shown that privacy perceptions are related to the use of mHealth apps, in which people with more concerns about the secondary use of their personal data were less likely to use certain mHealth apps [25]. In addition, the manufacturer of a health app can be considered a heuristic in the consumers' willingness to download a CTA. There is reason to assume that this factor may be an important, as the industry has struggled with its public image over the past few decades. Companies need to negotiate a tension between, on the one hand, striving for optimal health care and, on the other hand, striving for profit [26]. In the eyes of the public, it is not always clear that the

industry has the patients' interests at heart [27]. Therefore, people would probably prefer if an mHealth app is manufactured by the government than by a company.

By contrast, the Edelman Trust Barometer [28] reveals that business is the only institution that is understood to be competent and ethical by the public in 18 of the 27 countries evaluated. The public is more likely to trust in business as it acts as the guardian of information quality, embraces sustainable practices, and provides robust COVID-19 health and safety response [28]. However, the Netherlands noticed an increase of trust in both business and government [28]. The Edelman Trust Barometer did not capture data from Turkey, but the result of 2 studies conducted in 2020 show that the pandemic crisis revealed 2 different results. Bostan et al [29] conclude that the public trusts the authorities and the accuracy of the decisions taken by the state to combat the pandemic. However, Tanca et al [30] perceive the ambiguity and capriciousness of the public's reactions as a result of the uncertainty of both the government and national economy. A special report about trust and coronavirus published by Edelman Trust Barometer [31] reveals that 85% of the respondent want to hear more from scientists and less from politicians.

Lastly, it is worth emphasizing the importance of the design of the CTA when participants are downloading apps. Web-based gaming studies have shown that elements and features with clear goals at every step [32], immediate feedback [33-35], and balance between challenge and skill [36] grounded in the flow theory [37] are considered key elements of the willingness to download apps.

Finally, since last year, the majority of the public has prioritized increasing their own media and information literacy as well as science literacy, but only 26% of the respondents have good information hygiene—considering news engagement, avoiding information echo chambers, verifying information, and not amplifying unvetted information [28].

Hypotheses

In this study, we conducted 4 small experimental studies based on vignettes in 2 different countries (the Netherlands and Turkey). In the vignette study, we manipulated the manufacturer (government vs company), data protection (data protection vs no information), reward (no reward vs voucher as a reward), and gaming (no gaming elements vs gaming elements) variables and assessed the likelihood of downloading a COVID-19 CTA to test the following hypothesis:

1. Participants who are exposed to a COVID-19 CTA developed by a government will be more willing to download the app than participants who were exposed to a COVID-19 CTA that is developed by a company (Hypothesis [H]1).
2. Participants who are exposed to a COVID-19 CTA whereby there was communication that the data is protected by law will be more willing to download the app than participants who are exposed to a COVID-19 CTA whereby no information was given about the data protection (H2).
3. Participants who are exposed to a COVID-19 CTA and will receive a reward based on scientific evidence will be more willing to download the app than participants who are exposed to a COVID-19 CTA and will receive no rewards (H3).
4. Participants who are exposed to a COVID-19 CTA that includes gaming elements will be more willing to download the app than participants who are exposed to a COVID-19 CTA without any gaming elements (H4).

Methods

Design

In this study, we conducted 4 small experimental studies through a web-based questionnaire using vignettes to simulate real-life situations. The use of a web-based questionnaire provided completion time data to support the internal validity checks and enabled an accurate record of the time taken to complete the surveys. There were 2 rounds of cognitive testing (N=12) undertaken in the Netherlands to check the participants' comprehension of information when making choices. These pretests confirmed that a study based on the questionnaire was acceptable and understandable for participants, after some minor revisions in the explanation of the task.

Procedure

All survey participants were informed about the overall study goals and procedures. First, participants were asked to provide sociodemographic information, including their age group, gender, education, and employment status. Subsequently, an introduction to the vignette studies provided an explanation of what was expected from the participants. One example of the vignette is the following, whereby the last sentence is the part of the script that varied among the conditions.

> *Imagine that an app is presented to you to mitigate the pandemic, by providing you detailed information about the people that you have been around with, how close, and for how long, that are infected by the COVID virus, or not. Based on the collected data, the app will provide tailored advice to improve your health. We will present you some pairs of options and will ask you to select the one that you would prefer to use the app or not. Importantly, the app is developed by the government.*

Subsequently, participants were asked to score the likelihood that they would download the app and whether they would download the app. The other vignettes entailed exactly the same text, except the last sentence was replaced by the specific condition the participant was randomly allocated to (government vs company, data protection vs no information, no reward vs voucher as a reward, or no gaming elements vs gaming elements).

Ethics Approval

Only those who agreed to participate in the study gained access to the web-based survey. The approval of the Ethical Committee of the Tilburg School of Humanities and Digital Sciences to conduct the experiment was obtained (REDC 2021.73). Through signing informed consent, participants were ensured that their data would remain confidential, and they were told that they

could cease participation at any moment. All participants participated anonymously, and the collected data was stored in a dark archive at the Tilburg University.

Participants

For this study, we used data collected from 2 different countries (Turkey and the Netherlands). This study was part of a larger project, whereby multiple experiments were conducted. After the participants finished the discrete choice experiment, they also participated in a separate experiment that is reported in a separate paper. The data in the Netherlands (N=62) and Turkey (N=83) were collected through a web-based survey administered through Qualtrics (SAP America Inc). Participant characteristics are shown in Table 1.

Table 1. Descriptive information about the participants per country.

Characteristic	Netherlands (N=62)	Turkey (N=83)
Gender, women, n (%)	31 (50)	43 (52)
Age (year), mean (SD)	23.92 (7.45)	21.87 (2.46)
Educational level, n (%)		
Primary education	3 (5)	0 (0)
High school diploma	3 (5)	10 (12)
Some years of university	28 (45)	61 (73)
University degree	18 (29)	10 (12)
Post-graduate degree	10 (16)	2 (2)
Employment status, n (%)		
Employed/self-employed	11 (18)	12 (14)
Unemployed	7 (11)	0 (0)
Student	44 (71)	71 (86)
Retired	0 (0)	0 (0)
Not working due to illness or disability	0 (0)	0 (0)
Another reason for not being in the labor force	0 (0)	0 (0)
Health app use, n (%)		
No use	27 (44)	30 (36)
1 time	8 (13)	18 (22)
2 times	11 (18)	18 (22)
3 times	7 (11)	6 (7)
4 times	3 (5)	4 (5)
5 times	0 (0)	3 (4)
>5 times	6 (10)	4 (5)
Health consciousness, mean (SD)	3.77 (0.68)	3.93 (0.83)
Health information orientation, mean (SD)	2.99 (0.82)	3.48 (0.92)
eHealth literacy, mean (SD)	2.85 (0.96)	3.09 (1.00)

Measures

Dependent Variables

Willingness to download the app was measured through the question “Please indicate on a scale of 1 to 10 how likely it is that you would download the app. Please place yourself at a point on this scale where ‘0’ indicates that you would ‘definitely not download the app,’ ‘10’ indicates that you would ‘definitely download the app,’ and the remaining numbers indicate something in between these 2 positions.” whereby participants could answer on a visual analog scale.

Intention to download the app was measured through the question “Would you download the app?” whereby participants could answer yes or no.

Intrapersonal Factors

Health app use was measured by asking how often the participant used a health app, varying from 0 (never) to 6 (more than 5 times), and how much time the participant spent using a health app in the last week, varying from 0 (0 hours) to 6 (more than 1 hour).

Health consciousness was measured using 5 statements that were rated on a 5-point scale (from 1 representing strongly

disagree to 5 representing strongly agree) [38]. The reliability of the scale was high (α=.82).

Health information orientation was measured using 8 statements rated on a 5-point scale (from 1 representing strongly disagree to 5 representing strongly agree) [38]. The reliability of the scale was high (α=.87)

eHealth literacy was measured using 8 statements rated on a 5-point scale (from 1 representing strongly disagree to 5 representing strongly agree) [39]. The reliability of the scale was high (α=.95).

Statistical Analyses

Several analyses of covariance (ANCOVAs) were conducted to assess the participants' likelihood to download the app—first, for all participants and subsequently, for each country. Next, we conducted logistic regression analyses to assess the participants' probability to download the app.

Results

Developer Models

The first ANCOVA tested if participants who were exposed to a COVID-19 CTA developed by a government were more willing to download the app than participants who were exposed to a COVID-19 CTA developed by a company, whereby age, gender, education, health consciousness, health information orientation, and eHealth literacy were included as covariates in the analysis. No effects were found for the developer ($F_{7,137}$=1.307; *P*=.26). For age ($F_{1,137}$=9.848; *P*=.002) and eHealth literacy ($F_{1,137}$=7.047; *P*=.009), we found significant relationships. For the other factors, we found no significant differences.

Next, the logistic regression analysis showed that there was no effect of the developer on whether the participants would download the app (*P*=.32). For age (*P*<.001) and education (*P*=.02), we found significant relations to whether the participants would download the app.

A separate ANCOVA for the Netherlands showed that no effects were found for the developer ($F_{7,54}$=1.638; *P*=.21). We found a significant relationship for age (*P*=.01). For the other factors, we found no significant differences. In addition, the logistic regression analysis showed that there was no effect of the developer on whether the participants would download the app (*P*=.46). For age (*P*=.01) and education (*P*=.008), we found significant relations to whether the participants would download the app.

In Turkey, an ANCOVA showed that no effects were found for the developer as well ($F_{7,75}$=0.094; *P*=.76). For the other factors, we found no significant differences. Furthermore, the logistic regression analysis showed that there was no effect of the developer on whether the participants would download the app (*P*=.49). For age (*P*=.02) and eHealth literacy (*P*=.01), we found significant relations to whether the participants would download the app.

Data Protection Models

The second ANCOVA tested if participants who were exposed to a COVID-19 CTA developed whereby data was protected by European legislation were more willing to download the app than participants who were exposed to a COVID-19 CTA whereby no information was given about the data protection, whereby age, gender, education, health consciousness, health information orientation, and eHealth literacy were included as covariates in the analysis. The results showed that participants who were exposed to a COVID-19 CTA developed whereby data was protected by law (mean 6.93, SD 2.44) were significantly more willing to download the app than participants who were exposed to a COVID-19 CTA whereby no information was given about the data protection (mean 6.93, SD 2.44; $F_{7,137}$=19.125; *P*>.001). For the other factors, we found no significant differences. The logistic regression analysis showed that there was a significant effect of data protection on whether the participants would download the app (*P*<.001). The probability that a participant would download the app would be 3.886 times greater when the data was protected by EU legislation than when no information was provided about what would be done with the data. For the other factors, we found no significant relations.

Separate analyses for the Netherlands showed that participants who were exposed to a COVID-19 CTA developed whereby data was protected by law were not more likely to download the app than participants who were exposed to a COVID-19 CTA whereby no information was given about the data protection ($F_{7,54}$=1.130; *P*=.29). For the other factors, we found no significant differences. The logistic regression analysis showed that there was a significant effect of data protection on whether the participants would download the app (*P*=.03). The probability that a participant would download the app would be 3.327 times greater when the data was protected by law than when no information was provided about what would be done with the data. For the other factors, we found no significant relations.

In Turkey, participants who were exposed to a COVID-19 CTA developed whereby data was protected by law were significantly more willing to download the app (mean 7.27, SD 2.52) than participants who were exposed to a COVID-19 CTA whereby no information was given about the data protection (mean 4.17, SD 3.09; $F_{7,75}$=24.584; *P*>.001). For the other factors, we found no significant differences. The logistic regression analysis showed that there was a significant effect of data protection on whether the participants would download the app (*P*=.007). The probability that a participant would download the app would be 3.847 times greater when the data was protected by law than when no information was provided about what would be done with the data. For the other factors, we found no significant relations.

Reward Models

The next ANCOVA tested if participants who were exposed to a COVID-19 CTA and would receive a reward if they adopted the app were more willing to download the app than participants who were exposed to a COVID-19 CTA and would receive no

rewards, whereby age, gender, education, health consciousness, health information orientation, and eHealth literacy were included as covariates in the analysis. The results showed that participants who were exposed to a COVID-19 CTA and would receive a reward if they adopted the app were not more willing to download the app than participants who were exposed to a COVID-19 CTA and would receive no rewards ($F_{7,137}$=0.324; P=.57). For the other factors, we found no significant differences. In addition, the logistic regression analysis showed that there was no effect of reward on whether the participants would download the app (P=.93). For age (P=.002), health consciousness (P=.04), health information orientation (P<.001), and eHealth literacy (P=.003), we found significant relations to whether the participants would download the app.

Separate analyses for the Netherlands showed that participants who were exposed to a COVID-19 CTA and would receive a reward if they adopted the app were not more willing to download the app than participants who were exposed to a COVID-19 CTA and would receive no rewards ($F_{7,54}$=0.322; P=.57). For the other factors, we found no significant differences. The logistic regression analysis showed that there was no effect of reward on whether the participants would download the app (P=.24). For gender (P=.05), age (P=.003), health consciousness (P=.006), health information orientation (P=.004), and eHealth literacy (P=.003), we found significant relationships to whether the participants would download the app. In Turkey, we found similar results for the reward (P=.74) and no other significant relations.

Gaming Models

The next ANCOVA tested if participants who were exposed to a COVID-19 CTA that included gaming elements were more willing to download the app than participants who were exposed to a COVID-19 CTA without any gaming elements, whereby age, gender, education, health consciousness, health information orientation, and eHealth literacy were included as covariates in the analysis. The results showed that participants who were exposed to a COVID-19 CTA where gaming elements were included (mean 6.46, SD 2.28) were significantly more willing to download the app than participants who were exposed to a COVID-19 CTA without any gaming elements (mean 5.46, SD 2.73; $F_{7,137}$=6.603; P=.01). We found a significant relationship for age ($F_{1,137}$=6.008; P=.02). For the other factors, we found no significant differences. Next, the logistic regression analysis showed that there was no effect of gaming on whether the participants would download the app (P=.15). For education (P=.03) and health information orientation (P=.02), we found significant relations to whether the participants would download the app.

Separate analyses for the Netherlands showed that participants who were exposed to a COVID-19 CTA where gaming elements were included (mean 6.51, SD 2.07) were significantly more willing to download the app than participants who were exposed to a COVID-19 CTA without any gaming elements (mean 4.87, SD 2.27; $F_{7,54}$=26.755; P<.001). Furthermore, we found a significant relationship for age ($F_{1,54}$=13.399; P<.001). For the other factors, we found no significant differences. Next, the logistic regression analysis showed that there was a significant effect of gaming on whether the participant would download the app (P=.006). Participants who were exposed to a COVID-19 CTA with gaming elements were 20.516 times more likely to download the app than participants who were exposed to a COVID-19 CTA without gaming elements. For age (P=.03), education (P=.02), health consciousness (P=.04), and eHealth literacy (P=.03), we found significant relations to whether the participants would download the app. In Turkey, we found no significant differences between the gaming conditions ($F_{7,75}$=2.380; P=.57) and for the other factors. The logistic regression analysis showed that there was no effect of gaming on whether the participants would download the app (P=.62). No other factors were significantly related to whether the participants would download the app.

Discussion

Principal Findings

The results showed that data protection and gaming elements are important factors that affect the willingness to download a COVID-19 CTA, thereby supporting H2 and H4. For the developer and reward, we found no significant differences between the participants, thus rejecting H1 and H3. Governments might not have a second chance to get an intervention right, especially because in times of crises, trust in politicians is an important factor to mitigate the crisis at stake. Governments, developers, and deployers must ensure that COVID-19 CTAs satisfactorily address the ethical questions that were set out [3]. According to the Organisation for Economic Cooperation and Development [40], citizens expect integrity, openness, and fairness in communication and knowledge transfer, where responsiveness and reliability are crucial factors. As a result, a positive perception of comprehension leads to an increase in acceptance and cooperative engagement.

Comparison With Prior Work

Despite the worldwide growth in using COVID-19 CTAs and the potential benefits for all the actors within the Quintuple Helix to mitigate the pandemic more effectively and make us able to open up society sooner, only a few studies have examined the factors affecting the levels of willingness to download the apps. In this study, we investigated individuals' preferences in the willingness to download a health app. Digital contact tracing via smartphone apps was established as a new public health intervention in many countries in 2020. Most of these apps are now at a stage where they need to be evaluated as public health tools, especially because this could provide us with important lessons for future events such as the current pandemic [3,41].

Limitations

One of the limitations of the study is that we used a convenience sampling, whereby most of the participants were students from only 2 countries, thus reducing the generalizability of the outcomes. Considering that this study is part of a larger study where we will develop an actual CTA that will be tested among a larger and more representative study in 5 different countries

(Spain, the Netherlands, Finland, Germany, and Turkey), the generalizability will improve. Finally, gamification elements might have been judged differently if they were actually displayed or could be tried out in the experiment, increasing its effectiveness, which unfortunately was not possible. In the next phase of the study, we will test gaming elements in more detail.

Conclusions

Simply rolling out a CTA without ethical considerations is not acceptable and does not support the knowledge circulation in the Quintuple Helix. Even in a crisis, a "try everything" approach is dangerous when it ignores the real costs, including serious and long-lasting harms to fundamental rights and freedoms and the opportunity costs of not devoting resources to something else. As this pilot study shows, reassuring people that data protection regulations are put in place is an essential element for increasing the levels of willingness to download CTAs.

During the last few decades, mHealth initiatives have emerged at an accelerating pace; some have seen widespread willingness to download, whereas others have failed to provide sustained value [42]. These failings can be attributed to their design and implementation efforts that were initiated without a good understanding of the interdependencies between technology, societal and cultural values, and user experience in a health care setting, which has become highly apparent during the current pandemic and health care crisis. Many conceptual frameworks based on implementation science have been developed to evaluate and orient mHealth delivery. These frameworks highlight key factors that predict successful and sustainable mHealth technologies, although there is still limited understanding of the factors influencing the willingness to download these technologies. The urgency of the ongoing public health crisis stimulated the rapid development of CTAs and other mHealth innovations, and this generated a substantial number of related publications. Their coverage of the essential design and implementation characteristics for eHealth innovation remains under investigated and needs to be further researched in future studies [41].

Conflicts of Interest

None declared.

References

1. Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C, et al. The socio-economic implications of the coronavirus pandemic (COVID-19): a review. Int J Surg 2020 Jun;78:185-193 [FREE Full text] [doi: 10.1016/j.ijsu.2020.04.018] [Medline: 32305533]
2. Bassi A, Arfin S, John O, Jha V. An overview of mobile applications (apps) to support the coronavirus disease 2019 response in India. Indian J Med Res 2020 May;151(5):468-473 [FREE Full text] [doi: 10.4103/ijmr.IJMR_1200_20] [Medline: 32474557]
3. Morley J, Cowls J, Taddeo M, Floridi L. Ethical guidelines for COVID-19 tracing apps. Nature 2020 Jun 28;582(7810):29-31. [doi: 10.1038/d41586-020-01578-0] [Medline: 32467596]
4. Mahase E. COVID-19: hospital admission 50-70% less likely with omicron than delta, but transmission a major concern. BMJ 2021 Dec 24;375:n3151-n3170. [doi: 10.1136/bmj.n3151] [Medline: 34952835]
5. Davalbhakta S, Advani S, Kumar S, Agarwal V, Bhoyar S, Fedirko E, et al. A systematic review of smartphone applications available for Corona Virus Disease 2019 (COVID19) and the assessment of their quality using the Mobile Application Rating Scale (MARS). J Med Syst 2020 Aug 10;44(9):164 [FREE Full text] [doi: 10.1007/s10916-020-01633-3] [Medline: 32779002]
6. Oldeweme A, Märtins J, Westmattelmann D, Schewe G. The role of transparency, trust, and social influence on uncertainty reduction in times of pandemics: empirical study on the adoption of COVID-19 tracing apps. J Med Internet Res 2021 Feb 08;23(2):e25893 [FREE Full text] [doi: 10.2196/25893] [Medline: 33465036]
7. Wu S, Neill R, De Foo C, Chua AQ, Jung A, Haldane V, et al. Aggressive containment, suppression, and mitigation of covid-19: lessons learnt from eight countries. BMJ 2021 Nov 28;375:e067508 [FREE Full text] [doi: 10.1136/bmj-2021-067508] [Medline: 34840136]
8. Walensky RP, Del Rio C. From mitigation to containment of the COVID-19 pandemic: putting the SARS-CoV-2 genie back in the bottle. JAMA 2020 May 19;323(19):1889-1890. [doi: 10.1001/jama.2020.6572] [Medline: 32301959]
9. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the Mobile Application Rating Scale (MARS). PLoS One 2020 Nov 2;15(11):e0241480 [FREE Full text] [doi: 10.1371/journal.pone.0241480] [Medline: 33137123]
10. Carayannis EG, Campbell DFJ. Triple Helix, Quadruple Helix and Quintuple Helix and how do knowledge, innovation and the environment relate to each other? : a proposed framework for a trans-disciplinary analysis of sustainable development and social ecology. International Journal of Social Ecology and Sustainable Development 2010;1(1):41-69. [doi: 10.4018/jsesd.2010010105]
11. Walrave M, Waeterloos C, Ponnet K. Adoption of a contact tracing app for containing COVID-19: a health belief model approach. JMIR Public Health Surveill 2020 Sep 01;6(3):e20572 [FREE Full text] [doi: 10.2196/20572] [Medline: 32755882]

12. Anastasiadou D, Folkvord F, Lupiañez-Villanueva F. A systematic review of mHealth interventions for the support of eating disorders. Eur Eat Disord Rev 2018 Sep 21;26(5):394-416. [doi: 10.1002/erv.2609] [Medline: 29927004]
13. Anastasiadou D, Folkvord F, Serrano-Troncoso E, Lupiañez-Villanueva F. Mobile health adoption in mental health: user experience of a mobile health app for patients with an eating disorder. JMIR mHealth uHealth 2019 May 31;7(6):e12920 [FREE Full text] [doi: 10.2196/12920] [Medline: 31199329]
14. Jacob C, Sanchez-Vazquez A, Ivory C. Social, organizational, and technological factors impacting clinicians' adoption of mobile health tools: systematic literature review. JMIR mHealth uHealth 2020 Feb 20;8(2):e15935 [FREE Full text] [doi: 10.2196/15935] [Medline: 32130167]
15. Jacob C, Sezgin E, Sanchez-Vazquez A, Ivory C. Sociotechnical factors affecting patients' adoption of mobile health tools: systematic literature review and narrative synthesis. JMIR mHealth uHealth 2022 May 05;10(5):e36284 [FREE Full text] [doi: 10.2196/36284] [Medline: 35318189]
16. Ng JYY, Ntoumanis N, Thøgersen-Ntoumani C, Deci EL, Ryan RM, Duda JL, et al. Self-determination theory applied to health contexts: a meta-analysis. Perspect Psychol Sci 2012 Jul 29;7(4):325-340. [doi: 10.1177/1745691612447309] [Medline: 26168470]
17. Salomoni P, Prandi C, Roccetti M, Casanova L, Marchetti L, Marfia G. Diegetic user interfaces for virtual environments with HMDs: a user experience study with oculus rift. J Multimodal User Interfaces 2017 Jan 9;11(2):173-184. [doi: 10.1007/s12193-016-0236-5]
18. He D, Naveed M, Gunter CA, Nahrstedt K. Security concerns in Android mHealth apps. AMIA Annu Symp Proc 2014;2014:645-654 [FREE Full text] [Medline: 25954370]
19. Kahnbach L, Lehr D, Brandenburger J, Mallwitz T, Jent S, Hannibal S, et al. Quality and adoption of COVID-19 tracing apps and recommendations for development: systematic interdisciplinary review of European apps. J Med Internet Res 2021 Jun 02;23(6):e27989 [FREE Full text] [doi: 10.2196/27989] [Medline: 33890867]
20. Klar R, Lanzerath D. The ethics of COVID-19 tracking apps – challenges and voluntariness. Research Ethics 2020 Aug 05;16(3-4):1-9. [doi: 10.1177/1747016120943622]
21. Ryan M. In defence of digital contact-tracing: human rights, South Korea and COVID-19. International Journal of Pervasive Computing and Communications 2020 Aug 06;16(4):383-407. [doi: 10.1108/ijpcc-07-2020-0081]
22. Peschke L, Günes Peschke S, Ağca YG, Seyfafjehi S, Dündar I, Aydogdu Y. Reward mechanism in COVID-19 tracking apps and its impact on the voluntary participation of the public in sustainable innovation processes. Turkish Review of Communication Studies 2022(39):54-72. [doi: 10.17829/turcom.1019006]
23. Peschke L(, Çataloğlu B, Ceylan N, Dionio P, Dündar I, Folkvord F, et al. Review about good practices of Quadruple Helix collaboration processes and developments of COVID-19 tracking apps and platforms. PandeVITA. 2021. URL: https://www.pandevita.eu/images/D3-1_GoodPractices-CTAs.pdf [accessed 2022-07-22]
24. HES kodu nedir. Hayat Eve Sığar. 2020. URL: https://hayatevesigar.saglik.gov.tr/hes.html [accessed 2021-12-31]
25. Bol N, Helberger N, Weert JCM. Differences in mobile health app use: a source of new digital inequalities? The Information Society 2018 Apr 26;34(3):183-193. [doi: 10.1080/01972243.2018.1438550]
26. Bauchner H, Fontanarosa PB. Restoring confidence in the pharmaceutical industry. JAMA 2013 Feb 13;309(6):607-609. [doi: 10.1001/jama.2013.58] [Medline: 23403686]
27. Olsen AK, Whalen MD. Public perceptions of the pharmaceutical industry and drug safety: implications for the pharmacovigilance professional and the culture of safety. Drug Saf 2009;32(10):805-810. [doi: 10.2165/11316620-000000000-00000] [Medline: 19722724]
28. Edelman Trust Barometer 2021. Edelman. 2021. URL: https://www.edelman.com/sites/g/files/aatuss191/files/2021-03/2021%20Edelman%20Trust%20Barometer.pdf [accessed 2022-02-20]
29. Bostan S, Erdem R, Öztürk YE, Kılıç T, Yılmaz A. The effect of COVID-19 pandemic on the Turkish society. Electron J Gen Med 2020;17(6):em237. [doi: 10.29333/ejgm/7944]
30. Tanca D, Aydoğ E, Murphy A, Zinciroğlu Ö. The political and economic impact of the coronavirus pandemic in Turkey. Center for Economics and Foreign Policy Studies. 2020 Jan. URL: https://edam.org.tr/wp-content/uploads/2020/11/The-Political-and-Economic-Impact-of-the-Corona-Virus-in-Turkey.pdf [accessed 2022-07-22]
31. Special report: trust and the coronavirus. Edelman. 2020. URL: https://www.edelman.com/sites/g/files/aatuss191/files/2020-03/2020%20Edelman%20Trust%20Barometer%20Coronavirus%20Special%20Report_0.pdf [accessed 2022-02-02]
32. Csikszentmihalyi M. Flow: The Psychology of Optimal Experience. New York, NY: Harper & Row; 1990.
33. Kiili K, de Freitas S, Arnab S, Lainema T. The design principles for flow experience in educational games. Procedia Comput Sci 2012;15:78-91. [doi: 10.1016/j.procs.2012.10.060]
34. Wang H, Sun CT. Game reward systems: gaming experiences and social meanings. 2011 Jan Presented at: DiGRA 2011 Conference: Think Design Play; September 14-17, 2011; Hilversum, The Netherlands URL: http://www.digra.org/wp-content/uploads/digital-library/11310.20247.pdf
35. Bracken CC, Jeffres LW, Neuendorf KA. Criticism or praise? the impact of verbal versus text-only computer feedback on social presence, intrinsic motivation, and recall. Cyberpsychol Behav 2004 Jun;7(3):349-357. [doi: 10.1089/1094931041291358] [Medline: 15257835]
36. Koster R. A Theory of Fun for Game Design. Scottsdale, AZ: Paraglyph Press; 2005.

37. Lopez SJ, Snyder CR, editors. The Oxford Handbook of Positive Psychology. 2 ed. Oxford, United Kingdom: Oxford University Press; Jul 2009:88-102.
38. Dutta-Bergman MJ. Health attitudes, health cognitions, and health behaviors among internet health information seekers: population-based survey. J Med Internet Res 2004 May 28;6(2):e15 [FREE Full text] [doi: 10.2196/jmir.6.2.e15] [Medline: 15249264]
39. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
40. Organisation for Economic Cooperation and Development. Trust and Public Policy: How Better Governance Can Help Rebuild Public Trust. Paris, France: OECD Publishing; Mar 27, 2017.
41. Osmanlliu E, Rafie E, Bédard S, Paquette J, Gore G, Pomey M. Considerations for the design and implementation of COVID-19 contact tracing apps: scoping review. JMIR mHealth uHealth 2021 Jun 09;9(6):e27102 [FREE Full text] [doi: 10.2196/27102] [Medline: 34038376]
42. Patrick K, Hekler EB, Estrin D, Mohr DC, Riper H, Crane D, et al. The pace of technologic change: implications for digital health behavior intervention research. Am J Prev Med 2016 Nov;51(5):816-824. [doi: 10.1016/j.amepre.2016.05.001] [Medline: 27745681]

Abbreviations

ANCOVA: analysis of covariance
CTA: contact tracing app
H: hypothesis
mHealth: mobile health

Edited by A Mavragani; submitted 10.03.22; peer-reviewed by M Roccetti, C Jacob; comments to author 08.05.22; revised version received 29.05.22; accepted 31.05.22; published 28.07.22.

Please cite as:
Folkvord F, Peschke L, Ağca YG, van Houten K, Stazi G, Lupiáñez-Villanueva F
Preferences in the Willingness to Download a COVID-19 Contact Tracing App in the Netherlands and Turkey: Experimental Study
JMIR Form Res 2022;6(7):e37891
URL: https://formative.jmir.org/2022/7/e37891
doi:10.2196/37891
PMID:35867840

Short Paper

COVID-box Experiences of Patients and Health Care Professionals (COVID-box Project): Single-Center, Retrospective, Observational Study

Ebru Dirikgil[1], MD; Kim Brons[2]; Michael Duindam[2]; Geert H Groeneveld[3], MD, PhD; J J Miranda Geelhoed[4], MD, PhD; Christian Heringhaus[5], MD; Paul J M van der Boog[1], MD, PhD; Ton J Rabelink[1], MD, PhD; Willem Jan W Bos[1,6], MD, PhD; Niels H Chavannes[7,8], MD, PhD; Douwe E Atsma[8,9], MD, PhD; Y K Onno Teng[1], MD, PhD

[1]Department of Nephrology, Leiden University Medical Center, Leiden, Netherlands
[2]Directorate of Quality and Patient Safety, Leiden University Medical Center, Leiden, Netherlands
[3]Department of Infectious diseases, Leiden University Medical Center, Leiden, Netherlands
[4]Department of Pulmonology, Leiden University Medical Center, Leiden, Netherlands
[5]Department of Emergency Medicine, Leiden University Medical Center, Leiden, Netherlands
[6]Department of Internal Medicine, St Antonius Hospital, Nieuwegein, Netherlands
[7]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, Netherlands
[8]National eHealth Living Lab, Leiden University Medical Center, Leiden, Netherlands
[9]Department of Cardiology, Leiden University Medical Center, Leiden, Netherlands

Corresponding Author:
Y K Onno Teng, MD, PhD
Department of Nephrology
Leiden University Medical Center
Albinusdreef 2
Leiden, 2333 ZA
Netherlands
Phone: 31 071 52621
Email: y.k.o.teng@lumc.nl

Abstract

Background: During the COVID-19 pandemic, several home monitoring programs have described the success of reducing hospital admissions, but only a few studies have investigated the experiences of patients and health care professionals.

Objective: The objective of our study was to determine patients' and health care professionals' experiences and satisfaction with employing the COVID-box.

Methods: In this single-center, retrospective, observational study, patients and health care professionals were asked to anonymously fill out multiple-choice questionnaires with questions on a 5-point or 10-point Likert scale. The themes addressed by patients were the sense of reassurance and safety, experiences with teleconsultations, their appreciation for staying at home, and the instructions for using the COVID-box. The themes addressed by health care professionals who treated patients with the COVID-box were the characteristics of the COVID-box, the technical support service and general satisfaction, and their expectations and support for this telemonitoring concept. Scores were interpreted as *insufficient* (≤2 or ≤5, respectively), *sufficient* (3 or 6-7, respectively), or *good* (≥4 or ≥8, respectively) on a 5-point or 10-point Likert scale.

Results: A total of 117 patients and 25 health care professionals filled out the questionnaires. The median score was 4 (IQR 4-5) for the sense of safety, the appreciation for staying at home, and experiences with teleconsultations, with good scores from 76.5% (88/115), 86% (56/65), and 83.6% (92/110) of the patients, respectively. Further, 74.4% (87/117) of the patients scored the home monitoring program with a score of ≥8. Health care professionals scored the COVID-box with a minimum median score of 7 (IQR 7-10) on a 10-point scale for all domains (ie, the characteristics of the COVID-box and the technical support service and general satisfaction). For the sense of safety, user-friendliness, and additional value of the COVID-box, the median scores were 8 (IQR 8-10), 8 (IQR 7-9), and 10 (IQR 8-10), respectively, with good scores from 86% (19/22), 75% (15/20), and 96% (24/25) of the health care professionals, respectively. All health care professionals (25/25, 100%) gave a score of ≥8 for supporting this home monitoring concept, with a median score of 10 (IQR 10-10).

Conclusions: The positive experiences and satisfaction of involved users are key factors for the successful implementation of a novel eHealth solution. In our study, patients, as well as health care professionals, were highly satisfied with the use of the home monitoring program—the COVID-box project. Remote home monitoring may be an effective approach in cases of increased demand for hospital care and high pressure on health care systems.

(JMIR Form Res 2022;6(7):e38263) doi:10.2196/38263

KEYWORDS

COVID-19; digital health; eHealth; telemedicine; telemonitoring; hospital admission; health care professional; thematic analysis; user satisfaction; usability; home monitoring; health care system

Introduction

The COVID-19 pandemic resulted in an increased demand for hospital care. To keep up with this surging demand, home monitoring was implemented in many countries to avoid unnecessary hospital admissions and detect clinical deterioration in patients at an earlier stage to allow for timely admission and readmission [1].

At the Leiden University Medical Center (LUMC), the Netherlands, we developed the COVID-box project, which is a home monitoring program for patients with (suspected) COVID-19. After a hospital or emergency department visit, patients with (suspected) COVID-19 receive Bluetooth-connected devices (blood pressure monitor, pulse oximeter, and thermometer) and instructions for monitoring their vital parameters 3 times per day, combined with daily teleconsultations carried out by a health care professional. Once the patients get home, the COVID-box team calls the patients to help with the installation of devices and answer questions. The COVID-box team is reachable during office hours for solving logistics issues and answering questions from patients and health care professionals. A detailed description of the telemonitoring program was published previously [2].

The implementation of several home monitoring programs has resulted in a reduction in hospital admissions by allowing for the safe survey of clinical symptoms and vitals [3-9]. Although many observational studies have studied the effectiveness of home monitoring, few studies have reported on patients' and doctors' experiences with telemonitoring [5,10-13]. We focused on the COVID-box experiences of patients and health care professionals.

Methods

Ethics Approval

Central ethical approval was obtained for this study from the medical ethics committee of the LUMC.

Study Design

This retrospective observational study was conducted as part of the COVID-box project, which was initiated at the LUMC in May 2020 upon the first wave of the pandemic. A detailed operational description of the COVID-box project was previously reported [2]. In this study, we evaluated patients with COVID-19 and their experiences with telemonitoring by surveying patients after the completion of the telemonitoring phase and full recovery. Patients were asked to anonymously fill out questionnaires regarding the sense of reassurance and safety, experiences with teleconsultations, their appreciation for staying at home, and the instructions for the COVID-box. Health care professionals who treated patients with the COVID-box were given separate questionnaires regarding the characteristics of the COVID-box, the technical support service and general satisfaction (estimated patient satisfaction), and their expectations and support for this telemonitoring concept. All questions were multiple-choice questions on a 5-point or 10-point Likert scale. For this study, scores were interpreted as follows: on a 5-point Likert scale, scores of ≤2 were *insufficient* scores, scores of 3 were *sufficient* scores, and scores of ≥5 were *good* scores; on a 10-point scale, scores of ≤5 were marked as *insufficient*, scores of 6 to 7 were marked as *sufficient*, and scores of ≥8 were marked as *good*. Unvalidated questionnaires were developed by members of the Department of the Directorate of Quality and Patient Safety and clinicians from the Department of Internal Medicine. The questionnaires were web-based and anonymous.

Statistics

Descriptive statistics were used to summarize the results. The scores given by patients and health care professionals are presented as medians with IQRs, and the number of patients and health care professionals are presented as absolute numbers and percentages. We used IBM SPSS Statistics Version 25 (IBM Corporation).

Results

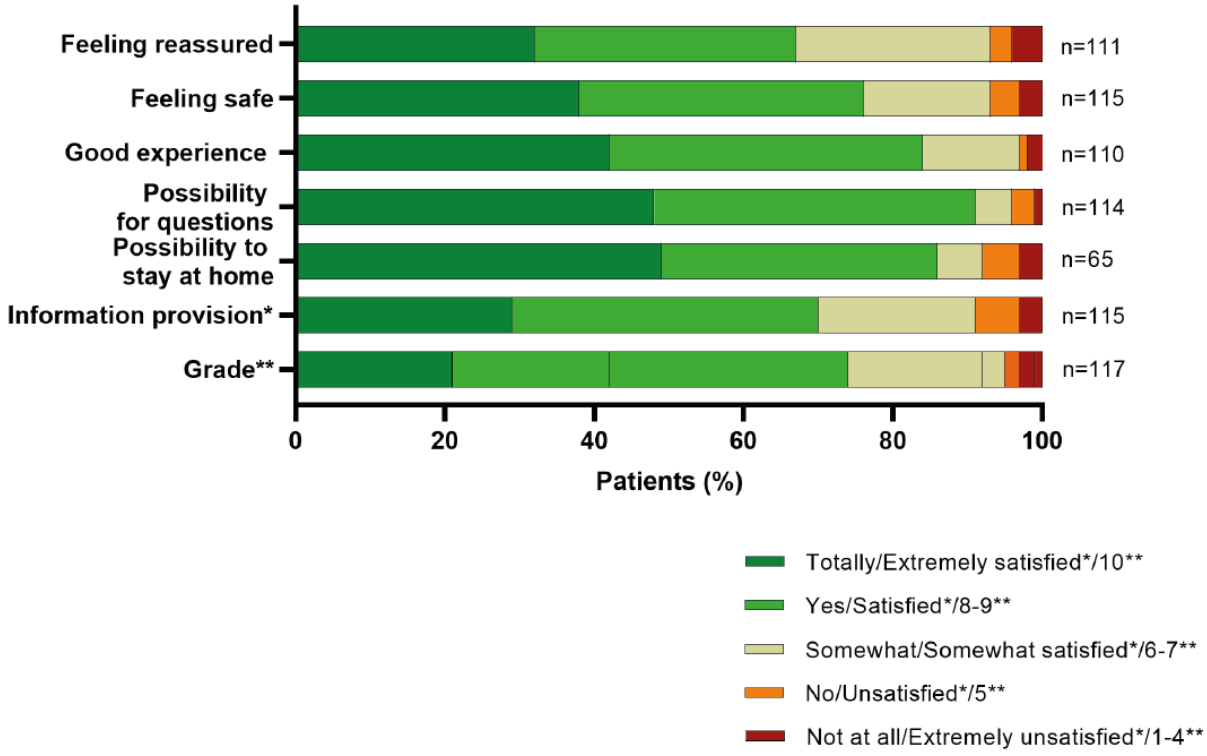
Patients

Of the first 300 patients who were monitored in the COVID-box project from June 2020 to March 2021, a total of 117 (39%) responded to a web-based survey (Multimedia Appendix 1). All of these patients underwent actual telemonitoring with at least 1 contact with a health care professional by using the COVID-box as a home monitoring tool. The results are summarized in Figure 1.

The median score for the sense of reassurance and for the instructions of the COVID-box was 4 (IQR 3-5), with good scores from 67.6% (75/111) and 70% (80/115) of patients, respectively. For the sense of safety, experiences with teleconsultations, the appreciation for staying at home, and the ability to ask questions about the disease and disease course, the median score was 4 (IQR 4-5); good scores were given by 76.5% (88/115), 83.6% (92/110), 86% (56/65), and 91.2% (104/114) of patients, respectively. Overall, 74.4% (87/117) of

patients scored the home monitoring program with the COVID-box as *good*, with a score of 8 or higher.

Figure 1. Patients' valuation. *"Information provision" ratings; **"Grade" ratings.

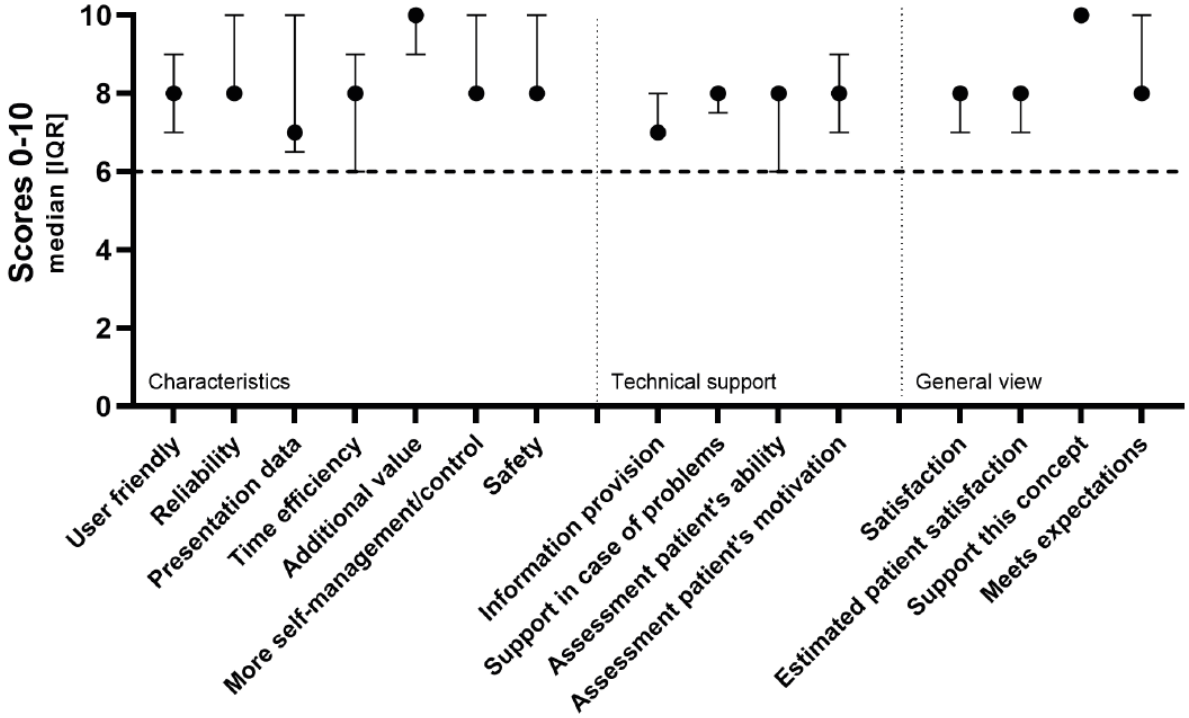


Health Care Professionals

Overview

Of the 60 health care professionals approached, 25 (42%) filled out the questionnaire (Multimedia Appendix 2). Of these, 6 (24%) were specialists who worked at the emergency department, 5 (20%) were specialists from the Department of Internal Medicine (20%), 12 (48%) were residents, and 2 (8%) were nurse practitioners. The results are summarized in Figure 2.

Figure 2. Health care professionals' valuation.



Characteristics of the COVID-box

The median score for the sense of safety, reliability, and more self-management/control in patients was 8 (IQR 8-10), with good scores from 86% (19/22), 89% (8/9), and 92% (22/24) of health care professionals, respectively. For user-friendliness, data presentation, and time efficiency, the median scores were 8 (IQR 7-9), 7 (IQR 7-10), and 8 (IQR 6-9), respectively; good scores were given by 75% (15/20), 44% (4/9), and 55% (6/11) of health care professionals, respectively. The additional value of the COVID-box had a median score of 10 (IQR 8-10), with good scores from 96% (24/25) of health care professionals.

Technical Support Service (COVID-box Team)

For information provision and technical support for problems, the median scores were 7 (IQR 7-8) and 8 (IQR 7.5-8), respectively, with good scores from 37% (8/22) and 76% (13/17) of health care professionals, respectively. The median scores for the assessment of a patient's ability and motivation were 8 (IQR 6-8) and 8 (IQR 7-9), respectively. A total of 57% (13/23) and 76% (16/21) of health care professionals, respectively, scored these items as *good*.

General Satisfaction

The general satisfaction of health care professionals and estimated patient satisfaction had a median score of 8 (IQR 7-8), and 60% (15/25) and 95% (18/19) of health care professionals scored these items as *good*, respectively. With regard to meeting the expectations for this home monitoring concept, good scores were given by 92% (22/24) of health care professionals, with a median score of 8 (IQR 8-10). With regard to supporting this home monitoring concept, the median score was 10 (IQR 10-10), with good scores from all health care professionals (25/25, 100%).

Discussion

Principal Findings

Our study demonstrates that the home monitoring of patients with COVID-19 is well appreciated by patients as well as health care professionals. Previous observational studies have shown the safety of the remote telemonitoring of patients with (suspected) COVID-19 and its efficacy in reducing hospitalization. Few studies have addressed user experience and patients' and health care professionals' satisfaction with telemonitoring. It is well established that the successful implementation of novel eHealth solutions is critically dependent on the positive experiences and satisfaction of involved users.

Comparison With Prior Work

During the extraordinary situation of the COVID-19 pandemic, remote telemonitoring has been quickly implemented in different ways [5-7,10-13]. In general, many patients and health care professionals are very positive about this concept. Remote telemonitoring is used for conducting disease triage; reducing hospital admissions; and providing reassurance, disease and disease course information, and psychological support to clinically stable patients at home. Several studies have reported that patients appreciate all forms of remote telemonitoring (eg, the measuring of vital parameters, symptom recording, and daily teleconsultations) in various settings (eg, in primary care and after an emergency department visit or hospital admission) [5-7,10-13]. The important aspects are easy access to the program, good information provision, and the good quality of the service for the onboarding process. Importantly, older age does not seem to be a problem, as different studies have successfully included patients aged >50 years [5,7,11]. Patients have pointed out that video consultations are also highly appreciated. Self-evidently, the adherence of patients to telemonitoring is critical to its success. Nonadherence to telemonitoring among patients with COVID-19 has been reported when they feel too sick, forget to measure vital parameters, feel insufficiently informed, or experience quick improvements in disease symptoms [13]. Health care professionals are largely convinced of the benefits of remote telemonitoring, as long as a program is easy to use and it is possible to receive patient data correctly.

Limitations

Given the observational and retrospective nature of our study, which was conducted during the COVID-19 pandemic, our study has several limitations that are noteworthy. First, the patient-reported experience measure questionnaires on telemonitoring a new disease, such as COVID-19, were not validated, and a formal validation of these questionnaires was out of the scope of this study. Second, the questionnaires were completed anonymously; therefore, internal consistency could not be reliably assessed. Lastly, the relatively low response rates (patients: 117/300, 39%; health care professionals: 25/60, 42%) could have introduced unwanted bias to the results of this study.

Conclusion

In conclusion, the home monitoring of patients with COVID-19 is well appreciated by patients as well as health care professionals. This study demonstrated that patients felt safe and reassured with the home monitoring and daily teleconsultations in the COVID-box project. Additionally, health care professionals were satisfied with the safety and user-friendliness of the COVID-box. The acceptation of the COVID-box is critical for the successful implementation and expansion of home monitoring for patients with COVID-19 to relieve the burden on health care systems. Our findings could be especially relevant to the current perspectives on oral antiviral agents for the out-of-hospital treatment of patients with COVID-19.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire for patients.
[DOCX File , 15 KB - formative_v6i7e38263_app1.docx]

Multimedia Appendix 2
Questionnaire for healthcare professionals.
[DOCX File , 15 KB - formative_v6i7e38263_app2.docx]

References

1. Vindrola-Padros C, Singh KE, Sidhu MS, Georghiou T, Sherlaw-Johnson C, Tomini SM, et al. Remote home monitoring (virtual wards) for confirmed or suspected COVID-19 patients: a rapid systematic review. EClinicalMedicine 2021 Jul;37:100965 [FREE Full text] [doi: 10.1016/j.eclinm.2021.100965] [Medline: 34179736]
2. 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]
3. Dirikgil E, Roos R, Groeneveld GH, Heringhaus C, Silven AV, Petrus AHJ, et al. Home monitoring reduced short stay admissions in suspected COVID-19 patients: COVID-box project. Eur Respir J 2021 Aug 05;58(2):2100636 [FREE Full text] [doi: 10.1183/13993003.00636-2021] [Medline: 33795321]
4. Casariego-Vales E, Blanco-López R, Rosón-Calvo B, Suárez-Gil R, Santos-Guerra F, Dobao-Feijoo MJ, Telea-Covid Lugo Comanagement Team. Efficacy of telemedicine and telemonitoring in at-home monitoring of patients with COVID-19. J Clin Med 2021 Jun 29;10(13):2893. [doi: 10.3390/jcm10132893] [Medline: 34209725]
5. Wurzer D, Spielhagen P, Siegmann A, Gercekcioglu A, Gorgass J, Henze S, et al. Remote monitoring of COVID-19 positive high-risk patients in domestic isolation: A feasibility study. PLoS One 2021 Sep 24;16(9):e0257095 [FREE Full text] [doi: 10.1371/journal.pone.0257095] [Medline: 34559832]
6. Drewett GP, Holmes NE, Trubiano JA, Vogrin S, Feldman J, Rose M. COVID-Care - a safe and successful digital self-assessment tool for outpatients with proven and suspected coronavirus-2019. Digit Health 2021 Sep 27;7:20552076211047382 [FREE Full text] [doi: 10.1177/20552076211047382] [Medline: 34868615]
7. Clarke J, Flott K, Crespo RF, Ashrafian H, Fontana G, Benger J, et al. Assessing the safety of home oximetry for COVID-19: a multisite retrospective observational study. BMJ Open 2021 Sep 14;11(9):e049235 [FREE Full text] [doi: 10.1136/bmjopen-2021-049235] [Medline: 34521666]
8. Dinh A, Mercier JC, Jaulmes L, Artigou JY, Juillière Y, Yordanov Y, AP-HP/Universities/INSERM COVID-19 Research Collaboration. Safe discharge home with telemedicine of patients requiring nasal oxygen therapy after COVID-19. Front Med (Lausanne) 2021 Nov 03;8:703017 [FREE Full text] [doi: 10.3389/fmed.2021.703017] [Medline: 34805196]
9. Pimlott N, Agarwal P, McCarthy LM, Luke MJ, Hum S, Gill S, et al. Clinical learnings from a virtual primary care program monitoring mild to moderate COVID-19 patients at home. Fam Pract 2021 Sep 25;38(5):549-555 [FREE Full text] [doi: 10.1093/fampra/cmaa130] [Medline: 33340398]
10. Panicacci S, Donati M, Lubrano A, Vianello A, Ruiu A, Melani L, et al. Telemonitoring in the Covid-19 era: The Tuscany Region experience. Healthcare (Basel) 2021 Apr 29;9(5):516 [FREE Full text] [doi: 10.3390/healthcare9050516] [Medline: 33946633]
11. 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 Sep 27;5(9):e20131 [FREE Full text] [doi: 10.2196/20131] [Medline: 34449404]
12. Lim HM, Abdullah A, Ng CJ, Teo CH, Valliyappan IG, Hadi HA, et al. Utility and usability of an automated COVID-19 symptom monitoring system (CoSMoS) in primary care during COVID-19 pandemic: A qualitative feasibility study. Int J Med Inform 2021 Nov;155:104567 [FREE Full text] [doi: 10.1016/j.ijmedinf.2021.104567] [Medline: 34536808]
13. Kerr C, O' Regan S, Creagh D, Hughes G, Geary U, Colgan MP, et al. Acceptability of and symptom findings from an online symptom check-in tool for COVID-19 outpatient follow-up among a predominantly healthcare worker population. BMJ Open 2021 Sep 28;11(9):e050444 [FREE Full text] [doi: 10.1136/bmjopen-2021-050444] [Medline: 34588254]

Abbreviations

LUMC: Leiden University Medical Center

Edited by A Mavragani; submitted 25.03.22; peer-reviewed by B McKinstry, A Torad; comments to author 04.05.22; revised version received 15.06.22; accepted 20.06.22; published 28.07.22.

Please cite as:
Dirikgil E, Brons K, Duindam M, Groeneveld GH, Geelhoed JJM, Heringhaus C, van der Boog PJM, Rabelink TJ, Bos WJW, Chavannes NH, Atsma DE, Teng YKO
COVID-box Experiences of Patients and Health Care Professionals (COVID-box Project): Single-Center, Retrospective, Observational Study
JMIR Form Res 2022;6(7):e38263
URL: https://formative.jmir.org/2022/7/e38263
doi:10.2196/38263
PMID:35816688

Corrigenda and Addenda

Correction:The Associations Between Racially/Ethnically Stratified COVID-19 Tweets and COVID-19 Cases and Deaths: Cross-sectional Study

Xiaohui Liu[1,2], PhD; Bandana Kar[3], PhD; Francisco Alejandro Montiel Ishino[1], MPH, PhD; Tracy Onega[2], MPAS, PhD; Faustine Williams[1], MPH, PhD

[1]National Institute on Minority Health and Health Disparities, National Institutes of Health, Bethesda, MD, United States

[2]Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, United States

[3]National Security Sciences Directorate, Oak Ridge National Lab, Knoxville, TN, United States

Corresponding Author:
Xiaohui Liu, PhD
National Institute on Minority Health and Health Disparities
National Institutes of Health
6707 Democracy Boulevard, Suite 800
Bethesda, MD, 20892
United States
Phone: 1 8016464179
Email: xiaohuiliugis@gmail.com

Related Article:

Correction of: https://formative.jmir.org/2022/5/e30371

(JMIR Form Res 2022;6(7):e40365) doi:10.2196/40365

In "The Associations Between Racially/Ethnically Stratified COVID-19 Tweets and COVID-19 Cases and Deaths: Cross-sectional Study" (JMIR Form Res 2022 May 30;6(5):e30371. doi: 10.2196/30371), one error was noted.

In the original article author Faustine Williams was incorrectly associated with affiliation 2:

Faustine Williams[2], MPH, PhD

[2]Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, United States.

The correct affiliation should be affiliation 1:

Faustine Williams[1], MPH, PhD

[1]National Institute on Minority Health and Health Disparities, National Institutes of Health, Bethesda, MD, United States.

The correction will appear in the online version of the paper on the JMIR Publications website on July 6, 2022, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 17.06.22; this is a non–peer-reviewed article;accepted 21.06.22; published 06.07.22.

Please cite as:
Liu X, Kar B, Montiel Ishino FA, Onega T, Williams F
Correction:The Associations Between Racially/Ethnically Stratified COVID-19 Tweets and COVID-19 Cases and Deaths: Cross-sectional Study
JMIR Form Res 2022;6(7):e40365
URL: https://formative.jmir.org/2022/7/e40365
doi:10.2196/40365
PMID:35793543

Corrigenda and Addenda

Correction: A Telehealth-Delivered Tai Chi Intervention (TaiChi4Joint) for Managing Aromatase Inhibitor–Induced Arthralgia in Patients With Breast Cancer During COVID-19: Longitudinal Pilot Study

Sameh Gomaa[1], MD; Carly West[1], MPH; Ana Maria Lopez[1], MD; Tingting Zhan[2], PhD; Max Schnoll[1], BA; Maysa Abu-Khalaf[1], MD; Andrew Newberg[3], MD; Kuang-Yi Wen[1], PhD

[1]Department of Medical Oncology, Thomas Jefferson University, Philadelphia, PA, United States

[2]Department of Pharmacology & Experimental Therapeutics, Thomas Jefferson University, Philadelphia, PA, United States

[3]Department of Integrative Medicine and Nutritional Sciences, Thomas Jefferson University, Philadelphia, PA, United States

Corresponding Author:
Kuang-Yi Wen, PhD
Department of Medical Oncology
Thomas Jefferson University
834 Chestnut Street Suite 300
Philadelphia, PA, 19106
United States
Phone: 1 2155034623
Email: kuang-yi.wen@jefferson.edu

Related Article:

Correction of: https://formative.jmir.org/2022/6/e34995

(JMIR Form Res 2022;6(7):e40830) doi:10.2196/40830

In "A Telehealth-Delivered Tai Chi Intervention (TaiChi4Joint) for Managing Aromatase Inhibitor–Induced Arthralgia in Patients With Breast Cancer During COVID-19: Longitudinal Pilot Study" (JMIR Form Res 2022;6(6): e34995), the authors made the following corrections in the corresponding authorship:

1. In the originally published article, author Sameh Gomaa was erroneously listed as the corresponding author. The corresponding authorship is now correctly attributed to Kuang-Yi Wen.

2. Accordingly, the address and contact details of the corresponding author have been changed as follows:

Kuang-Yi Wen, PhD
Department of Medical Oncology
Thomas Jefferson University
834 Chestnut Street Suite 300
Philadelphia, PA, 19106
United States
Phone: 1 2155034623

3. In the originally published article, author Tingting Zhan was incorrectly associated with the following affiliation:

Department of Medical Oncology, Thomas Jefferson University, Philadelphia, PA, United States

The author is now correctly associated with newly added Affiliation 2:

Department of Pharmacology & Experimental Therapeutics, Thomas Jefferson University, Philadelphia, PA, United States

4. In the originally published article, author Andrew Newberg was incorrectly associated with the following affiliation:

Department of Medical Oncology, Thomas Jefferson University, Philadelphia, PA, United States

The author is now correctly associated with newly added Affiliation 3:

Department of Integrative Medicine and Nutritional Sciences, Thomas Jefferson University, Philadelphia, PA, United States

The correction will appear in the online version of the paper on the JMIR Publications website on July 19, 2022, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 07.07.22; this is a non–peer-reviewed article;accepted 07.07.22; published 19.07.22.

Please cite as:
Gomaa S, West C, Lopez AM, Zhan T, Schnoll M, Abu-Khalaf M, Newberg A, Wen KY
Correction: A Telehealth-Delivered Tai Chi Intervention (TaiChi4Joint) for Managing Aromatase Inhibitor–Induced Arthralgia in Patients With Breast Cancer During COVID-19: Longitudinal Pilot Study
JMIR Form Res 2022;6(7):e40830
URL: https://formative.jmir.org/2022/7/e40830
doi:10.2196/40830
PMID:35853219

Corrigenda and Addenda

Correction: Evaluating the Quality of Asynchronous Versus Synchronous Virtual Care in Patients With Erectile Dysfunction: Retrospective Cohort Study

Lauren Broffman[1], DPhil; Melynda Barnes[1], MD; Kevin Stern[1], BS; Amy Westergren[1], JD

Ro, New York, NY, United States

Corresponding Author:
Lauren Broffman, DPhil
Ro
116 West 23rd Street
New York, NY, 10011
United States
Phone: 1 888 798 8686
Email: lauren.broffman@ro.co

Related Article:

Correction of: https://formative.jmir.org/2022/1/e32126

(JMIR Form Res 2022;6(7):e41121) doi:10.2196/41121

In "Evaluating the Quality of Asynchronous Versus Synchronous Virtual Care in Patients With Erectile Dysfunction: Retrospective Cohort Study" (JMIR Form Res 2022;6(1):e32126), the authors noted the following corrections:

1. In Table 1, the values for each side effect - headache, dizziness, flushing, congestion, dyspepsia, back pain, and blurry vision - were mislabeled. Values categorized as 'asynchronous' should have been labeled 'synchronous' and vice versa. No values were incorrectly referenced in the text of the article and this error does not impact the overall interpretation of results.

2. In the original paper, the following sentence was present in the *Discussion* section:

> *The recent widespread adoption of telehealth as an acceptable treatment modality and the exploration for potential expansion of traditional methods outside the scope of synchronous care have prompted deeper exploration of the downstream effects of these approaches to care distribution.*

This has been changed to:

> *The recent widespread adoption of telehealth as an acceptable treatment modality and the potential expansion of asynchronous care have prompted deeper exploration of the downstream effects.*

3. The phone number of the corresponding author, Lauren Broffman, has been updated to 1 888 798 8686.

Table 1. Rates of reported side effects by modality.

Side effects	Synchronous (n=2150)	Asynchronous (n=7850)
Any side effect, n (%)	24 (1.12)	113 (1.44)
Headache, n (%)	10 (0.47)	56 (0.71)
Dizziness, n (%)	0 (0)	3 (0.04)
Flushing, n (%)	2 (0.09)	31 (0.39)
Congestion, n (%)	12 (0.56)	17 (0.22)
Dyspepsia, n (%)	7 (0.33)	10 (0.13)
Back pain, n (%)	0 (0)	6 (0.08)
Blurry vision, n (%)	2 (0.09)	8 (0.10)
Other, n (%)	2 (0.09)	7 (0.09)

The correction will appear in the online version of the paper on the JMIR Publications website on July 27, 2022, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text

repositories, the corrected article has also been resubmitted to those repositories.

Submitted 15.07.22; this is a non–peer-reviewed article;accepted 15.07.22; published 27.07.22.

Please cite as:
Broffman L, Barnes M, Stern K, Westergren A
Correction: Evaluating the Quality of Asynchronous Versus Synchronous Virtual Care in Patients With Erectile Dysfunction: Retrospective Cohort Study
JMIR Form Res 2022;6(7):e41121
URL: https://formative.jmir.org/2022/7/e41121
doi:10.2196/41121
PMID:35896026

Original Paper

Exploring Public Perceptions of Dental Care Affordability in the United States: Mixed Method Analysis via Twitter

Shahen Yashpal[1*], MSc; Ananditha Raghunath[2*], BA; Nihan Gencerliler[3], MPH; Lorel E Burns[3], MS, DDS

[1]College of Dentistry, New York University, New York, NY, United States

[2]Computer Science & Engineering, University of Washington, Seattle, WA, United States

[3]Department of Endodontics, College of Dentistry, New York University, New York, NY, United States

[*]these authors contributed equally

Corresponding Author:
Lorel E Burns, MS, DDS
Department of Endodontics
College of Dentistry
New York University
345 E 24th Street
New York, NY, 10010
United States
Phone: 1 212 998 9332
Email: leb409@nyu.edu

Abstract

Background: Dental care expenses are reported to present higher financial barriers than any other type of health care service in the United States. Social media platforms such as Twitter have become a source of public health communication and surveillance. Previous studies have demonstrated the usefulness of Twitter in exploring public opinion on aspects of dental care. To date, no studies have leveraged Twitter to examine public sentiments regarding dental care affordability in the United States.

Objective: The aim of this study is to understand public perceptions of dental care affordability in the United States on the social media site, Twitter.

Methods: Tweets posted between September 1, 2017, and September 30, 2021, were collected using the Snscrape application. Query terms were selected a priori to represent dentistry and financial aspects associated with dental treatment. Data were analyzed qualitatively using both deductive and inductive approaches. In total, 8% (440/5500) of all included tweets were coded to identify prominent themes and subthemes. The entire sample of included tweets were then independently coded into thematic categories. Quantitative data analyses included geographic distribution of tweets by state, volume analysis of tweets over time, and distribution of tweets by content theme.

Results: A final sample of 5314 tweets were included in the study. Thematic analysis identified the following prominent themes: (1) general sentiments (1614 tweets, 30.4%); (2) delaying or forgoing dental care (1190 tweets, 22.4%); (3) payment strategies (1019 tweets, 19.2%); (4) insurance (767 tweets, 14.4%); and (5) policy statements (724 tweets, 13.6%). Geographic distributions of the tweets established California, Texas, Florida, and New York as the states with the most tweets. Qualitative analysis revealed barriers faced by individuals to accessing dental care, strategies taken to cope with dental pain, and public perceptions on aspects of dental care policy. The volume and thematic trends of the tweets corresponded to relevant societal events, including the COVID-19 pandemic and debates on health care policy resulting from the election of President Joseph R. Biden.

Conclusions: The findings illustrate the real-time sentiment of social media users toward the cost of dental treatment and suggest shortcomings in funding that may be representative of greater systemic failures in the provision of dental care. Thus, this study provides insights for policy makers and dental professionals who strive to increase access to dental care.

(JMIR Form Res 2022;6(7):e36315) doi:10.2196/36315

KEYWORDS

dentistry; oral health; social media; access to care; healthcare reform; COVID-19; dental care; health care service; twitter; public health; health communication; dental treatment; health policy; dental professional; thematic analysis

Introduction

A lack of access to dental care can lead to lower levels of systemic health, quality of life, and economic outcomes [1]. Yet, those who are most in need of dental care are often the least likely to receive it [2]. Low-income, working-age adults report the highest levels of financial barriers to needed dental care [3]. In the United States, the financial barriers to accessing dental care are higher than any other type of health care service [3-5]. The percentage of the population without dental insurance is more than twice that of those who are medically uninsured [6]. Spending on dental care results in a high percentage of out-of-pocket expenses because of a lack of insurance or high insurance deductibles and copayments [7].

While comprehensive dental coverage for children is an essential benefit under the Affordable Care Act, dental coverage for adults remains optional [7]. At the time of this study, 3 states provide no dental coverage, and 12 states provide emergency-only dental services to Medicaid beneficiaries [8]. In a comparison of public dental coverage for older adults in high-income countries, the United States had the shallowest dental coverage for older adults [9]. In an effort to mitigate access inequalities, there was a recent push in the US federal government to provide for Medicare coverage of dental and oral health services. In January of 2021, the Medicare Dental Benefit Act of 2021 (H.R.502 and S.97) was introduced into Congress, and President Biden's budget-reconciliation package proposed funding for a Medicare dental benefit. In August 2021, the Centers for Medicare and Medicaid Services appointed a chief dental officer to guide it in advancing oral health in Medicare [10]. To best develop a policy to address financial barriers to dental care, perspectives at the individual level are needed.

Social media is increasingly becoming an essential tool for public health communication [11]. Twitter is a free social media service where people communicate their daily thoughts and behaviors in short, 280-character messages called "tweets." With over 68 million active monthly users in the United States, Twitter offers rich, population-based data for tracking concerns of public health significance [12]. Twitter data emerge from real-world social environments, which encompass a large and diverse range of people, without any prompting from researchers. This contrasts with traditional approaches of public surveillance where responses are elicited in the form of semistructured interviews and web-based surveys with open-ended questions [13]. In addition, Twitter is a compelling data source for public health researchers because of the real-time nature of the content and high level of correlation with user sentiment and consumer confidence indices [14]. These qualities have contributed to a growing number of studies examining the use of Twitter for public health research [15-21], including the investigation of aspects of oral health [22-26]. Several studies have aimed to assess the influence of societal events on Twitter content related to oral health [27,28].

This study aims to explore the sentiments of Twitter users in the United States on dental care affordability in order to summarize trends and perceptions that describe how the cost of dental treatment impacts access to dental care. To date, no studies have leveraged Twitter to examine public sentiments regarding dental care affordability in the United States.

Methods

Ethical Considerations

This infodemiological study used a convergent mixed methods approach [29,30] to analyze publicly available Twitter content related to dental care affordability. This study was submitted to the Institutional Review Boards of New York University (IRB-FY2021-5634) and the University of Washington (STUDY00013725). In alignment with federal regulations regarding the use of publicly available data for research, both institutional reviews determined that the study did not meet the criteria for research involving human subjects and that no further review was required.

Data Collection and Preprocessing

Data were obtained from Twitter, a free social media website created in 2006. Tweets can remain visible to the public or can be made visible only to approved followers, at the discretion of the user. Only publicly available tweets were used in this study, and usernames were removed for privacy protection.

Search terms were generated to identify tweets that discussed financial considerations associated with dental treatment. These search terms were tested and expanded through pilot queries and assessments until they were refined to the following word stems: "dental," "dentist," "tooth," "teeth," "root canal" AND "expensive," "pay," "afford," and "money."

Using a newly created account on Twitter, tweets were collected between September 1, 2017, and September 17, 2021, using Snscrape [31-33], an open-source web scraper written in Python (Python Software Foundation). The data collection script included a specific query that consisted of any combination of the search terms. Duplicate tweets, foreign language tweets, and retweets were then excluded. In addition to the tweets, metadata was collected as follows: "*url*," "*date*," "*renderedContent*," "*user*," "*replyCount*," "*retweetCount*," "*likeCount*," "*quoteCount*," "*retweetedTweet*," "*quotedTweet*," "*mentionedUsers*," "*coordinates*," "*place*," "*hashtags*," and "*cashtags*."

Metadata related to the users' location were used in order to limit the included tweets to those in the United States. Content was excluded from the data set of tweets for any of the following reasons: (1) content was unrelated to dental-treatment needs or experiences; (2) content was determined to be an advertisement; (3) content pertained to purely cosmetic or orthodontic dental procedures; (4) content pertained to veterinary dentistry; (5) content was classified as a joke or sarcasm; and (6) content was in reference to dental policies outside of the United States.

Data Analysis

Data were analyzed qualitatively using both deductive and inductive approaches. Two members of the research team with clinical dental knowledge (SY and LB) co-coded all of the tweets. In total, 8% (440/5500) of all of the included tweets were coded in order to identify prominent themes and subthemes. A final codebook was developed through consensus

among the members of the research team. The themes identified for the codebook included (1) general sentiments; (2) delaying or forgoing dental care; (3) payment strategies; (4) insurance; and (5) policy statements. Using the codebook, the entire sample of the included tweets was independently coded by the aforementioned researchers. In instances where the coders disagreed, they reached a consensus through discussion.

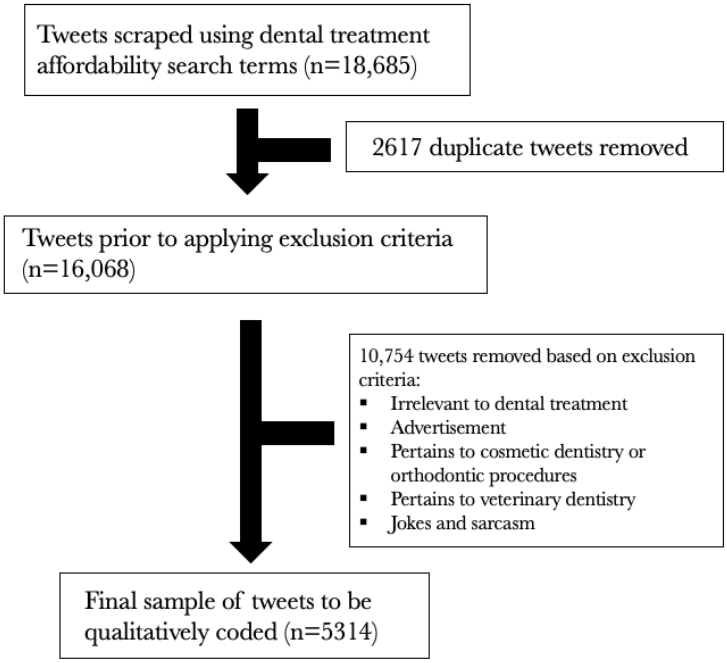
Quantitative data analysis included geographic distribution of tweets by state, volume analysis of tweets over time, and distribution of tweets by content theme. The tweets were mapped to an individual state within the United States, and the volume of tweets over time were depicted using Tableau (Tableau Software Inc), a data visualization software.

Results

Data Collection

Using the search terms, we collected a total of 18,685 tweets from September 1, 2017, to September 30, 2021. Of these tweets, 2617 (14%) were removed as duplicates, and 10,754 (57.6%) were removed based on preestablished exclusion criteria. A final sample of 5314 (28.4%) tweets (from 4963 unique users) were included in the study (Figure 1). The top 5 pairs of search terms were as follows: dental_afford (520/5314, 9.8% tweets); teeth_pay (461/5314, 8.7% tweets); dentist_pay (436/5314, 8.2% tweets); dental_money (356/5314, 6.7%, tweets); and dental_expensive (353/5314, 6.6% tweets).

Figure 1. Data collection flow chart. Tweets about dental care affordability.



Thematic Distribution of Tweets

Thematic analysis identified the following prominent themes (Table 1): (1) general sentiments (1614/5314 tweets, 30.4%); (2) delaying or forgoing dental care (1190/5314 tweets, 22.4%); (3) payment strategies (1019/5314 tweets, 19.2%); (4) insurance (767/5314 tweets, 14.4%); and (5) policy statements (724/5314 tweets, 13.6%). The identified subthemes and illustrative quotes from the data are presented below and in Table 2.

Table 1. Volume of tweets over time by thematic category.

Main category	Overall (n=5314), n (%)	2017 (n=277), n (%)	2018 (n=1227), n (%)	2019 (n=1341), n (%)	2020 (n=1090), n (%)	2021 (n=1379), n (%)
General sentiments	1614 (30.4)	110 (39.2)	399 (32.5)	405 (30.2)	320 (29.4)	380 (27.6)
Delaying or forgoing care	1190 (22.4)	64 (23.1)	304 (24.8	286 (21.3)	266 (24.4)	270 (19.6)
Payment strategies	1019 (19.2)	34 (12.3)	206 (16.8)	298 (22.2)	249 (22.8)	232 (16.8)
Insurance	767 (14.4)	44 (15.9)	173 (14.1)	187 (13.9)	155 (14.2)	208 (15.1)
Policy statements	724 (13.6)	25 (9.0)	145 (11.8)	165 (12.3)	100 (9.2)	289 (21.0)

Table 2. Representative tweets for theme or subtheme as described in the codebook.

Theme and subtheme	Tweet
Insurance	• "My dental insurance deductible was so high, I couldn't have dental work done last year. And my knee surgery got cancelled as a result. Damn shame. Still can't afford it. [URL]" • "And... 19's root canal and crown cost more than the annual dental insurance max. There go $837 of tuition money. I'm really just over this whole century." • "According to most health insurance companies teeth and eyes are luxury items that I must pay more to continue enjoying because they're a cosmetic privilege."
Payment strategies	• "@username @username hey. I know it's a long shot but I just had to use my rent money to have a very badly infected tooth removed and now I don't know how I'm doing to pay rent. PLEASE help if you can  God bless ya'll doing God's work"[a] • "Nervous to get my tooth pulled tomorrow but mainly because I'm afraid it's gonna cost more than what I have left on the credit card I'm using to pay for it…."
Delaying or foregoing care	• "I canceled my root canal that was supposed to be Monday because my tooth quit hurting but here we are in ridiculous pain again and my husband has no job so no money to spend on it right now  " • "@username Yes it is I have several teeth I need to get pulled/ worked on but can't afford it..and also take pain meds that don't seem to help much of anything..."[a] • "@username Do you live near a dental school? When I didn't have dental insurance, I did an expensive project at a dental school for about 1/5 the cost. It took more time, but it was worth paying less."[a]
General sentiments	
Dentist mistrust	• "I'm positive the dentist doesn't really find cavities in my teeth they just want to get my money and torture me. How do I brush every day and floss and still get cavities every time"
Dentistry is expensive	• "@username Dental work is horribly expensive. I have had my very last $1700 root canal."[a]
General statements	• "@username @username This is why I never want to hear another American make fun of British teeth again. People in this country cannot afford the astronomical prices charged by dentists for dental care. So they just go without dental care at all."[a]
Positive sentiments	• "i spent a lotta money on my teeth and it was worth it  "
Policy statements	• "@DeptVetAffairs @SenateGOP @HouseGOP I cannot afford regular dental coverage as a disabled veteran with a $200 a month income." • "My mom doesn't have insurance for dental care...now, she has an infection in her tooth... now she needs to pay $1504 before seeing the specialist... we don't have $1504 in the bank... This is why we need Medicare For All!!! I'm angry because of this greedy corruption #Bernie2020" • "@username @username When someone has a dental problem serious enough for a root canal, they are in pain and it will be an emergency treatment. But only people with money or dental insurance can get one. Most states' Medicaid doesn't cover adult dental care, and if it does, they pull the tooth."[a]

[a]Username was removed to maintain the privacy of the Twitter user.

General Sentiments

The most prominent theme included tweets containing general sentiments that expressed that dental care was expensive or not affordable but did not additionally suggest the cost that prohibited the user from accessing needed dental care. The overwhelming majority of these tweets expressed negative sentiments. Tweets in this thematic category were subcoded into the following subthemes: mistrust of dentists, expressions that dental care is expensive for the tweeting individual or a personal acquaintance (eg, a family member), impersonal statements about the affordability of dental care, and positive sentiments about accessing dental care.

The most prevalent subtheme was "expensive dental care," representing 70.3% (n=1135) of the tweets in this theme. Individuals frequently expressed displeasure related to the direct cost of dental care and surcharges related to administrative costs.

> *The consultation alone was $250... out of pocket at that ...why the dentist gotta be so expensive?? I just want to be beautiful and healthy for the low [URL].* [User #677]

Tweets about "dentist mistrust" represented the second most prevalent subtheme of general sentiments. Users expressed sentiments of being financially duped or perceptions of dentists prioritizing financial gain over the patient's health.

> *@username @username Another common practice to make more money, is to remove all of the wisdom teeth, when often times patients do not need all of*

them removed. It's just like, we're in there, so let's do them all. [User #1106]

Delaying or Forgoing Dental Care

Tweets about delaying or forgoing dental care due to an inability to afford treatment were categorized into the second most prominent theme. Strategies resulting from delaying or forgoing dental care mentioned within this theme included dental tourism, visits to the emergency room, visits to free clinics or dental schools, and self-treatment.

Users tweeted about not being able to afford the upfront cost of care and as a result delaying care despite being in pain. In some cases, users shared experiences of death resulting from an inability to access dental care.

@username I had a friend who didn't get treated, he died from sepsis. I don't want to scare you but it can be infected and that can spread really fast. No dentist should refuse someone because they can't pay when it comes to infection. I wish her well. [User #1126]

Users reported visiting the emergency room to address dental pain or infection when they could not afford treatment by a dentist. They often expressed dissatisfaction with their experience, which often left them with palliative care such as pain killers and antibiotics rather than treatment. In some cases, visits to the emergency room resulted in unexpected costs when medical insurances denied claims for dental-related conditions.

@username will not cover the ER visit because dental related, and I will have to pay close to $2000 hospital bill. All they did was examine me and prescribe painkiller and antibiotics. How is this right? Everything needs to change. [User #4995]

Travel to countries with lower-cost dental care was expressed as a way to access care when dental treatment in the United States was considered to be too expensive. Domestically, users considered going to dental schools or public dental clinics when they could not afford the cost of a private dentist.

The fact that I'm driving three hours into Mexico tomorrow to get my teeth done at an 1/8 of the price I was going to pay at a "nonprofit" dentist in America is fucking ridiculous. [User #1990]

Payment Strategies

The theme about "payment strategies" for needed dental care most frequently included tweets asking for donations from others. Certain twitter accounts were frequently referenced in donation requests including @pulte, @TeamPulte, and @JefferyStar. Twitter users often asked for funds to be sent to them via PayPal, Venmo, GoFundMe, and CashApp.

I'm raising money to have my teeth pulled. Click to Donate: [URL] via @gofundme. [User #241]

A limited number of tweets indicated having to make sacrifices, such as foregoing groceries or rent, to afford dental care. Other payment strategies included going into debt and using stimulus checks or tax refunds to pay for care.

Really hoping I get stimulus money. I have teeth that hurt and dentists are expensive. [User #1091]

Insurance

Tweets discussing experiences with private-payer dental insurance were included in this less prominent theme. Many tweets discussed dental insurance as an employee benefit, and some expressed gratitude for occupation-granted insurance facilitating access to dental care.

However, tweets about insurance for dental care expressed negative sentiments more frequently, including a fear of losing coverage or not being able to afford care despite having insurance. In certain instances, individuals expressed displeasure with the complexity of the insurance system and dissatisfaction for having to pay for a portion of dental care in addition to their monthly insurance premium. Some users also expressed displeasure with dental care being excluded from medical insurance.

@username I have fancy employee dental insurance, but still have to pay a bunch for crowns and root canals with the added bonus of being really limited in dentists who accept the insurance. Instead of paying $1,000 for a crown I pay $300. So basically paying $30/month for a discount card. [User #4934]

Policy Statements

Tweets categorized as "policy statements" were the least prominent theme. The included tweets were those that mentioned Medicaid or Medicare, health care reform, were directed at politicians, or were about veterans. The content of these tweets was often related to the desired policy changes related to dental coverage.

Time and Geographic Distribution of Tweets

Overall, 5314 tweets about dental affordability were collected (Table 1), with 277 (5.2%) in 2017, 1227 (23.1%) in 2018, 1341 (25.2%) in 2019, 1090 (20%) in 2020, and 1379 (26%) in 2021. The period of data collection was less than a year in 2017 (September 1 to December 31; 4 months) and 2021 (January 1 to September 30; 9 months).

The volume of tweets over the study period is depicted in Figure 2. There was a monthly average of 116 included tweets. The lowest number of tweets was observed in April 2020 (58 tweets, 1.1%), and the highest number of tweets was observed in September 2021 (673 tweets, 12.7%). Overall and across each year of collection, "general sentiments" was the most prevalent thematic category (Table 1). "Delaying or foregoing care" was the second most prevalent thematic category, except in 2019 and 2021, when "payment strategies" and "policy statements," respectively, were the more frequently observed categories. "Policy statements" was consistently the least frequently observed thematic category, except for a significant increase in 2021.

The geographic location of tweets was determined and mapped in Figure 3. Tweets originated the most from the most populous states in the United States, which are California, Texas, Florida, and New York [34].

Figure 2. Volume of tweets by month and year.

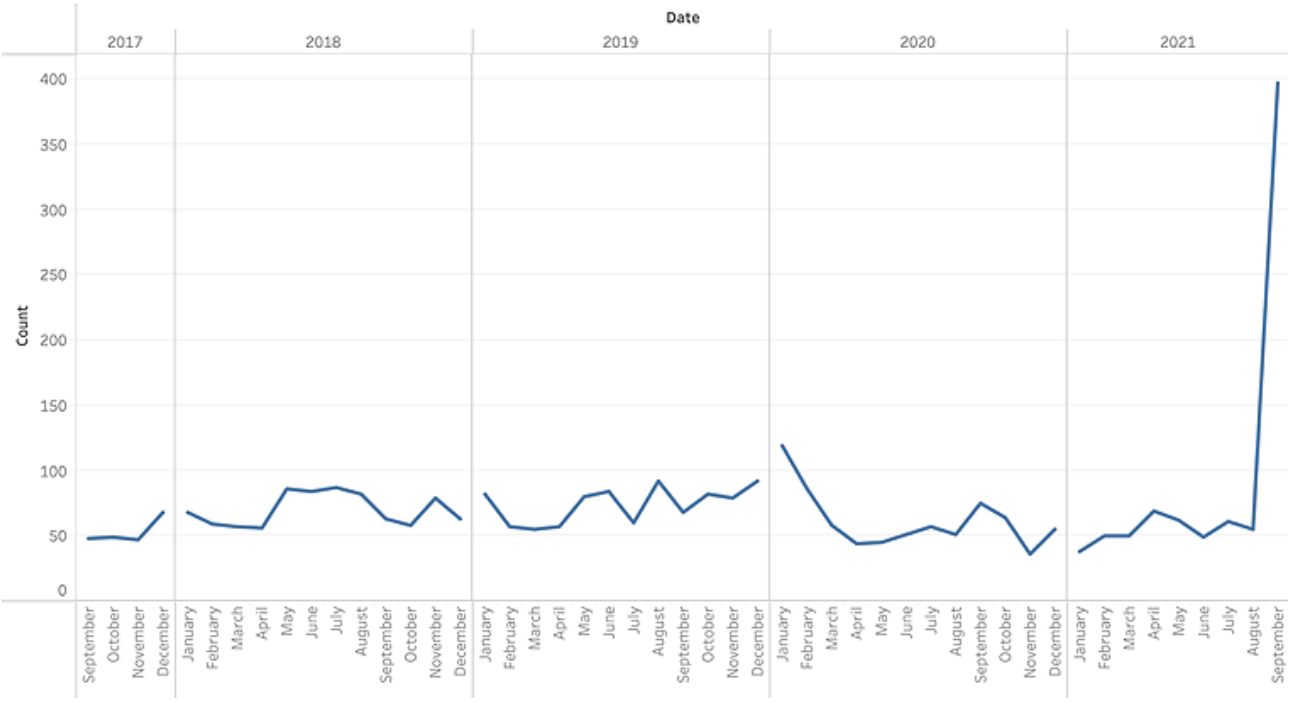
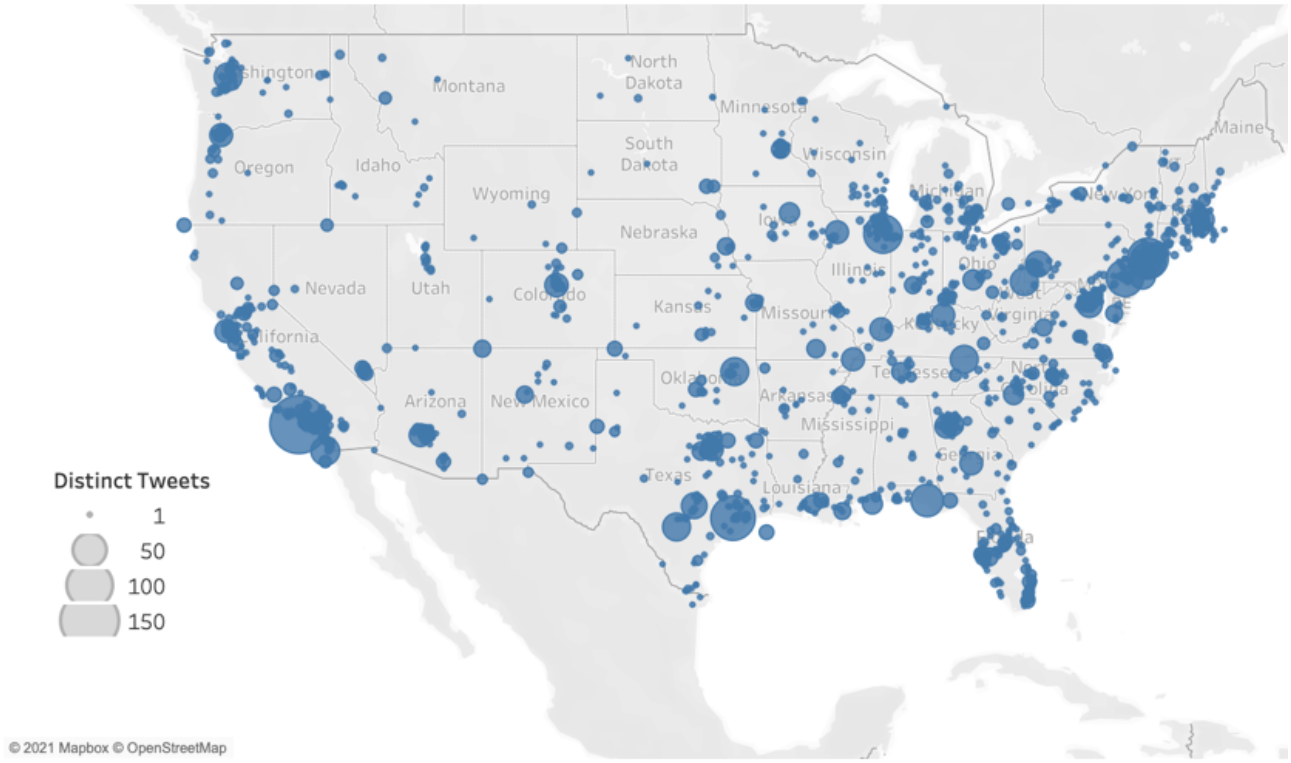


Figure 3. Map of tweet geolocations. Size of the dot represents the volume of tweets at a given latitude-longitude pair.



Discussion

Principal Results and Comparisons With Prior Work

This study leveraged Twitter to examine public sentiments toward dental care affordability in the United States. Twitter users expressed dissatisfaction with the cost of dental treatment and the ability to access dental care. The overwhelmingly negative sentiments found in this study provide insight into how individuals are coping with financial barriers, including delaying or foregoing care and pursuing various payment strategies. Our

study findings support the conclusions of the existing dental literature on dental care affordability [2-5,7]. In the 2019 National Health Interview Survey, it was found that 19.2% of women and 15.6% of men did not access their needed dental care because of cost in the prior 12 months [35]. Rates of forgoing dental care are particularly high for uninsured adults, where 1 in 2 had not seen a dentist because of costs [36].

Moreover, our study findings highlight the fact that financial barriers to accessing dental care are prevalent even among those with dental insurance. The expressed sentiments of discontent with dental insurance payment structures echoed those of the existing reports in the literature. In a study on US health care spending, among 154 conditions examined, oral disorders requiring dental care had the highest out-of-pocket costs [37]. Another study that supports the sentiments of insured tweeters in this study reported that among individuals who were insured all year, US adults were significantly more likely than adults in other developed countries to go without care because of costs, the possibility of facing high out-of-pocket spending, or the financial burden of medical bills [36].

In an effort to come up with the money required to be paid out of pocket, users reported a variety of payment strategies that included credit cards, loans, stimulus checks, tax refunds, and donations. Crowdfunding, the web-based solicitation of public donations, has become a major financer of health care–related costs [38-41]. This trend was reflected in our study findings and reinforces sentiments that the use of crowdfunding to cover direct health care expenses may be a sign of a failing system [42]. Future research may explore crowdfunding for dental procedures.

Two relevant societal events occurred during the study period, which were the COVID-19 pandemic and the election of President Biden. COVID-19 resulted in dental office closures across the country in the spring and summer of 2020, with the greatest decline in weekly visits compared with 2019 observed in the week of April 12, 2020 [43]. This trend corresponds with the quantitative findings of this study, in which the lowest number of tweets were observed in April 2020. Further, COVID-19 had an economic impact that appeared to exacerbate financial barriers to accessing dental care for some people. One user tweeted, "…I can't wait to get an eye exam and new glasses, not to mention long overdue dental care! Or just not because I can't afford it … because of #COVID19". For others, economic stimulus checks distributed during the COVID-19 pandemic helped reduce financial barriers to dental care; for example, "@politico Got my stimulus money today. Thinking of getting much-needed dental work. Several teeth fell out when I had COVID in Dec. and Jan. Thank you President Joe!" Lastly, once dental offices reopened after COVID-19–related closures, there were tweets expressing displeasure with increased costs of care due to safety measures; for example, "…Had to pay an up charge for PPE for a dental appointment over the summer. Was told insurance was not likely to cover that aspect of the cleaning...." The election of President Biden in 2021 renewed conversations about health care reform, and the volume of tweets coded as "policy statements" increased in correspondence. In September 2021, there were 3.4 times the monthly average number of tweets. We hypothesize that this increase is a reaction to the announcement in August 2021 that President Joe Biden's budget-reconciliation package included funding for a standard Medicare dental benefit [44-46]. Ultimately, the Medicare dental benefit was not included in the House of Representatives' passed legislation.

Limitations

The trends in the volume of tweets over time as well as the geographic distribution of the tweets support the generalizability and transferability of the findings of this study. However, this study is not without limitations. While all demographic categories have been shown to engage with social media to varying extents in the United States, Twitter users are not necessarily representative of the US population [47]. On average, Twitter users are younger, are more likely to identify as Democrats, are more highly educated, and have higher incomes than US adults overall [48]. Further, the sample of tweets collected may have some sampling bias, as the data set was not necessarily limited to a single tweet per unique user.

Conclusions

The findings illustrate the real-time sentiment of Twitter users toward the cost of dental treatment and suggest shortcomings in funding, which may be representative of a greater systemic failure in the provision of dental care. Thus, this study provides insights for policy makers and dental professionals who strive to increase access to dental care. Limitations of dental insurance payment models, both public and private, are one such area that may be explored.

Acknowledgments

The authors acknowledge Jennifer Gibbs, MAS, DDS, PhD for her guidance on the pilot study for this project.

Conflicts of Interest

None declared.

References

1. Oral Health in America: A Report of the Surgeon General. Department of Health and Human Services. 2000. URL: https://www.nidcr.nih.gov/sites/default/files/2017-10/hck1ocv.%40www.surgeon.fullrpt.pdf [accessed 2022-06-06]
2. Bersell CH. Access to Oral Health Care: A National Crisis and Call for Reform. J Dent Hyg 2017 Feb;91(1):6-14. [Medline: 29118145]

3. Vujicic M, Buchmueller T, Klein R. Dental Care Presents The Highest Level Of Financial Barriers, Compared To Other Types Of Health Care Services. Health Aff (Millwood) 2016 Dec 01;35(12):2176-2182. [doi: 10.1377/hlthaff.2016.0800] [Medline: 27920304]
4. Economic Well-Being of U.S. Households in 2021. Board of Governors of the Federal Reserve System. URL: https://www.federalreserve.gov/newsevents/pressreleases/files/other20220523a1.pdf [accessed 2022-06-06]
5. 2019 National Health Interview Survey (NHIS) Data. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhis/2019nhis.htm [accessed 2022-06-06]
6. Understanding Dental Benefits. National Association of Dental Plans. URL: https://www.whydental.org/about/understanding-dental-benefits [accessed 2022-01-02]
7. Mertz EA. The Dental-Medical Divide. Health Aff (Millwood) 2016 Dec 01;35(12):2168-2175. [doi: 10.1377/hlthaff.2016.0886] [Medline: 27920303]
8. Medicaid Adult Dental Benefits: An Overview. Center for Health Care Stategies. 2019. URL: https://www.chcs.org/media/Adult-Oral-Health-Fact-Sheet_091519.pdf [accessed 2022-06-06]
9. Allin S, Farmer J, Quiñonez C, Peckham A, Marchildon G, Panteli D, et al. Do health systems cover the mouth? Comparing dental care coverage for older adults in eight jurisdictions. Health Policy 2020 Sep;124(9):998-1007 [FREE Full text] [doi: 10.1016/j.healthpol.2020.06.015] [Medline: 32712013]
10. Simon L, Giannobile WV. Is It Finally Time for a Medicare Dental Benefit? N Engl J Med 2021 Dec 02;385(23):e80. [doi: 10.1056/nejmp2115048] [Medline: 34670037]
11. Schein R, Wilson K, Keelan J. Literature Review on Effectiveness of the Use of Social Media: a Report for Peel Public Health. Region of Peel. 2010. URL: https://www.peelregion.ca/health/resources/pdf/socialmedia.pdf [accessed 2022-06-06]
12. Number of monthly active Twitter users in the United States from 1st quarter 2010 to 1st quarter 2019. Statista. 2019. URL: https://www.statista.com/statistics/274564/monthly-active-twitter-users-in-the-united-states/ [accessed 2022-06-06]
13. Andreotta M, Nugroho R, Hurlstone MJ, Boschetti F, Farrell S, Walker I, et al. Analyzing social media data: A mixed-methods framework combining computational and qualitative text analysis. Behav Res Methods 2019 Aug;51(4):1766-1781. [doi: 10.3758/s13428-019-01202-8] [Medline: 30941697]
14. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a Tool for Health Research: A Systematic Review. Am J Public Health 2017 Jan;107(1):e1-e8. [doi: 10.2105/AJPH.2016.303512] [Medline: 27854532]
15. Scanfeld D, Scanfeld V, Larson EL. Dissemination of health information through social networks: twitter and antibiotics. Am J Infect Control 2010 Apr;38(3):182-188 [FREE Full text] [doi: 10.1016/j.ajic.2009.11.004] [Medline: 20347636]
16. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov 29;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
17. McClellan C, Ali MM, Mutter R, Kroutil L, Landwehr J. Using social media to monitor mental health discussions - evidence from Twitter. J Am Med Inform Assoc 2017 May 01;24(3):496-502 [FREE Full text] [doi: 10.1093/jamia/ocw133] [Medline: 27707822]
18. Tang L, Bie B, Zhi D. Tweeting about measles during stages of an outbreak: A semantic network approach to the framing of an emerging infectious disease. Am J Infect Control 2018 Dec;46(12):1375-1380 [FREE Full text] [doi: 10.1016/j.ajic.2018.05.019] [Medline: 29929837]
19. Doan S, Yang EW, Tilak SS, Li PW, Zisook DS, Torii M. Extracting health-related causality from twitter messages using natural language processing. BMC Med Inform Decis Mak 2019 Apr 04;19(Suppl 3):79 [FREE Full text] [doi: 10.1186/s12911-019-0785-0] [Medline: 30943954]
20. Gohil S, Vuik S, Darzi A. Sentiment Analysis of Health Care Tweets: Review of the Methods Used. JMIR Public Health Surveill 2018 Apr 23;4(2):e43. [doi: 10.2196/publichealth.5789] [Medline: 29685871]
21. Zhang Z, Ahmed W. A comparison of information sharing behaviours across 379 health conditions on Twitter. Int J Public Health 2019 Apr;64(3):431-440 [FREE Full text] [doi: 10.1007/s00038-018-1192-5] [Medline: 30585297]
22. Heaivilin N, Gerbert B, Page JE, Gibbs JL. Public health surveillance of dental pain via Twitter. J Dent Res 2011 Sep;90(9):1047-1051 [FREE Full text] [doi: 10.1177/0022034511415273] [Medline: 21768306]
23. Ahlwardt K, Heaivilin N, Gibbs J, Page J, Gerbert B, Tsoh JY. Tweeting about pain: comparing self-reported toothache experiences with those of backaches, earaches and headaches. J Am Dent Assoc 2014 Jul;145(7):737-743 [FREE Full text] [doi: 10.14219/jada.2014.30] [Medline: 24982280]
24. Chan A, Antoun JS, Morgaine KC, Farella M. Accounts of bullying on Twitter in relation to dentofacial features and orthodontic treatment. J Oral Rehabil 2017 Apr;44(4):244-250. [doi: 10.1111/joor.12487] [Medline: 28128466]
25. Hanna K, Sambrook P, Armfield JM, Brennan DS. Exploring and modelling impacts of third molar experience on quality of life: a real-time qualitative study using Twitter. Int Dent J 2017 Oct;67(5):272-280. [doi: 10.1111/idj.12298] [Medline: 28338226]
26. Emmott R, Barber SK, Thompson W. Antibiotics and toothache: a social media review. Int J Pharm Pract 2021 May 25;29(3):210-217. [doi: 10.1093/ijpp/riaa024] [Medline: 33880539]
27. Oliveira LM, Zanatta FB. Self-reported dental treatment needs during the COVID-19 outbreak in Brazil: an infodemiological study. Braz Oral Res 2020 Sep 04;34:e114 [FREE Full text] [doi: 10.1590/1807-3107bor-2020.vol34.0114] [Medline: 32901729]

28. Tao Z, Chu G, McGrath C, Hua F, Leung YY, Yang W, et al. Nature and Diffusion of COVID-19-related Oral Health Information on Chinese Social Media: Analysis of Tweets on Weibo. J Med Internet Res 2020 Jun 15;22(6):e19981 [FREE Full text] [doi: 10.2196/19981] [Medline: 32501808]
29. Creswell JW, Clark VLP. Designing and Conducting Mixed Methods Research. Thousand Oaks, California, USA: Sage; 2007.
30. Creswell JW. A concise introduction to mixed methods research. Thousand Oaks, California, USA: Sage; 2014:36-37.
31. snscrape: A social networking service scraper in Python. GitHub. 2020. URL: https://github.com/JustAnotherArchivist/snscrape [accessed 2022-06-06]
32. Yousefinaghani S, Dara R, Mubareka S, Papadopoulos A, Sharif S. An analysis of COVID-19 vaccine sentiments and opinions on Twitter. Int J Infect Dis 2021 Jul;108:256-262 [FREE Full text] [doi: 10.1016/j.ijid.2021.05.059] [Medline: 34052407]
33. Mohamed Ridhwan K, Hargreaves CA. Leveraging Twitter data to understand public sentiment for the COVID‐19 outbreak in Singapore. International Journal of Information Management Data Insights 2021 Nov;1(2):100021 [FREE Full text] [doi: 10.1016/j.jjimei.2021.100021]
34. U.S. and World Population Clock. United States Census Bureau. URL: https://www.census.gov/popclock/ [accessed 2022-01-02]
35. Cohen RA. QuickStats: Percentage* of Adults Who Did Not Get Needed Dental Care Because of Cost in the Past 12 Months, by Age Group and Sex - National Health Interview Survey, United States, 2019. MMWR Morb Mortal Wkly Rep 2021 Jun 25;70(25):935 [FREE Full text] [doi: 10.15585/mmwr.mm7025a5] [Medline: 34166335]
36. Schoen C, Osborn R, Squires D, Doty MM. Access, affordability, and insurance complexity are often worse in the United States compared to ten other countries. Health Aff (Millwood) 2013 Dec;32(12):2205-2215. [doi: 10.1377/hlthaff.2013.0879] [Medline: 24226092]
37. Dieleman JL, Cao J, Chapin A, Chen C, Li Z, Liu A, et al. US Health Care Spending by Payer and Health Condition, 1996-2016. JAMA 2020 Mar 03;323(9):863-884 [FREE Full text] [doi: 10.1001/jama.2020.0734] [Medline: 32125402]
38. Cerullo M. As medical costs soar, more Americans turn to crowdfunding. CBS News. 2020 Feb 21. URL: https://www.cbsnews.com/news/health-care-costs-crowdfunding-medical-bills [accessed 2022-01-02]
39. McClanahan C. People are raising $650 million on GoFundMe each year to attack rising healthcare costs. Forbes. 2018 Aug 13. URL: https://www.forbes.com/sites/carolynmcclanahan/2018/08/13/using-gofundme-to-attack-health-care-costs/?sh=658330282859 [accessed 2021-12-15]
40. Heller N. The hidden cost of GoFundMe health care. New Yorker. 2019 Jun 24. URL: https://www.newyorker.com/magazine/2019/07/01/the-perverse-logic-of-gofundme-health-care [accessed 2021-12-15]
41. Wooley S. American health care tragedies are taking over crowdfunding. Bloomberg. 2017 Jun 12. URL: https://www.bloomberg.com/news/articles/2017-06-12/america-s-health-care-crisis-is-a-gold-mine-for-crowdfunding#xj4y7vzkg [accessed 2021-12-15]
42. Snyder J. Crowdfunding FOR MEDICAL CARE: Ethical Issues in an Emerging Health Care Funding Practice. Hastings Cent Rep 2016 Nov 22;46(6):36-42. [doi: 10.1002/hast.645] [Medline: 27875643]
43. Kranz AM, Chen A, Gahlon G, Stein BD. 2020 trends in dental office visits during the COVID-19 pandemic. J Am Dent Assoc 2021 Jul;152(7):535-541.e1 [FREE Full text] [doi: 10.1016/j.adaj.2021.02.016] [Medline: 34023093]
44. H.R.502 - Medicare Dental Benefit Act of 2021. Congress.gov. URL: https://www.congress.gov/bill/117th-congress/house-bill/502?s=1&r=6#:~:text=Introduced%20in%20House%20(01%2F28%2F2021)&text= [accessed 2021-12-15]
45. Freed M, Ochieng N, Sroczynski N, Damico A, Amin K. Medicare and Dental Coverage: A Closer Look. Kaiser Fam Found. 2021 Jul 28. URL: https://www.kff.org/medicare/issue-brief/medicare-and-dental-coverage-a-closer-look/ [accessed 2022-06-06]
46. Sanger KM. Five Decades Later, Medicare Might Cover Dental Care. New York Times. 2021 Aug 29. URL: https://www.nytimes.com/2021/08/29/upshot/medicare-dental-care.html [accessed 2022-06-06]
47. Social Media Fact Sheet. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/?menuItem=b14b718d-7ab6-46f4-b447-0abd510f4180 [accessed 2022-01-11]
48. Wojcik S, Hughes A. Sizing Up Twitter Users. Pew Research Center. 2019. URL: https://www.pewinternet.org/wp-content/uploads/sites/9/2019/04/twitter_opinions_4_18_final_clean.pdf [accessed 2022-06-06]

Edited by A Mavragani; submitted 18.01.22; peer-reviewed by JA Johnsen, A Zain; comments to author 25.04.22; revised version received 16.05.22; accepted 16.05.22; published 01.07.22.

Please cite as:
Yashpal S, Raghunath A, Gencerliler N, Burns LE
Exploring Public Perceptions of Dental Care Affordability in the United States: Mixed Method Analysis via Twitter
JMIR Form Res 2022;6(7):e36315
URL: https://formative.jmir.org/2022/7/e36315
doi:10.2196/36315
PMID:35658090

Original Paper

Sex Workers' Lived Experiences With COVID-19 on Social Media: Content Analysis of Twitter Posts

Ahmed Al-Rawi[1], PhD; Kiana Zemenchik[1], BA

School of Communication, Simon Fraser University, Burnaby, BC, Canada

Corresponding Author:
Ahmed Al-Rawi, PhD
School of Communication
Simon Fraser University
8888 University Dr.
Burnaby, BC, V5A 1S6
Canada
Phone: 1 778 782 4419
Email: aalrawi@sfu.ca

Abstract

Background: The COVID-19 pandemic has drawn attention to various inequalities in global societies, highlighting discrepancies in terms of safety, accessibility, and overall health. In particular, sex workers are disproportionately at risk due to the nature of their work and the social stigma that comes alongside it.

Objective: This study examines how public social media can be used as a tool of professional and personal expression by sex workers during the COVID-19 pandemic. We aimed to explore an underresearched topic by focusing on sex workers' experiences with the ongoing COVID-19 pandemic on the social media platform Twitter. In particular, we aimed to find the main issues that sex workers discuss on social media in relation to the COVID-19 pandemic.

Methods: A literature review followed by a qualitative analysis of 1458 (re)tweets from 22 sex worker Twitter accounts was used for this study. The tweets were qualitatively coded by theme through the use of intercoder reliability. Empirical, experimental, and observational studies were included in this review to provide context and support for our findings.

Results: In total, 5 major categories were identified as a result of the content analysis used for this study: concerns (n=542, 37.2%), solicitation (n=336, 23.0%), herd mentality (n=231, 15.8%), humor (n=190, 13.0%), and blame (n=146, 10.0%). The concerns category was the most prominent category, which could be due to its multifaceted nature of including individual concerns, health issues, concerns for essential workers and businesses, as well as concerns about inequalities or intersectionality. When using gender as a control factor, the majority of the results were not noteworthy, save for the blame category, in which sexual and gender minorities (SGMs) were more likely to post content.

Conclusions: Though there has been an increase in the literature related to the experiences of sex workers, this paper recommends that future studies could benefit from further examining these 5 major categories through mixed methods research. Examining this phenomenon could recognize the challenges unique to this working community during the COVID-19 pandemic and potentially reduce the widespread stigma associated with sex work in general.

(JMIR Form Res 2022;6(7):e36268) doi:10.2196/36268

KEYWORDS

sex work; social media; COVID-19; pandemic; Twitter; infodemiology; social stigma; sex worker; risk; public health

Introduction

Background

There is some ethnographic research discussing the relationship between the COVID-19 pandemic and female sex workers' experience as a marginalized group in society, but there is a clear lack of literature discussing how public social media can be used as a tool of professional and personal expression amongst sex workers. We focus here on public social media in contrast to closed or restricted social media sites, such as OnlyFans. Though sex workers are a vulnerable group, social media provides people with vocal autonomy and advocacy, while also finding a sense of belonging and community [1]. Thus, social media is an important tool for drawing attention to some of the major concerns and issues experienced by sex

workers during the COVID-19 pandemic. Though there are studies that have focused on the challenges experienced by sex workers more generally, there is a lack of literature examining the ways that sex workers use social media.

This gap in academia reinforces the argument that sex workers are both vulnerable and devalued and fails to acknowledge how social media can act as an important tool for developing an online community, encouraging safety, and developing a dialogue between vulnerable groups, such as sex workers, and the public. It is important to note that most of the literature focuses on female sex workers, but some studies have examined the implications and stigma associated with male sex workers [2-4]. Regardless of the gender of participants, it is undeniable that the stigma associated with sex work reinforces the vulnerability of this group before the pandemic and continues to render them even more vulnerable today.

Despite the ground that existing research has covered, there is still a lack of studies exploring how sex workers use social media to amplify their voices, discuss their personal experiences, and demonstrate support for other vulnerable groups. As a result, this study aims at filling some of these gaps by exploring how sex workers use social media to disseminate information about COVID-19 and their experiences while working during the pandemic.

Literature Review

In general, existing academic studies concerning sex workers' use of social media platforms have not fully explored the implications of social media usage during precarious times, such as during the ongoing COVID-19 pandemic. Although several available studies have examined the implications of sex workers' use of social media [1,2,4-6], there is a limited number of academic studies that have empirically explored and analyzed existing social media posts from sex workers themselves to determine how social media platforms function as important sites for self-identified sex workers in a multitude of ways during the COVID-19 pandemic.

The available studies highlight the importance of ensuring that sex workers have adequate access to health care resources, particularly during the COVID-19 pandemic. Indeed, almost all types of sex work remain highly stigmatized [7], which has led to the creation of various organizations that fight for sex workers' basic rights to ensure that their positions are seen not only as work but also as essential work [1,8]. In fact, some studies and reports concluded with calls of action that specifically identified the discrepancies between work that is, on the one hand, considered essential and, on the other hand, considered sex work, thus drawing attention to the stigma associated with the latter [3,7,9]. Many sex workers are unable to access their health care needs due to government restrictions, long wait times, and limitations on social gatherings [8]. Though the demand for sex work during the pandemic remains strong, sex workers are not granted the same precautions that workers in other industries are, especially with regard to economic and labor support [10], a lack of available and accessible COVID-19 testing [7], and a higher risk for the spread of sexually transmitted diseases [8]. Due to inadequate resources and health care, many sex workers have to rely on their own strategies for their safety and advocacy.

Several published academic studies concerning the implications of COVID-19 and sex workers have been explored worldwide. In Ghana and Kenya, sex work itself is not deemed illegal, though it is not considered essential work [3,11]. The Kenyan and Ghanaian governments have put several regulations in place to prevent the spread of the virus. Though this has reduced the spread of the virus, it has threatened many sex workers' stability and only added to their vulnerability [3,11]. Through various social media channels, sex workers in Kenya are able to share information with one another and discuss safety strategies [3]. The instability of work for sex workers is not limited to the West; thus, it is a reality for sex workers worldwide that can be somewhat mitigated through the use of social media channels. This highlights that although the nature of sex work has depended on being in close physical contact with others, the integration of social media platforms provides sex workers with alternative and safer options for their line of work.

Several studies have highlighted the various purposes social media usage can serve for sex workers, sex work organizations, and other vulnerable groups. One recent study found that social media can function as a tool for advocacy amongst the sex worker community [1]. In the study, the authors highlighted the many benefits of encouraging sex workers to use social media, as well as the potential risks involved with using these platforms [1]. Though activities that are both online and offline have similar characteristics, it is worth noting that not all digital spaces are accessible to all sex workers, even if there has been an influx of sex workers relying on these online spaces [12,13]. In addition, another study found that social media advocacy can increase the efficiency of moving between offline and online spaces, enhance safety, and strengthen communication between sex workers [1]. In particular, the authors found that sex workers who were engaged in dialogue with one another would have many more positive encounters with clientele, as they could exchange information about potential clients and discuss problems concerning specific clientele [1]. In the past, sex workers have been limited to who they encounter in the streets [4], but now, given the ubiquity of social media, sex workers can instantly connect with one another and transmit information in a much more efficient, safe, and effective manner.

In general, the advent of social media has changed how sex workers engage with and meet their clients, promote their services, and explore their agency online. A different study found that online spaces provide extensive opportunities for sex workers to explore and curate their online identities to collaborate with one another, to avoid conflict, and to develop a stronger role of agency [6]. Although social media has proven to have many benefits for sex workers, its use is not without risks. One study found that sex workers struggle with receiving payment, especially since social media can act as a transparent space where content, such as photos and videos, is already visible on one's profile for free [4]. Although social media offers many benefits for sex workers, these sites call for a lot more emotional and mental labor to maintain and manage one's online profile.

One of the challenges not mentioned in these aforementioned studies is that of shadowbanning, which is a phenomenon vital to understanding how sex workers navigate online spaces. Shadowbanning is a process that reduces and limits the visibility of a user's social media platform, even to their followers [14,15]. Shadowbanning has become a lot more prominent for sex workers on social media platforms with an audio-visual focus, such as Instagram, OnlyFans, TikTok, and Facebook. According to 1 sex worker who identifies as a camgirl who has worked in the industry for over 10 years, most social media platforms frequently discriminate against sex workers' persons, work, and jobs, rather than their content [14]. This is believed to be in relation to the US Fight Online Sex Trafficking Act–Stop Enabling Sex Traffickers Act (FOSTA-SESTA) bill, which was enacted to prevent online sex trafficking, though sex work is consensual and legal for many sex workers online. This bill has increased social stigma toward sex work as it does not differentiate between consensual and nonconsensual sex work sectors [14]. However, Twitter has proven to be much less extreme in shadowbanning sex workers, whereas Instagram has made it extremely difficult for users to search for sex workers' accounts through the search function, as well as limiting the visibility of posts and stories of sex workers, even for their followers [15,16]. Thus, the shadowbanning techniques implemented by Twitter appear to be less unrelenting compared to other social media platforms, but the platform is still not completely accepting and inclusive of sex workers.

What is largely missing in these studies is an empirical examination of sex workers' social media posts referencing the pandemic. Therefore, this study attempts to answer the following main research question: *What are the main issues that sex workers discuss on social media in relation to the COVID-19 pandemic?*

Methods

Data Collection

Social media data for this study were collected from December 29, 2019, until July 7, 2021, which is when the study was conducted. The data collection began in late 2019 because this is when coronavirus emerged. Using convenience sampling, 2 researchers examined the Twitter profiles of 22 sex workers, which were identified using several English search terms, such as "sexwork" and "sex workers." The identified users included 15 (68.2%) women; 5 (22.7%) sexual and gender minorities (SGMs), including 1 (20%) individual who is openly transgender; and 2 (9.1%) men. For this study, SGMs should be understood as those who identify as being part of the two-spirit, lesbian, gay, bisexual, transgender, queer or questioning, intersex, asexual (2SLGBTQIA+) community, those who identify as nonbinary, and those who are not cisgender. Most of the sex workers have self-identified their gender on their Twitter profiles. Interestingly, 9 (40.9%) sex workers on our list publicly declared on their Twitter profiles that they are vaccinated.

In this respect, we cannot claim that we managed to map all the available sex workers on Twitter, as this is a prodigious if not an impossible task, but we believe that the number of users was adequate to conduct this study. We used convenience sampling for this study, where users were selected based on their availability; the chosen profiles have identified that they are sex workers in their descriptions, and their profiles are public, in which their posts are available to all users and viewers. We chose to use Twitter for this study due to the higher prominence of shadowbanning on different social media platforms and its higher level of accessibility for sex workers. Though not without controversy, Twitter acts as a more accessible and less problematic platform for sex workers as it is one of the few that does not censor users solely based on their career, and the guidelines are much easier to navigate [15,17].

Data Extraction

To extract the data, we used Twitter application programming interface (API) v2, which allows full data extraction, except for deleted posts or private accounts. In total, 134,025 tweets were downloaded, representing all the available messages posted on the accounts of the 22 sex workers, including what they retweeted. We included the retweets in our analysis because they also represent what the sex workers intended to highlight, such as violations or work restrictions. On average, each sex worker posted 6092 (re)tweets during the study period; the highest number of (re)tweets was 31,802 by 1 (4.5%) sex worker, and the lowest was only 3 (re)tweets. Multimedia Appendix 1 shows that March 14, 2021, had the highest number of (re)tweets.

As the collected data were rather substantial for a manual content analysis, we used a Python script [18] to extract tweets that referenced 33 terms related to the pandemic: "coronavirus," "pandemic," "coronapocalypse," "COVID," "COVID-19," "COVID19," "vaccin*," "vaxxed," "mRNA," "Pfizer," "Moderna," "Johnson & Johnson," ""Johnson&Johnson," "Astra-Zeneca," "Covidiot," "essential worker," "frontline worker," "herd immunity," "virus," "quarantine," "epidemic," "asymptomatic," "personal protective equipment," "PPE," "contact tracing," "lockdown," "new normal," "social distancing," "social distance," and "N95." Another Python script was used to extract the most referenced hashtags in the overall data set. Once again, these search terms do not represent all the available terms that are related to the pandemic, but they cover the main issues concerning COVID-19. The filtered data comprised 1458 (re)tweets that received 11,035 likes, 960 replies, and 10,517,689 retweets. To design the codebook, 2 coders examined a sample data set using emergent coding [19,20] to identify the major categories in the social media posts. In total, 6 main categories were identified as follows:

- Blame: Tweets that blame and criticize the government, rich people, and those in positions of power, such as health authorities, for how the virus is being handled, how the vaccine is being distributed, etc.
- Humor: This is generally used as a coping mechanism or to increase people's ability to relate to what others are feeling. Many people will use the retweet or like function to show that they agree with it or to share these tweets with their followers to make others laugh about some unique circumstances occurring during the pandemic.

- Concerns: Tweets that express health concerns or concerns about their loved ones or people they know, as well as concerns about fake news, media reliability, race, class, and gender—intersectionality—across social groups, including concerns about essential workers, etc.
- Solicitation: Tweets that promote sex workers' availability, vaccine status, times that they will be online or available for meetings, or cities that they will be traveling to. This category also includes when users post photos of themselves.
- Herd mentality: This category refers to the way people will try to create an emotional plea or rational argument to conform to mainstream ideas and policy guidelines; tweets will normally say "retweet if you're still social distancing" or "have you taken the vaccine yet?". This is similar to shaming but is based on being socially desirable by following public health regulations.
- Other: Any other minor issues that are not listed in the previous categories.

To test the codebook, 2 coders examined over 10% (n=150) of the total data set, and in the second attempt after a few rounds of deliberation and discussion, intercoder reliability was measured using Krippendorff $\alpha \geq 0.78$, which was acceptable [21].

Ethical Considerations

As mentioned, sex workers are a highly vulnerable group that is also highly stigmatized. When researching such a population, it should be ensured that the privacy and confidentiality of the participants are protected. Given that the data for this study were collected through users' public profiles on social media, ethics clearance was not needed from our university. That being said, though these users have their profiles set to "public," that does not mean that it is any less incriminating if their identities are revealed [22]. To address this ethical concern, this study did not provide any personal details about any of the participants, where names, locations, and specific details from their profiles were all omitted. The method for this study was broad enough that others using similar search terms in the Twitter search bar may not receive the same results—due to algorithms or other factors. To further protect the identities of the participants, our method included both original tweets and retweeted posts, so it is impossible to identify the sex workers examined in this study as the increased number of posts further protects the participants.

Results

Main Findings

As mentioned before, this study aims to identify the major themes and topics sex workers discuss on the social media platform Twitter. From the 1458 tweets and retweets, 6 main categories emerged: blame, humor, concerns, solicitation, herd mentality, and other. These categories are exhaustive and mutually exclusive, and the coders identified the most dominant underlying message in each tweet, especially due to the brevity of tweets. To ensure the validity of the codebook, 2 coders examined a sample and the intercoder reliability test using Krippendorff $\alpha \geq 0.78$ was acceptable [21].

Blame

Blaming others is a reactionary retort that can address one's frustration. For this study, the blame category is understood as posts that explicitly and critically identify an actor responsible for the cause of concern. These actors are criticized and can include government representatives, political parties, the elite, police departments, and others in positions of authority. Although the rationale behind blaming or criticizing others may vary, it can be agreed that blaming others in times of uncertainty is a common reaction, especially if individuals are experiencing a heightened level of fear and anxiety [23,24]. This category came as the fifth-most frequent one, with only 146 (10%) of the social media posts referencing it (Multimedia Appendix 2). Some examples of the blame category include statements resembling criticism, such as:

> *How ICE helped Spread the Coronavirus...*
>
> *Rich people did not experience the same pandemic as working class people and now they get the vaccine first. It's actually twisted.*

In terms of gender, we found that SGM sex workers had the highest percentage of tweets that blame and criticize other parties (n=62, 11.2%), followed by female sex workers (n=83, 9.3%), denoting the urgent issues these groups discuss. Twitter audiences did not show active engagement with messages that contained blame, for this category came fifth, with 1,790,953 (17%) of the most retweeted posts.

Humor

Humor is a powerful coping mechanism that takes on multiple forms on social media and is especially prominent in times of crisis, such as during a global pandemic. With this in mind, humor is understood as the way that users rely on funny posts in their many forms—including sarcasm, satire, and irony—as a coping mechanism and to produce relatable content. Humor is a vital tool for bringing hope to people while they are experiencing times of crisis; humor is also understood as 1 of the only "available option[s]" for openly criticizing others, releasing one's frustration, and providing a narrative for living through a pandemic [25,26]. This category came in fourth at 13%, with 190 of the social media posts referencing it (Multimedia Appendix 2), and male sex workers (n=4, 28.5%) seem to rely on this category more often than other gender groups (Table 1); however, the data on men are extremely low, so definite conclusions cannot be reached. Some examples of humor in this data set include:

> *Feel bad for people who got fit in quarantine but now have to STAY fit for like 7 more months before anyone sees. Shoulda paced yourselves*
>
> *Whoever smoked mid from a soda can as a teenager is immune to coronavirus*

Interestingly, social media audiences found this category more appealing than other ones for they mostly retweeted posts in the humor category, at 3,205,669 (30.4%) of the most retweeted posts.

Table 1. Major categories (N=6) discussed by sex workers by gender.

Category	Women (N=886), n (%)	Men (N=14), n (%)	SGM[a] (N=558), n (%)
Blame	83 (9.3)	0	63 (11.2)
Humor	120 (13.5)	4 (28.5)	66 (11.8)
Concerns	302 (34.0)	6 (42.8)	234 (41.9)
Solicitation	221 (24.9)	3 (21.4)	112 (20.0)
Herd mentality	155 (17.4)	0	76 (13.6)
Other	5 (0.5)	1 (7.1)	7 (1.2)

[a]SGM: sexual and gender minority.

Concerns

The number of concerns that have emerged since the outbreak of coronavirus have been multifaceted and expansive. Thus, for the purposes of research, concerns are understood as being associated with individuals' worries about their loved ones, health issues, essential workers and businesses, and intersectionality [27]. The latter refers to how the lived experiences of someone's life can impact one's "social determinants of health" due to their identity [28]. This category of concerns therefore also refers to those whose identities render them marginal, such as those who are immunocompromised or racialized. Concerns not only refer to fears associated with contracting or spreading coronavirus but also include concerns about public safety, public health, and those who are especially vulnerable to the virus.

The findings showed that this is the top category in the whole data set, constituting 542 (37.2%) of the posts (Multimedia Appendix 2). It is also the highest category across all gender groups (men: n=6, 42.8%; SGMs: n=234, 41.9%; and women: n=302, 34.0%; see Table 1). One example is:

> *in the 8 days since the first COVID-19 case on the navajo nation, our nation now has 49 cases. that number is growing every day. i'm starting a thread of all the ways you can support the navajo nation during the COVID-19 pandemic.*

Another example states:

> *The Covid-19 pandemic particularly has demonstrated how so many people who already face marginalisation are excluded from vital financial resources - government schemes, access to bank accounts and cards. Denying people access to these denies them access to society as a whole.*

In terms of the most referenced hashtags, we found that sex workers used relevant terms to express different types of legal, safety, and health concerns, such as #covid19 (n=52), which is ranked amongst the top 10 most used hashtags, as well as other ones, such as #sexworkerdemands (n=18); #antitrafficking (n=15); #sextrafficking (n=15); #coronavirus (n=14); #decrimqld (n=12), used in relation to decriminalizing sex work in Queensland, Australia; and #antiprostitution (n=6). Some of the other concerns expressed by sex workers were related to their online safety, for a few of them mentioned the importance of avoiding doxxing, which is the online disclosure of personal information, such as sex workers' phone numbers and addresses. For example, 1 (4.5%) sex worker retweeted:

> *Harm Reduction: Doxxing Prevention Tips for Sex Workers and Protesters is now live!*

In terms of the audience's reaction, this category came second in relation to the retweeted posts, comprising 2,787,188 (26.5%) of the retweeted posts.

Solicitation

The COVID-19 pandemic has presented sex workers with many challenges, which include the way that they promote and solicit their work. Although there were some sex workers who had already moved to online platforms, the accessibility of Twitter provides sex workers with an online platform where they can solicit online clients by promoting their services or performing their gender in a specific way [6,29]. This category of promotion and solicitation, therefore, includes content that promotes the availability of sex workers' services, such as photographs of the worker in question, information about their vaccine status, cities that they will be traveling to, and specific dates and times that they will be available for meetings. This is an expansive and comprehensive category that not only includes sex workers promoting themselves but could also include the promotion of other organizations that need financial or social support that aligns with sex workers' individual values. This is the second-most frequent category, with 336 (23%) posts (Multimedia Appendix 2), and female sex workers posted more tweets around this issue (n=221, 24.9%; see Table 1). Some examples of this include:

Finally, this category of posts received the least amount of attention from Twitter audiences, with 32,047 (0.3%) retweeted posts.

Herd Mentality

As mentioned before, herd mentality refers to how the actions and behaviors of a group can influence other individuals. During the COVID-19 pandemic, herd mentality has often been associated with promoting messages that align with government recommendations and public health authorities' guidelines to reduce the number of cases of COVID-19 among the population. Those on social media will often promote government restrictions, information about how to social distance at protests, or promote getting vaccinated. Although there is no central

authority or specified leader to herd mentality, this kind of behavior and its associated actions are transmitted to local networks with the help of others who explicitly promote this way of thinking [30,31]. In this case, the local actions can refer to social media posts, such as in the form of (re)tweets, to promote these messages. This category came third, with 231 (15.8%) posts, as the most important issue discussed by sex workers (Multimedia Appendix 2), and female sex workers are ahead of other groups in referencing it (n=154, 17.4%), followed by SGMs (n=76, 13.6%; see Table 1. Some examples of herd mentality include:

In terms of audience engagement with these types of tweets, this category only came third, with 2,653,060 (25.2%) of the retweeted posts.

Finally, the other category is not discussed here, because it deals with only minor issues that are mostly personal or irrelevant to this study, comprising 13 (0.8%) of the retweeted posts.

Discussion

Principal Findings

As described in this paper, most tweets in our sample align with the issue of concerns. The ongoing COVID-19 pandemic has drawn significant attention to the inequitable structures in society that disproportionately impact vulnerable groups. As a result, the pandemic has influenced how sex workers advocate for themselves on social media platforms, as they are more likely to express concerns about how COVID-19 has impacted their own work and other communities worldwide. Although the concerns have been localized to certain national contexts, such as the United States and Australia, the nature of this online community through the digitization of advocacy has highlighted how COVID-19 is a global phenomenon, thus causing sex workers to advocate for themselves and their own work alongside other disparate groups worldwide. For example, we found that the hashtag #blacklivesmatter was mentioned 50 times by sex workers, together with #BLM (n=31) and #blacklivesmatteraustralia (n=19), to express solidarity with the movement. Though this could be due to the large wave of support for the Black Lives Matter movement following the murder of George Floyd, this highlights that sex workers, as a marginalized community, will not only use social media as a site to discuss their own difficulties but also draw attention to how other social groups have been impacted [28]. The same point applies to intersectional issues and politics.

Given that concerns are closely associated with protecting vulnerable populations, we found that herd mentality was the third-most frequent category. The implicit part of addressing concerns about the vulnerability of certain groups is ensuring that one is taking the necessary steps to protect themselves and others from contracting and spreading the virus. By following the official COVID-19 guidelines, sex workers examined in this study are highlighting the importance of following restrictions to keep others safe and, by showing that they are taking part in this behavior, others should too. Since this category was the third-most frequent one, it is closely associated with the most recurrent category of concerns because, to protect those who are more vulnerable, everyone needs to follow the official health guidelines.

Solicitation was the second-most frequent category. Given that COVID-19 is highly contagious and sex work is a profession that often requires close physical contact between parties, sex workers are at a much higher risk for contracting COVID-19. The codebook sample provided insight into how sex workers were struggling to find work during the pandemic, especially since sex work is not decriminalized or legalized in some countries [11]. Thus, the high frequency of tweets that were coded as solicitation demonstrates how sex workers are using social media platforms to promote their services and have an income to rely on.

The category of humor was not only used by sex workers during the COVID-19 pandemic but also relied upon by many individuals as a coping mechanism [25,26]. Through the use of humor, the severity of the pandemic can be momentarily forgotten as users share relatable information in order to make light of the situation. Though humor was the fourth-most frequent category, its presence shows that the sex workers from this sample engaged with these posts during times of crisis, though there was a much stronger presence of other serious themes.

Finally, the blame category was the least frequent, where those in positions of power are criticized or identified as the cause of concern. During a global health crisis, it is only natural that vulnerable populations will point out further issues of inequity with regard to the specific challenges faced by their community (as is the case with the concerns category) or through the identification of a single actor or representative engaging in activities that are a cause for concern. By identifying different actors that are believed to be responsible for social inequalities or, at the very least, perpetrating them, sex workers are drawing specific attention to the root causes of inequity and how it is manifested in the society during the COVID-19 pandemic.

Limitations

One of the limitations of this study is that we only used posts and retweets that were in the English language. It is likely that there are many other sex workers' accounts on Twitter that are not written in English, but they were not included in this study. This was not the only factor that limited our scope: While conducting our search, we used a limited number of search terms and only used Twitter, while some sex workers might prefer to use other online platforms, such as OnlyFans and Instagram. That being said, Twitter was the most reliable platform for this study because other social media platforms are a lot stricter with shadowbanning content, thereby making it more difficult to find and identify sex workers on those platforms.

Though the authors believed that they reached saturation with 22 sex workers' accounts for this study, it is likely not representative of the sex worker community as a whole. There are many unique challenges that sex workers in different countries experience, which may not be summed up in only 280 characters. Furthermore, there are sex workers who may not

feel comfortable disclosing their profession on their Twitter account for personal or safety reasons, just as there are likely sex workers who do not use Twitter at all.

Conclusion

Although the existing literature vaguely highlights how sex workers use public social media as tools of advocacy, education, and community support in this community, this study used a selected sample of tweets to examine the specific issues that arise from the content of tweets on sex workers' public accounts. The 6 main issues of blame, humor, concerns, solicitation, herd mentality, and other discerned from this sample highlight how Twitter functions as an uncensored space for sex workers to openly discuss their lived experiences while working in an essential, yet high-risk, career throughout the pandemic. These issues, some of which are pressing ones, do not only highlight sex workers' individual concerns but demonstrate that sex workers are advocating for and supporting other vulnerable communities through retweets and likes and posting their own content about other essential workers and those with marginal identities.

Future research on the topic of sex workers' online advocacy could benefit from examining these issues in detail through mixed methods research, such as through online surveys, interviews, or focus groups that explicitly seek answers to questions about sex workers' online advocacy strategies. Examining this phenomenon is not only significant for recognizing the challenges unique to this community during the COVID-19 pandemic but can also highlight how online spaces function as alternative locations of inclusion without the normally ubiquitous stigma attached to sex work. Indeed, social media can be effectively used as a tool for advocacy, education, and community support.

Data Availability

The data set and the coded posts are available with the researchers and can be shared with the reviewers upon request. However, we cannot post them online in a public data repository in order to protect the identity of sex workers.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Frequency of tweets posted by 22 sex workers.
[PNG File , 153 KB - formative_v6i7e36268_app1.png]

Multimedia Appendix 2
Percentages and frequencies of tweets by category.
[PNG File , 78 KB - formative_v6i7e36268_app2.png]

References

1. Duke EE, Sitter KC, Boggan N. Sex work and social media: online advocacy strategies. Cult Pedagog Inq 2018 Oct 01;10(1):49-61. [doi: 10.18733/cpi29372]
2. Argento E, Taylor M, Jollimore J, Taylor C, Jennex J, Krusi A, et al. The Loss of Boystown and transition to online sex work: strategies and barriers to increase safety among men sex workers and clients of men. Am J Mens Health 2018 Nov 28;12(6):1994-2005 [FREE Full text] [doi: 10.1177/1557988316655785] [Medline: 27352925]
3. Macharia P, Moore S, Mathenge J, Ndunda E, Lazarus L, McKinnon LR, et al. Sexual health among Kenyan male sex workers in a time of COVID-19. Health Educ J 2020 Aug 13;80(1):119-127. [doi: 10.1177/0017896920950774]
4. Ryan P. #Follow: exploring the role of social media in the online construction of male sex worker lives in Dublin, Ireland. Gender Place Cult 2016 Oct 31;23(12):1713-1724. [doi: 10.1080/0966369x.2016.1249350]
5. Bernier T, Shah A, Ross LE, Logie CH, Seto E. The use of information and communication technologies by sex workers to manage occupational health and safety: scoping review. J Med Internet Res 2021 Jun 24;23(6):e26085 [FREE Full text] [doi: 10.2196/26085] [Medline: 34185001]
6. Chib A, Nguyen H, Lin D. Provocation as agentic practice: gender performativity in online strategies of transgender sex workers. J Comput Mediat Commun 2021;26(2):55-71. [doi: 10.1093/jcmc/zmaa017]
7. Singer R, Crooks N, Johnson AK, Lutnick A, Matthews A. COVID-19 prevention and protecting sex workers: a call to action. Arch Sex Behav 2020 Nov 14;49(8):2739-2741 [FREE Full text] [doi: 10.1007/s10508-020-01849-x] [Medline: 33057832]
8. Gichuna S, Hassan R, Sanders T, Campbell R, Mutonyi M, Mwangi P. Access to healthcare in a time of COVID-19: sex workers in crisis in Nairobi, Kenya. Glob Public Health 2020 Oct 20;15(10):1430-1442. [doi: 10.1080/17441692.2020.1810298] [Medline: 32816628]
9. Milan S, Treré E. The rise of the data poor: the COVID-19 pandemic seen from the margins. Soc Media Soc 2020 Jul 11;6(3):2056305120948233 [FREE Full text] [doi: 10.1177/2056305120948233] [Medline: 34192035]

10. Lam E. Pandemic sex workers' resilience: COVID-19 crisis met with rapid responses by sex worker communities. Int Soc Work 2020 Oct 27;63(6):777-781. [doi: 10.1177/0020872820962202]
11. Gbagbo FY. Experiences of commercial sex workers during COVID-19 restrictions in selected metropolises in Ghana. Health Care Women Int 2020 Oct 01;41(11-12):1398-1409. [doi: 10.1080/07399332.2020.1822362] [Medline: 33001717]
12. Johansson H, Scaramuzzino G. The logics of digital advocacy: between acts of political influence and presence. New Media Soc 2019 Feb 04;21(7):1528-1545. [doi: 10.1177/1461444818822488]
13. Gil C, Ramaiah M, Mantsios A, Barrington C, Kerrigan D. Best practices and challenges to sex worker community empowerment and mobilisation strategies to promote health and human rights. In: Goldenberg SM, Thomas RM, Forbes A, Baral S, editors. Sex Work, Health, and Human Rights. New York, NY: Springer; 2021:189-206.
14. Tierney A. Sex Workers Say They're Being Pushed off Social Media Platforms. URL: https://www.vice.com/en/article/3kjawb/sex-workers-say-theyre-being-pushed-off-social-media-platforms [accessed 2022-04-25]
15. Are C. How Instagram's algorithm is censoring women and vulnerable users but helping online abusers. Fem Media Stud 2020 Jun 25;20(5):741-744. [doi: 10.1080/14680777.2020.1783805]
16. Blunt D, Wolf A, Coombes E, Mullin S. Posting into the Void: Studying the Impact of Shadowbanning on Sex Workers and Activists. URL: https://hackinghustling.org/posting-into-the-void-content-moderation/ [accessed 2022-04-23]
17. Cotter K. Playing the visibility game: how digital influencers and algorithms negotiate influence on Instagram. New Media Soc 2018 Dec 14;21(4):895-913. [doi: 10.1177/1461444818815684]
18. Al-Rawi A, Shukla V. Bots as active news promoters: a digital analysis of COVID-19 tweets. Information 2020 Sep 27;11(10):461. [doi: 10.3390/info11100461]
19. Wimmer R, Dominick J. Mass Media Research. 10th edition. New York, NY: Cengage Learning; 2013:1133307337.
20. Riffe D, Lacy S, Watson B, Fico F. Analyzing Media Messages: Using Quantitative Content Analysis in Research. London, UK: Routledge; 2019.
21. Krippendorff K. Krippendorff's alpha. In: Inalkind N, editor. Encyclopedia of Research Design. Thousand Oaks, CA: SAGE Publications; 2010.
22. Shaver FM. Sex work research: methodological and ethical challenges. J Interpers Violence 2005 Mar 02;20(3):296-319. [doi: 10.1177/0886260504274340] [Medline: 15684139]
23. Kumar A, Nayar KR. COVID-19: stigma, discrimination, and the blame game. Int J Mental Health 2020 Aug 31;49(4):382-384. [doi: 10.1080/00207411.2020.1809935]
24. Boukala S, Dimitrakopoulou D. Absurdity and the "blame game" within the Schengen area: analyzing Greek (social) media discourses on the refugee crisis. J Immigr Refug Stud 2017 Apr 13;16(1-2):179-197. [doi: 10.1080/15562948.2017.1303560]
25. Ridanpää J. Crisis and humorous stories: laughing at the times of COVID-19. Literacy Geogr 2020;6(2):296-301.
26. Outley C, Bowen S, Pinckney H. Laughing while black: resistance, coping and the use of humor as a pandemic pastime among blacks. Leis Sci 2020 Jun 26;43(1-2):305-314. [doi: 10.1080/01490400.2020.1774449]
27. Schück S, Foulquié P, Mebarki A, Faviez C, Khadhar M, Texier N, et al. Concerns discussed on Chinese and French social media during the COVID-19 lockdown: comparative infodemiology study based on topic modeling. JMIR Form Res 2021 Apr 05;5(4):e23593 [FREE Full text] [doi: 10.2196/23593] [Medline: 33750736]
28. Ryan NE, El Ayadi AM. A call for a gender-responsive, intersectional approach to address COVID-19. Glob Public Health 2020 Sep 07;15(9):1404-1412. [doi: 10.1080/17441692.2020.1791214] [Medline: 32633628]
29. Cunningham S, Sanders T, Scoular J, Campbell R, Pitcher J, Hill K, et al. Behind the screen: commercial sex, digital spaces and working online. Technol Soc 2018 May;53:47-54. [doi: 10.1016/j.techsoc.2017.11.004]
30. Loxton M, Truskett R, Scarf B, Sindone L, Baldry G, Zhao Y. Consumer behaviour during crises: preliminary research on how coronavirus has manifested consumer panic buying, herd mentality, changing discretionary spending and the role of the media in influencing behaviour. J Risk Finan Manag 2020 Jul 30;13(8):166. [doi: 10.3390/jrfm13080166]
31. Keller SN, Honea JC, Ollivant R. How social media comments inform the promotion of mask-wearing and other COVID-19 prevention strategies. Int J Environ Res Public Health 2021 May 25;18(11):5624 [FREE Full text] [doi: 10.3390/ijerph18115624] [Medline: 34070305]

Abbreviations

SGM: sexual and gender minority

Edited by A Mavragani; submitted 07.01.22; peer-reviewed by J Anderson, I Skafle; comments to author 05.04.22; revised version received 28.04.22; accepted 16.06.22; published 14.07.22.

Please cite as:
Al-Rawi A, Zemenchik K
Sex Workers' Lived Experiences With COVID-19 on Social Media: Content Analysis of Twitter Posts
JMIR Form Res 2022;6(7):e36268
URL: https://formative.jmir.org/2022/7/e36268
doi:10.2196/36268
PMID:35767693

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org