JMIR Formative Research

Impact Factor (2023): 2.0
Volume 4 (2020), Issue 12   ISSN 2561-326X   Editor in Chief: Amaryllis Mavragani, PhDc

Contents

Original Papers

Viewpoint

Original Paper

Care Me Too, a Mobile App for Engaging Chinese Immigrant Caregivers in Self-Care: Qualitative Usability Study

Mandong Liu[1,2], MSW; Tongge Jiang[3], MSG; Kexin Yu[1,2], MSW; Shinyi Wu[1,2,4], PhD; Maryalice Jordan-Marsh[1], PhD; Iris Chi[1,2], DSW

[1]Suzanne Dworak-Peck School of Social Work, University of Southern California, Los Angeles, CA, United States

[2]Edward R Roybal Institute on Aging, University of Southern California, Los Angeles, CA, United States

[3]Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, United States

[4]Epstein Department of Industrial and Systems Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, CA, United States

Corresponding Author:
Mandong Liu, MSW
Suzanne Dworak-Peck School of Social Work
University of Southern California
669 W 34th St
Los Angeles, CA, 90089
United States
Phone: 86 176 6259 0259
Email: mandongl@usc.edu

Abstract

Background: Caregiving and self-care are challenging for Chinese immigrants in the United States due to limited accessible support and resources. Few interventions exist to assist Chinese immigrant caregivers in better performing self-care. To address this gap in the literature, our team developed the Care Me Too app to engage Chinese immigrant caregivers in self-care and conducted a user experience test to assess its usability and acceptability.

Objective: This paper aims to report the results of the app's usability and acceptability testing with Chinese immigrant caregivers and to solicit participants' feedback of the app design and functions.

Methods: A total of 22 Mandarin-speaking Chinese caregivers participated in the study, which consisted of 2 parts: the in-lab testing and the 1-week at-home testing. In-depth face-to-face interviews and follow-up phone interviews were used to assess user experience of the app's usability and acceptability and to solicit feedback for app design and functions. Directed content analysis was used to analyze the qualitative data.

Results: Among the 22 participants, the average age was 60.5 (SD 8.1) years, ranging from 46 to 80 years; 17 (77%) participants were women and 14 (64%) had an associate degree or higher. Participants reported uniformly positive ratings of the usability and acceptability of the app and provided detailed suggestions for app improvement. We generated guidelines for mobile health (mHealth) app designs targeting immigrant caregivers, including weighing flexibility versus majority preferences, increasing text sizes, using colors effectively, providing engaging and playful visual designs and functions, simplifying navigation, simplifying the log-in process, improving access to and the content on the help document, designing functions to cater to the population's context, and ensuring offline access.

Conclusions: The main contribution of this study is the improved understanding of Chinese caregivers' user experiences with a language-appropriate mHealth app for a population that lacks accessible caregiving and self-care resources and support. It is recommended that future researchers and app designers consider the proposed guidelines when developing mHealth apps for their population to enhance user experience and harness mHealth's value.

(JMIR Form Res 2020;4(12):e20325) doi:10.2196/20325

KEYWORDS

mHealth; co-design; usability; acceptability; immigrant; caregiver; mobile phone

Introduction

Asian American and Pacific Islander Caregivers' Needs

Caring for others can be rewarding but also straining. The negative impact of caregiving is well documented in the literature, and it contributes to higher levels of stress and depression and lower levels of physical health and self-efficacy compared with noncaregivers [1-3]. Caregiving is especially challenging for Asian American and Pacific Islander populations. Asian American and Pacific Islander populations are the second fastest growing ethnic group among the aging population, with the number of older adults projected to increase by 145% from 2010 to 2030 [4]. However, stereotyped as the "model minority," Asian American and Pacific Islander individuals' caregiving and service needs are often overlooked [5]. According to a meta-analysis study in 2005, Asian American and Pacific Islander caregivers performed a higher number of caregiving tasks, were more depressed, had worse physical health, and used less formal support than White caregivers [6].

Language barriers, employment situations, and cultural beliefs heighten Asian American and Pacific Islander caregivers' challenges. Asian American and Pacific Islander caregivers reported a lack of language-appropriate and culturally sensitive formal services [7-9]. Almost 1 in 5 Asian American and Pacific Islander caregivers preferred non-English to English materials [10]. Asian American and Pacific Islander family caregivers were more likely to work (67%) and work full-time (61%) compared with other ethnic groups [10]. It is common in Asian American and Pacific Islander cultures to not speak up about challenges faced or ask for external caregiving help due to a strong sense of pride in self-managing [8,11].

The Chinese population comprises the largest Asian group in the United States. [12]. Literature shows that Chinese immigrant caregivers face severe caregiving challenges [9,13]. Providing culturally and linguistically appropriate caregiving and self-care support is crucial for enhancing the well-being of Chinese immigrant caregivers, but resources specific to their needs are lacking. In an evidence-mapping study, only 7 training programs were found to be provided to Chinese immigrant caregivers; not all the programs included self-care topics [14].

To address this research gap, our team, consisting of experts from the fields of social work, nursing, and gerontology, developed an in-person Chinese caregiver self-care training program and pilot tested it in Los Angeles county. The program was based on the body-mind-spirit (BMS) model to address a caregiver's holistic well-being [15]. The BMS model finds its roots in the Eastern philosophies of Taoism (eg, yin-yang perspectives) and Buddhism and focuses on the dynamics between individuals and the world and on mind-body and human-nature harmony. It also adopts relevant knowledge from traditional Chinese medicine. The model regards the physical, emotional, and spiritual as indivisible yet distinctly different aspects of the same reality. With these concepts being familiar to Chinese populations, the BMS model is believed to cater to the unique needs of Chinese immigrants, with cultural adaptability and consistency [15]. Researchers have used the BMS model in developing interventions and obtained preliminary evidence that sheds light on the feasibility of using the BMS model to tackle the stress and vulnerability inherent in the process of caring [16].

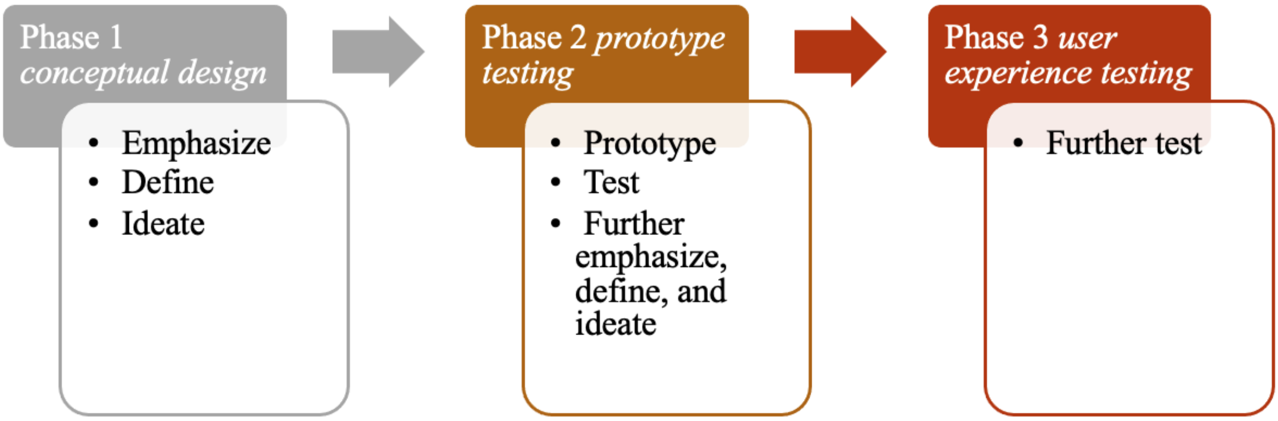
During the pilot testing of the in-person training, the team encountered challenges in recruitment and scheduling because of caregivers' time constraints and transportation barriers, making the training difficult to access for our target audience. This calls for training delivery that allows flexibility in access options to respond to the special needs of Chinese immigrants.

Mobile Health Apps for Caregivers

An increasing number of web-based interventions and mobile health (mHealth) apps have been developed to deliver more accessible resources for older adults and caregivers [17]. According to 2 systematic review studies (one with 7 studies and the other with 14 studies), web-based interventions and mHealth apps were effective in decreasing caregivers' depression and anxiety levels and increasing self-efficacy [18,19]. A study found that 7 out of 10 caregivers were somewhat or very receptive to using smartphone apps to assist them in caregiving [20]. However, in 2017, only 44 out of over 200,000 mHealth apps were designed especially for caregivers, and only 10 included self-care support functions, such as app-based support groups, a stress tracker, and health maintenance advice [17]. To the best of our knowledge, there is no app designed for Chinese immigrant caregivers of limited English proficiency to help them better care for their physical and psychological well-being.

Using a co-design approach in collaboration with local Chinese immigrant caregivers, our team of researchers, consisting of experts from the fields of social work, nursing, gerontology, engineering, and information technology, developed the Care Me Too app in Chinese for promoting caregivers' self-care. The team followed a traditional 5-stage design thinking process in co-designing the app: emphasize (develop understanding of the users' needs), define (frame the issue in a human-centric manner), ideate (develop a breadth of ideas), prototype (produce a sample or prototype), and test (test the components of the overall app design and study the feasibility of the app as a solution for caregivers' self-care) [21,22]. We conducted the design thinking process in 3 phases of participant co-design. In phase 1, the conceptual design phase, we interviewed 7 Chinese immigrant caregivers to emphasize, define, and ideate user needs. The research team then designed a Care Me Too app prototype. In the phase 2, the prototype testing phase, we interviewed the same 7 participants plus 5 additional Chinese immigrant caregivers to review the prototype design and gauge feasibility to help caregiver self-care. The interviews helped the research team further emphasize, define, and ideate user needs to refine the app design. The research team then developed the functional Care Me Too app. In phase 3, the user experience testing phase, the team assessed the usability, acceptability, and at-home use of the Care Me Too app and its specific functions. Figure 1 illustrates these 3 phases. This paper aims to report the results of the app's user experience testing with Chinese immigrant caregivers, in which existing literature is limited, and to solicit participants' feedback for app design and functions.

Figure 1. The 3 phases of participant co-design based on the design thinking process.



When designing the Care Me Too app, the team took into consideration Nielsen's 10 usability heuristics [23] to enhance the interaction design. Those heuristics most relevant to our app included match between the system and the real world (speak the users' language and use words, phrases, and concepts familiar to them), user control and freedom (a clearly marked "emergency exit" to leave the unwanted state or to undo), consistency and standards (use the same words and actions for the same thing to decrease confusion), aesthetic and minimalist design (decrease irrelevant information), and help and documentation (easy-to-search help and documentation should be provided).

Methods

Care Me Too App Description

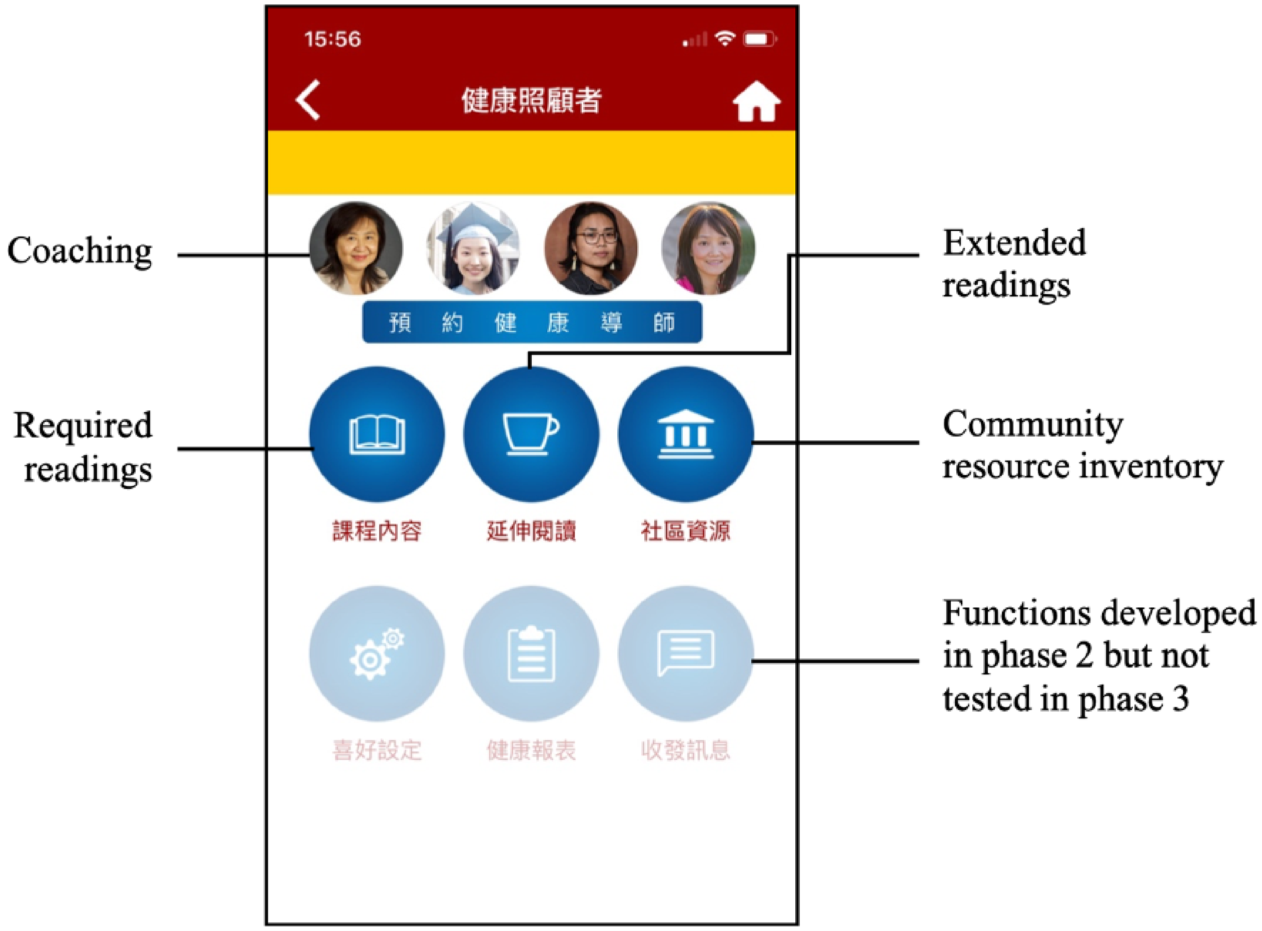
In the current literature, the most widely adopted functions in web-based interventions for caregivers included information and resources, assistance in problem solving, peer psychosocial support, professional psychosocial support, and family communication and care coordination [17,19]. Based on the research team's health promotion and digital expertise, we designed the app with 6 functions: required readings, BMS exercise videos, extended readings, a chat room, coaching, and a community resource inventory. The content of the required readings, which focused on self-care knowledge and skills, was modified after evaluating our BMS in-person caregiver training program and was divided into 4 chapters: (1) the role and stress of caregivers (eg, different caregiving duties, boundary issues, sources of stress, and an overview of using the BMS model to achieve holistic health); (2) diet, exercise, sleep, and medication management (eg, how to achieve a balanced diet, how to correctly store and use medication); (3) problem-solving skills (eg, how to increase self-enhancing thinking, a 7-step guide to problem solving with 2 real-life examples); and (4) how-to information on using formal and informal support resources (eg, what formal and informal resources are, things to pay attention to when immigrants use these resources). Interactive graphs were also used to better illustrate the information. For example, a healthy eating plate was used to visually assist caregivers in achieving a balanced diet [24]. After a user clicks a type of food on the plate (eg, fruits), a textbox pops up and describes its nutrition and recommended serving per day.

The BMS videos were visual demonstrations of exercises aiming to promote exercise and help relieve caregiving stress, such as head and neck self-massage, breathing exercises, and guided imagery. Extended readings addressed caregiving knowledge and skills, such as communication, oral health, legal issues, and more.

Social features have been employed in mHealth apps and have been shown to be effective in helping users gain social support and decrease stress related to their illnesses [25-27]. Therefore, we developed 2 functions, a chat room and coaching, to promote communication among caregivers and between caregivers and health coaches (health care professionals or researchers). To reduce learning burden in the testing stage, instead of building a new in-app chat room, we chose to use WeChat, a commonly used messaging and social media app among Chinese populations. For the coaching function, caregivers could schedule a meeting to consult a coach about health or caregiving-related questions; however, no real coaching took place during the testing.

The community resource inventory contained a variety of local and national resources available to Chinese caregivers and older adults, such as the Chinese American Coalition for Compassionate Care. Each organization was presented with a brief introduction, their website link, and their contact information. Figure 2 shows the app's home screen.

Figure 2. Home screen of the Care Me Too app.



Participants and Recruitment

Caregivers were recruited in the Los Angeles area from January to April 2019. The inclusion criteria were (1) 18 years or older, (2) ability to speak and understand Mandarin Chinese, (3) experience caring for someone aged 65 or older for at least 3 months, (4) experience assisting the care recipient with at least one activity of daily living or a medical task, and (5) possession of an iPhone (the team could provide one if a caregiver did not have an iPhone). A participant could either be a formal caregiver (eg, home care aide) or an informal caregiver (eg, family member, friend).

The team sent out flyers to community-based organizations with Chinese immigrant clients (eg, senior housing, community service centers, adult day health care centers) and asked their staff members to help recruit participants. Contact information for those interested in participating was directed to the project coordinator for eligibility screening. The coordinator then scheduled an interview time with caregivers who were eligible. Potential participants were also encouraged to refer their friends to participate. A total of 22 caregivers were recruited and participated in the testing.

User Experience Testing

The testing consisted of 2 parts: the 1- to 2-hour in-lab testing and the 1-week at-home testing. In the in-lab testing, which was observed by team members, participants could verbalize their feelings and thoughts while using the app prototype. The at-home testing provided participants with an opportunity to fully explore the app prototype and read the health literacy content.

Each in-lab testing was conducted in Mandarin Chinese either in a private room in a university or a private office at a local senior housing site by 2 Chinese-speaking graduate student research assistants, one serving as the instructor and interviewer and the other as the notetaker. There was a total of 5 research assistants involved, all trained by senior researchers in the team on qualitive interviews. They strictly followed the predesigned task lists and interview guides (described below) to ensure fidelity. The instructor explained the project once again before the participant signed the informed consent and reported their sociodemographic information. After that, the instructor helped download the app to the participant’s cell phone and introduced app functions based on a predesigned task list (eg, read a portion of the content, turn pages, view in-text graphics). This 12-item task list was developed by team members with an engineering background based on the app prototype functions. A hard copy help document in Chinese was presented to the participants for their reference. After hearing instructions, the participant went through the same task list on their own to show mastery of each task. During this process, the instructor provided assistance only if needed. For example, if a participant forgot how to go back to the home screen, the instructor reminded the participant of the icon that looked like a house. The participant needed to master all tasks before moving on to the next step. In the end, the interviewer interviewed the participant using a predesigned interview guide (see Multimedia Appendix 1). Specifically, a 10-item app usability scale modified after the System Usability Scale [28] was used to measure usability; open-ended interview questions were also used to elicit participants' opinions on the app prototype design and functions, such as the usefulness and ease of use of each app function, the visual appeal, the advantages and disadvantages of using an app to receive training, and the barriers to and facilitators of using the app. Sample questions were "Do you feel it was easy or difficult to use the following functions? Why?" and "What will help you use this app more conveniently?" Participants also provided detailed feedback on the health literacy reading content and its cultural sensitivity, with results reported in a future paper. The whole process was audio- and videorecorded. The notetaker took notes during the whole process. Upon completion of the in-lab testing, each participant was given a US $50 gift card for a commonly used grocery store among Chinese immigrants.

After the in-lab testing, the participant went home with a different 8-item app use task list (eg, complete one assigned chapter of the required readings, watch exercise videos) to be completed in the next 7 days. After a week, either the instructor or the notetaker who conducted the in-lab testing with the participant called the participant to complete a 30-minute phone interview for further evaluation of the app. Closed- and open-ended questions were used regarding the participant's actual use of the app, ratings of and more detailed suggestions for each app function, and the overall app's helpfulness, usefulness, and visual appeal (see Multimedia Appendix 2). Each participant was given a US $100 grocery store gift card upon completion of the 1-week at-home testing and the phone interview. After finishing all interviews, the research team summarized findings in a table and forwarded the table to participants for confirmation and further feedback, if any.

Analysis

All 22 in-lab testing interviews were transcribed verbatim in Chinese. The team adopted the directed content analysis approach proposed by Hsieh and Shannon [29] for the analysis of the qualitative data. Two investigators (ML and TJ), who were graduate students and had been trained by senior researchers on qualitative analysis, developed the initial coding scheme based on existing app functions and Nielsen's 10 usability heuristics [23]. After that, they began to independently code transcripts using the predetermined codes. As new information emerged from the transcripts, the coders either modified their current coding scheme or created new codes. The 2 coders discussed their coding results and identified disagreements twice a week. Any unresolvable discrepancies were discussed at the weekly meetings with the research team until a final decision was achieved. All 22 transcripts were separately and independently coded by 2 coders, with a κ coefficient of 0.92. NVivo 12.1.0 software (QSR International) was used for analysis. After deciding the final findings and making results tables, ML translated the findings into English; findings tables in English were then reviewed by the whole research team, including the investigator MJM, a native English speaker. Questions regarding English terms and phrases used were resolved in team meetings.

All activities were approved by the university's institutional review board.

Results

Participant Characteristics

The participants' average age was 60.5 (SD 8.1) years, ranging from 46 to 80 years. Of the 22 participants, 17 (77%) were women and 14 (64%) had an associate degree or higher. A total of 16 of the 22 (73%) participants were formal nonfamily caregivers and 6 (27%) were family caregivers. On average, the participants had lived in the United States for 20.4 (SD 10.2) years, ranging from 5 to 40 years. In addition, 14 of the 22 (64%) participants spoke little or no English. All had experiences using an iPhone even though 2 were using an Android phone during the interview. Of the 22 caregivers, 11 (50%) used an iPad or a tablet and 10 (45%) used a computer or a laptop in daily life.

Feedback for Specific App Functions During In-Lab Testing

Participants' positive feedback and suggestions for specific app functions during the in-lab testing are summarized in Table 1 (quotes are available upon request). Participants reported that the current 5 functions could be beneficial to them in gaining knowledge on caregiving, self-care, exercising, enhancing their communication, connection, and obtaining support from other caregivers and the coaches. They provided detailed improvement suggestions for each function, such as making the content audible; being able to highlight, copy, save, and share the content; being reminded to exercise; categorizing coaches based on their sociodemographic background or expertise; and categorizing community resources based on the type of service. They also suggested alternative format preferences. For example, to communicate with coaches, instead of using video chat (as originally considered), some preferred talking on the phone and some preferred texting. Participants hoped to add a certificate function that would allow them to be awarded a certificate whenever they completed a skill training (eg, blood pressure measuring) to show to a potential employer. One participant suggested adding in-app games that care recipients and caregivers could play together to gain reward points. See Table 1 for a detailed list of comments.

Table 1. Feedback for app functions.

Function	Positive feedback	Suggestions
Required and extended readings	1. Participants could learn how to better care for themselves and care recipients. - 1.1. Participants gained a better understanding of holistic health: body, mind, and spirit.	1. Be able to highlight, copy, and save content in the app and share it with others. 2. Be able to print the content from one's phone. 3. Have a place in the app to show all the latest content updates. 4. Use less text. 5. Some participants preferred listening to over viewing the text; some suggested using videos to illustrate the content; some suggested having multiple formats (text, audio, and video) from which participants could choose. - 5.1. If audio is used, it should be in a storytelling style.
BMS[a] videos	1. Short BMS videos would assist caregivers in exercising conveniently. 2. The video content was credible. 3. Watching videos was more fun than reading the text.	1. Encourage caregivers and care recipients to exercise together. 2. Some participants wanted to be reminded to exercise. However, reminders should be fun and entertaining.
Coaching	1. Participants could ask questions and get expert advice. 2. Participants could gain connection and emotional support.	1. Provide a clear function description, such as instructions on when to contact coaches and how much it costs. 2. Add an introduction for each coach. - 2.1. Categorize coaches based on sociodemographic information. 3. Some participants preferred talking on the phone; some preferred texting. If using video chat, participants preferred WeChat over FaceTime. Finally, some mentioned that participants should be allowed to choose to use different platforms (texting, phone call, video chat) for different needs.
Chat room (WeChat as platform)	1. WeChat group could enhance caregivers' communication on caregiving and health-related issues. 2. It was a platform for asking app-related questions and providing feedback to the research team.	1. Provide a clear function description, such as the purpose of having a caregiver WeChat group. 2. Monitor the content of the discussion. 3. Change name tags in the chat group to indicate whether the user is a research team member or caregiver. 4. Some did not want to talk to strangers and thus might not use the chat room function; some would only talk when asking questions; some preferred one-on-one communication to group chat.
Community resource inventory	The function could assist caregivers in seeking help and gaining support.	1. Provide Chinese resources. 2. Categorize resources. 3. Add emergency information and show it on the front page of this function.
Certificate (brought up by participants)	A certificate would be proof of skill training.	1. Divide the training into several sessions. 2. Provide a certificate for each skill that is learned. 3. Set a high bar for obtaining a certificate.
Games and rewards (brought up by participants)	N/A[b]	1. In-app entertaining games could be designed for care recipients and caregivers to play and spend time together. 2. Points could be gained through game playing and used to exchange for rewards.

[a] BMS: body-mind-soul.

[b] N/A: not applicable.

Feedback and Reflection on Overall App Design During In-Lab Testing

Participants' feedback for the overall app design during the in-lab testing is summarized in Table 2 under the categories of visibility, navigation and error prevention, password and privacy, language and terms, and help and documentation (participants' direct quotes are available upon request). Overall, the majority of participants reported that it was easy to navigate the app and use the current help document. Many liked the current color theme (ie, cardinal, gold, and blue). Participants provided detailed feedback to improve the app design, such as enlarging the font size in different ways, using lighter colors, having a clear log-out icon, and using terms and phrases that are less professional. Again, participants expressed alternative design preferences. For example, there were suggestions regarding the format of the help documentation. In-app text, in-app instruction videos, and a PDF file on the computer were mentioned to replace the current hard copy help document. For the password, some preferred to have a password while others stated that a password was not necessary.

Table 2. Feedback for overall app design.

Themes	Design heuristics	Positive feedback	Suggestions
Visibility	1. Aesthetic and minimalist design 2. Match between system and the real world	Some participants liked the current color theme.	1. Font: - 1.1. Use a bigger font size. - 1.2. Have a function to change the font size. - 1.3. Be able to view horizontally - 1.4. Develop an iPad version of the app. - 1.5. Connect the phone to a television. 2. Color: - 2.1. Use lighter colors. - 2.2. Use more green. - 2.3. Use comforting colors. - 2.4. Increase contrast. 3. Add the university's logo. 4. Make the appearance more interesting, such as adding pictures.
Navigation and error prevention	1. Flexibility and efficiency of use 2. Match between system and the real world 3. Error prevention 4. Help users recognize, diagnose, and recover from errors.	1. It was easy to navigate the app. - 1.1. Easy to use the home button, go forward and backward, and scroll up and down 2. Navigation was familiar from other apps.	1. Have a table of contents to go directly to a specific page. 2. Have a clear log-out icon. 3. Use pop-ups for additional information in the readings. 4. Use the scrolling bar instead of a drop-down list to select date and time for the meeting. 5. For the "reselect" icon in answers for case studies: - 5.1. Darken its color. - 5.2. Make it more visible. 6. For clickable text in graphics: - 6.1. Have a clear "close page" icon. - 6.2. Use arrows or other ways to highlight what to click.
Password and privacy	Recognition rather than recall	The password requirement was valued by some participants.	1. Password should be automatically saved. 2. Participants should be able to change their password. 3. Regarding the password format, some liked to use their phone number as the password (current version) and some preferred to use a fingerprint. 4. Some said a password was not needed. - 4.1. A password would be needed if the app is not free.
Language and terms	Match between systems and real world	Some participants said the terms and expressions used were understandable.	1. Pay attention to different expressions in different regions of China. 2. Use language and terms that are less professional.
Help and documentation	Help and documentation	The current help document (hard copy) was easy to use.	1. Some preferred a hard copy; some preferred the in-app text version; some wanted an in-person orientation; some preferred an instruction video; some wanted to have a PDF document on the computer. 2. Some participants wanted a more comprehensive help document, while some said a help document was not needed.

Participants reported that three main advantages of using a mobile app for receiving training were convenience, having thorough educational content in one place, and enhancing communication with others through content sharing. The disadvantage was the lack of personal contact with other participants. Their perceived biggest barrier to using the app was a caregiver's busy work and personal life; other barriers included having no Wi-Fi or cellular data plan and the potential cost of use in the future. Perceived facilitators to app use to be enhanced in future versions included (1) an app introduction that clearly explains the app's purposes and potential benefits to caregivers, (2) enhancing the app's relevance to care recipients, (3) delivering different training content based on caregivers' backgrounds, (4) maintaining a reasonable reading load or letting caregivers learn at their own pace, and (5) providing a sense of accomplishment, which could be fulfilled by the certificate function.

Usability Scale Results

As assessed by the 10 items of the usability scale, participants were generally positive about the app's usability, with the mean score for each item ranging from 3.86 to 4.32 (score range of 1-5, with higher scores being better). The 3 items that scored below 4 were “Most people would learn to use this app very quickly,” “I need a technical person’s support to be able to use the app,” and “There was too much inconsistency in this app” (reversed score).

Participants’ App Use at Home and Follow-up Assessment

The 1-week at-home testing and follow-up assessment were completed by 21 of the 22 participants; 1 participant did not complete the testing due to her busy schedule. On average, participants reported that they used the app 5.6 days in 1 week (range of 1-7, median of 7) and spent 26.8 minutes on it per day (range of 10-90, median of 20). All but 1 participant rated the app “good” or “very good.” A total of 76% (16/21) of participants thought the app’s color was appealing. Almost half (10/21), however, thought the font size was too small. In addition, 90% (19/21) of participants were interested in using this app in the future after it is fully developed, and all participants were willing to recommend the app to other people. See Table 3 for detailed information.

Table 3. App ratings after the 1-week at-home use.

Questions	Values
How many days in the past week did you use the app, mean (SD)	5.6 (2.0)
On average, how many minutes per day did you use the app, mean (SD)	26.8 (19.4)
Overall, how would you rate our app, n (%)	
Very good	11 (52)
Good	9 (43)
So-so	1 (5)
Is the app's font size appropriate, n (%)	
Appropriate	11 (52)
Too small	10 (48)
Is the app's color appropriate, n (%)	
Yes	16 (76)
No	5 (24)
Did you use the app all by yourself, n (%)	
All by myself	20 (95)
Needed a lot of help from others	1 (5)
What do you like the most about this app, n (%)	
Curriculum	14 (67)
Functions	2 (10)
Navigation	1 (5)
Speed	1 (5)
Other (convenience, can communicate issues with others, like all above)	3 (14)
What do you dislike the most about this app, n (%)	
Visual appeal and feeling	6 (29)
Other (prefer to listen; need more updates of the content; would like to have more consultations regarding health, diet, and healthy recipes; exercise videos not easy to find)	6 (29)
Nothing disliked	9 (43)
Have you ever used other similar health apps, such as "Health" on an iPhone, n (%)	
Yes	4 (19)
No	17 (81)
(If answered "yes" to previous question) How is our app compared to other health apps, n (%)	
As good as other apps	2 (50)
Better than other apps	2 (50)
Are you interested in using this app in the future, n (%)	
Interested	19 (91)
Not sure	1 (5)
Not interested (reason: already learned the content)	1 (5)
Will you recommend this app to others in the future (yes), n (%)	21 (100)

Discussion

Principal Findings

In this study, participants reported uniformly positive ratings of usability and acceptability of the Care Me Too app and its functions. Besides the current functions, participants hoped the team could add a certificate function and a game function in future versions to increase the app's usefulness and ability to engage. They reported that convenience was the biggest advantage of learning through an app because the format catered to their busy schedules. This is consistent with the literature, in which time restraints have been well documented for Chinese caregivers [30].

Participants provided a range of ideas and suggestions for designing a self-care app for Chinese immigrant caregivers. We summarized their suggestions into general mHealth app development guidelines to enrich the app design literature and benefit other immigrant caregiver populations.

General: Weigh Flexibility Versus Majority Preferences

Participants had various preferences for the functions and designs of this app. They reported that it would be convenient if we provided different options so that they could choose the most appropriate one depending on their context. For example, one participant stated she would prefer video chatting with the coach if she were not busy; if she were busy, she would text the coach her questions and wait for the reply while continuing with her work. This is consistent with previous research, in which Mortenson and colleagues [26] reported that formal caregivers desired flexibility in the tools within a user self-management app so that the tools could be tailored to users' individual needs. Thus, researchers need to weigh flexibility versus majority preferences; they can either provide as many options as possible or choose the most widely accepted option, whichever is more appropriate after considering the team's resources and technical challenges. In the future, researchers and app developers might begin by inventorying resource constraints and weighting them when choosing functions, add-ons, and the degree of tailoring possible.

Visibility

Increase Text Sizes

In the app, we used 14-point font for titles and 12-point font for text, but some participants reported these font sizes to be small. There could be two explanations. First, our participants were middle-aged and older adults. Only 5 of 22 were 50 years or younger, and several reported declining vision. Morey et al [31] reviewed mHealth apps for older adults and suggested that apps for older adults use at least 20-point font [31]. Second, Chinese characters use strokes, which make characters complex and detailed. Also, unlike in English, where there is space between words, no space exists between Chinese characters in one sentence. Thus, text can be cluttered if the font size is too small [32]. A font size larger than 14 points for caregiver mHealth apps in the future might make the app more accessible, considering that many family caregivers are older care recipients' spouses [10]. Overtly providing functions for choosing the font size or zooming in on the text would be an alternative.

Use Colors Effectively

Participants provided different opinions on alternatives to our current app color theme. Some liked the current colors, as they were simple and plain, while some hoped for lighter colors, indicating a desire to allow customization. Participants reported low contrast in some textboxes, which made text difficult to read. For example, for the healthy eating plate, participants reported a lack of contrast between text and some textboxes' background colors, as well as between the exit icon and the background colors. Morey et al [31] also noted that color contrast for text and background and for icons and background should be high enough that text and icons are easily seen [31]. In short, colors for an app should be selected to enhance user interface design and support usability and engagement.

Provide Engaging and Playful Visual Designs and Functions

Another suggestion from participants was to increase user engagement and the app's playfulness by using more pictures, using kittens and puppies for BMS exercise reminders instead of an alarm, and adding in-app games that caregivers and care recipients could play together to gain points. Adinugroho [33] observed that playfulness was important for improving the "stickiness" of digital products among Asian populations, with icons and emoticons being comical, cute, and cartoony [33]. For game-oriented designs, our finding is consistent with Jessen et al [34], which found that app users welcomed "gameful" designs, such as having one's own avatars, setting goals to accomplish, and earning rewards, such as points and badges [34]. Meanwhile, gamefully designed tools need to fit users and the context in which they will be used [34]. In our case, participants spend much time with care recipients, and thus, in-app games could be designed for mHealth apps to allow collaboration between caregivers and care recipients so that they learn and have fun together.

App Navigation

Simplify Navigation

Participants confirmed that our app had simple navigation design features that could facilitate universal user uptake. The match between the system for a new app and the real world is clearly significant for users to master the use of a new app and reduce cognitive burden. Our participants reported many of our app designs to be consistent with those in apps that they had used previously, which increased the efficiency of learning to use our app. These features included a home page to allow seeing everything on one screen, a clear home button to quickly get back to the home page if a user gets lost, a scroll bar to move up and down, and arrows to go forward and backward between chapters and sections. Participants also preferred to use a scroll bar, as in the iPhone's alarm function, instead of using the current drop-down list to select the date and time for coach meetings. Features for us to improve and for others to consider included adding a clear app log-out icon and ensuring crucial information is seen without the need to scroll down. For example, in the current case study textboxes, the reselect icon was sometimes outside of one screen, and participants did not know to scroll down to see the icon without the team members' assistance. Morey et al [31] also suggested using a back icon instead of a back arrow to decrease confusion, although our participants appeared satisfied with the back arrows in our app.

Simplify Log-in Process

Privacy has been a concern in previous mHealth literature for apps in which participants' information was recorded in the app and shared with health care professionals [35]. However, in this study, privacy did not appear to be an issue, perhaps because our current app does not have a health tracker function and participants did not need to input their personal information (except for their mobile phone number, which was used to log in to the app). Thus, some participants stated there was no need for a log-in password. Some other participants welcomed the use of their phone number as the log-in password. This is similar to smart phone users' habits in China, where many websites and apps also use mobile phone numbers as the primary identifier for logging in and password recovery [32]. Some participants hoped the app would automatically remember the password if one were required. In brief, a simple and clear log-in process would save users time and make the app use more efficient for caregivers if no privacy concerns exist.

Improve Access to and Content of the Help Document

Instead of using a hard copy of the help document, participants proposed that the next version could use a variety of formats, including a hard copy option, in-app text, introduction videos, and a PDF document that they could save on their computer. Some participants also wanted to see a more detailed help document. Considering that future users of our app and other mHealth apps could be novice or inexperienced smartphone users, it is important to provide a step-by-step guide for first-time users. Additional walk-through instructions, such as bubbles for using a new function, could be used, as suggested in previous research [31].

Function

Design Functions to Cater to the Population's Context

It is crucial to consider locomotion (ie, mobile users use their device while on the move) in the app design to enhance user experiences [36,37]. Our caregivers reported a desire to listen to education materials instead of reading the text so that they could learn while doing household chores. Another context dimension to consider is the immediacy, meaning that mobile users expect to have the right app at the right time [36]. This was manifested in our study when participants reported wanting to add emergency information, such as a 911 link, to the front page of the community resource inventory because of health and medical emergency concerns when caring for older adults. Finally, for mHealth apps, it is important to design app functions with a social circle beyond direct users in mind. In our case, caregiver participants wanted to use the app as a tool to meaningfully spend time with care recipients and suggested adding in-app multiplayer games, as discussed above, as well as changing the title of "BMS exercises" to "BMS exercises that caregivers can do with their care recipients." They also wanted to have a function to share the educational content with others, such as family members or caregiver colleagues, even though they were not directly related to their caregiving job. This finding is consistent with previous literature on web-based support for Chinese caregivers of dementia patients, in which participants liked to share information with other people in their social network to enhance mutual support [30].

Ensure Offline Access

One reported barrier to the use of our app was having no data or Wi-Fi at home or in certain locations. Ensuring offline access would entail that information and resources be downloadable once participants logged in to the app. The BMS videos could not be watched without Wi-Fi or data use because YouTube was the platform used. Chat room and coaching functions could not be used without Wi-Fi or data streaming, either. In the future, there may be legal ways to store YouTube videos to increase accessibility across geographic locations.

Strengths and Limitations

This study had multiple strengths. One was using a sample consisting of both family and nonfamily caregivers. Previous Chinese caregiver literature tended to focus on the needs of family caregivers only, but nonfamily caregivers are crucial to providing support for older Chinese immigrants, and they experience indistinct boundaries between themselves and their care recipients, with shifts from an occupational relationship to a personalized connection, such that the caregiving relationship becomes similar to a mother-daughter relationship [38]. This creates similarities between Chinese family and nonfamily caregivers. The second was the research team's expectation of a co-design model, which was shared with participants. The research team had extensive experience in working with older adults and Chinese immigrants.

There were also limitations. The app is currently only available on the iOS platform and in Chinese; due to time restraints of the pilot study, some standard features were not included in the tested version, such as in-app adjustment of the font sizes and a text-to-speech function, which generated much feedback from participants. The assessments were self-reported, and thus, our participants may have reported more positive findings. We did not collect any information on behavior changes activated by app use.

Conclusions

There is a need to design and develop mHealth apps for caregivers of all populations to help them take care of themselves. This will require attention not only to content but to functionality, ease of use, and pleasure in the use of the app. In this study, using both in-person and at-home testing, we elicited Chinese caregiver participants' feedback on the functions and design of the Care Me Too app. Caregivers gave positive ratings to the overall experience of using the app. The researcher–expert designers made adjustments in a formative design model and developed guidelines intended to enhance the design, usage, and usability characteristics of subsequent versions and, potentially, of other mHealth apps for immigrants and caregivers. In the future, we will include ways to measure the impact of using the app on caregivers' health promotion behaviors and changes in their work with care recipients. We will examine ways to incorporate changes proposed by the co-designer immigrant participants to increase engagement, minimize burden, and promote behavior change. Later, we will test the effectiveness of the Care Me Too app in increasing caregivers' caregiving and self-care knowledge, skills, and self-efficacy and on decreasing negative psycho-socio-spiritual outcomes. We recommend that future researchers and app designers consider the proposed guidelines when developing mHealth apps for their populations to enhance user experience and harness mHealth's value.

Acknowledgments

This study was funded by a pilot grant from the Southern California Clinical and Translational Science Institute (12-5176-3304) and a grant from that Department of Adult Mental Health and Wellness, University of Southern California Suzanne Dworak-Peck School of Social Work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
In-lab Testing Interview Guide.
[DOCX File , 18 KB - formative_v4i12e20325_app1.docx]

Multimedia Appendix 2
At-Home Testing Interview Guide.
[DOCX File , 25 KB - formative_v4i12e20325_app2.docx]

References

1. Capistrant BD. Caregiving for Older Adults and the Caregivers' Health: an Epidemiologic Review. Curr Epidemiol Rep 2016 Feb 3;3(1):72-80. [doi: 10.1007/s40471-016-0064-x]
2. Duffy B, Oyebode JR, Allen J. Burnout among care staff for older adults with dementia. Dementia 2009 Dec 04;8(4):515-541. [doi: 10.1177/1471301209350285]
3. Pinquart M, Sörensen S. Differences between caregivers and noncaregivers in psychological health and physical health: a meta-analysis. Psychol Aging 2003 Jun;18(2):250-267. [doi: 10.1037/0882-7974.18.2.250] [Medline: 12825775]
4. The emerging needs of Asian American and Pacific Islander older adults: What we know and what we have learned. National Asian Pacific Center on Aging. 2017. URL: https://www.napca.org/wp-content/uploads/2017/10/NAPCA-The-Emerging-Needs-of-AAPI-Older-Adults_Final-Report_Feb2017.pdf [accessed 2020-11-24]
5. Wong P, Lai CF, Nagasawa R, Lin T. Asian Americans as a Model Minority: Self-Perceptions and Perceptions by other Racial Groups. Sociol Perspect 2016 Aug 02;41(1):95-118. [doi: 10.2307/1389355]
6. Pinquart M, Sörensen S. Ethnic differences in stressors, resources, and psychological outcomes of family caregiving: a meta-analysis. Gerontologist 2005 Feb;45(1):90-106. [doi: 10.1093/geront/45.1.90] [Medline: 15695420]
7. A community of contrasts: Asian Americans 50 and older. AARP, Asian Americans Advancing Justice. 2014. URL: https://www.aarp.org/content/dam/aarp/home-and-family/asian-community/2016/04/2016-aaaj-report-aarp.pdf [accessed 2020-11-24]
8. Li H. Barriers to and Unmet Needs for Supportive Services: Experiences of Asian-American Caregivers. J Cross Cult Gerontol 2004 Sep;19(3):241-260. [doi: 10.1023/b:jccg.0000034221.07028.9a]
9. Miyawaki CE. A Review of Ethnicity, Culture, and Acculturation Among Asian Caregivers of Older Adults (2000-2012). Sage Open 2015 Feb;5(1):e1 [FREE Full text] [doi: 10.1177/2158244014566365] [Medline: 26229736]
10. Caregiving in the U.S. 2015. National Allegiance for Caregiving, AARP Public Policy Institute. 2015. URL: https://www.aarp.org/content/dam/aarp/ppi/2015/caregiving-in-the-united-states-2015-report-revised.pdf [accessed 2020-11-24]
11. Napoles AM, Chadiha L, Eversley R, Moreno-John G. Reviews: developing culturally sensitive dementia caregiver interventions: are we there yet? Am J Alzheimers Dis Other Demen 2010 Aug;25(5):389-406 [FREE Full text] [doi: 10.1177/1533317510370957] [Medline: 20508244]
12. 2017 American Community Survey 1-Year Estimates. US Census Bureau. 2017. URL: https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2017/1-year.html [accessed 2020-11-24]
13. Echeverria-Estrada C, Batalova J. Chinese immigrants in the United States. Migration Policy Institute. 2020 Jan 15. URL: https://www.migrationpolicy.org/article/chinese-immigrants-united-states [accessed 2020-11-24]
14. Chi I, Liu M, Wang S. Developing the Asian Pacific Islander Caregiver Train-the-Trainer Program in the United States. J Ethnic Cult Diversity Soc Work 2018 Oct 31;29(6):490-507. [doi: 10.1080/15313204.2018.1536906]
15. Chan C, Ho PS, Chow E. A body-mind-spirit model in health: an Eastern approach. Soc Work Health Care 2001;34(3-4):261-282. [doi: 10.1300/j010v34n03_02] [Medline: 12243428]
16. Ho RTH, Sing CY, Wong VPY. Addressing holistic health and work empowerment through a body-mind-spirit intervention program among helping professionals in continuous education: A pilot study. Soc Work Health Care 2016;55(10):779-793. [doi: 10.1080/00981389.2016.1231153] [Medline: 27805500]
17. Grossman MR, Zak DK, Zelinski EM. Mobile Apps for Caregivers of Older Adults: Quantitative Content Analysis. JMIR Mhealth Uhealth 2018 Jul 30;6(7):e162 [FREE Full text] [doi: 10.2196/mhealth.9345] [Medline: 30061093]

18. Parra-Vidales E, Soto-Pérez F, Perea-Bartolomé MV, Franco-Martín MA, Muñoz-Sánchez JL. Online interventions for caregivers of people with dementia: a systematic review. Actas Esp Psiquiatr 2017 May;45(3):116-126 [FREE Full text] [Medline: 28594057]
19. 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]
20. eConnected family caregiver: Bringing caregiving into the 21st century. National Alliance for Caregiving, United Healthcare. 2011. URL: https://www.caregiving.org/data/FINAL_eConnected_Family_Caregiver_Study_Jan%202011.pdf [accessed 2020-11-24]
21. Chen E, Leos C, Kowitt SD, Moracco KE. Enhancing Community-Based Participatory Research Through Human-Centered Design Strategies. Health Promot Pract 2020 Jan;21(1):37-48. [doi: 10.1177/1524839919850557] [Medline: 31131633]
22. What is design thinking? The Interaction Design Foundation. URL: https://www.interaction-design.org/literature/topics/design-thinking [accessed 2020-06-17]
23. Nielsen J, Mack RL, editors. Usability Inspection Methods. 1 ed. New York, NY: Wiley; 1994.
24. Healthy eating plate. Harvard T.H. Chan School of Public Health. URL: https://www.hsph.harvard.edu/nutritionsource/healthy-eating-plate/ [accessed 2020-01-05]
25. McColl LD, Rideout PE, Parmar TN, Abba-Aji A. Peer support intervention through mobile application: An integrative literature review and future directions. Can Psychol/Psychologie canadienne 2014 Nov;55(4):250-257. [doi: 10.1037/a0038095]
26. Mortenson WB, Singh G, MacGillivray M, Sadeghi M, Mills P, Adams J, et al. Development of a Self-Management App for People with Spinal Cord Injury. J Med Syst 2019 Apr 22;43(6):145. [doi: 10.1007/s10916-019-1273-x] [Medline: 31011881]
27. Vorderstrasse A, Lewinski A, Melkus GD, Johnson C. Social Support for Diabetes Self-Management via eHealth Interventions. Curr Diab Rep 2016 Dec;16(7):56. [doi: 10.1007/s11892-016-0756-0] [Medline: 27155606]
28. Brooke J. SUS: A quick and dirty usability scale. Usab Eval Indus 1996;189(194):4-7 [FREE Full text] [doi: 10.1201/9781498710411-35]
29. 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]
30. Chiu T, Marziali E, Colantonio A, Carswell A, Gruneir M, Tang M, et al. Internet-based caregiver support for Chinese Canadians taking care of a family member with Alzheimer disease and related dementia. Can J Aging 2009 Dec;28(4):323-336. [doi: 10.1017/S0714980809990158] [Medline: 19925698]
31. Morey SA, Stuck RE, Chong AW, Barg-Walkow LH, Mitzner TL, Rogers WA. Mobile Health Apps: Improving Usability for Older Adult Users. Ergonomics Design 2019 Apr 01;27(4):4-13. [doi: 10.1177/1064804619840731]
32. Berg S. Chinese mobile app design: Tips for westerners. Prototypr. URL: https://blog.prototypr.io/chinese-mobile-app-design-tips-for-westerners-4e3b36e4739f [accessed 2019-12-29]
33. Adinugroho S. Designing digital products for Asian markets. Medium. URL: https://medium.com/@hellosigit/designing-digital-products-for-the-asian-markets-9b6246872a0d [accessed 2019-12-29]
34. 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]
35. Davis B, Nies M, Shenk D. Developing a pilot e-mobile app for dementia caregiver support: Lessons learned. Online J Nurs Informatics 2014 Feb;18(1):e1 [FREE Full text]
36. Gualtieri M. Forrester's mobile app design context: Location, locomotion, immediacy, intimacy, and device. Forrester. URL: https://go.forrester.com/blogs/forresters-mobile-app-design-context-location-locomotion-immediacy-intimacy-and-device/ [accessed 2020-01-05]
37. Palalas A, Wark N, Pawluk P. Design principles for an adult literacy mobile learning solution in a blended learning context. 2017 Presented at: 2nd World Conference on Blended Learning (IABL); Apr 26-28, 2017; Toronto, ON, Canada. [doi: 10.1145/3136907.3136934]
38. Wu J, Liu M, Ouyang Y, Chi I. Beyond Just Giving Care: Exploring the Role of Culture in Chinese American Personal Care Aides' Work. J Cross Cult Gerontol 2020 Sep;35(3):255-272. [doi: 10.1007/s10823-020-09404-w] [Medline: 32583186]

Abbreviations

BMS: body-mind-spirit
mHealth: mobile health

Edited by G Eysenbach; submitted 16.05.20; peer-reviewed by J Lee, M Leung, I Schiering, T Alessa; comments to author 08.06.20; revised version received 05.08.20; accepted 20.11.20; published 02.12.20.

Please cite as:
Liu M, Jiang T, Yu K, Wu S, Jordan-Marsh M, Chi I
Care Me Too, a Mobile App for Engaging Chinese Immigrant Caregivers in Self-Care: Qualitative Usability Study
JMIR Form Res 2020;4(12):e20325
URL: https://formative.jmir.org/2020/12/e20325
doi:10.2196/20325
PMID:33263552

Original Paper

User Perceptions and Experiences of an Interactive Voice Response Mobile Phone Survey Pilot in Uganda: Qualitative Study

Raymond Tweheyo[1,2], MBChB, MPH, PhD; Hannah Selig[3], MPH; Dustin G Gibson[3], MSc, PhD; George William Pariyo[3], MBChB, PhD; Elizeus Rutebemberwa[1], MBChB, MPH, PhD

[1]Department of Health Policy Planning and Management, Makerere University School of Public Health, Kampala, Uganda

[2]Department of Public Health, Lira University, Lira, Uganda

[3]Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

Corresponding Author:
Raymond Tweheyo, MBChB, MPH, PhD
Department of Health Policy Planning and Management
Makerere University School of Public Health
Mulago Hill Road
P O Box, 7072
Kampala, 256
Uganda
Phone: 256 772466695
Email: rtweheyo@musph.ac.ug

Abstract

Background: With the growing burden of noncommunicable diseases in low- and middle- income countries, the World Health Organization recommended a stepwise approach of surveillance for noncommunicable diseases. This is expensive to conduct on a frequent basis and using interactive voice response mobile phone surveys has been put forth as an alternative. However, there is limited evidence on how to design and deliver interactive voice response calls that are robust and acceptable to respondents.

Objective: This study aimed to explore user perceptions and experiences of receiving and responding to an interactive voice response call in Uganda in order to adapt and refine the instrument prior to national deployment.

Methods: A qualitative study design was used and comprised a locally translated audiorecorded interactive voice response survey delivered in 4 languages to 59 purposively selected participants' mobile phones in 5 survey rounds guided by data saturation. The interactive voice response survey had modules on sociodemographic characteristics, physical activity, fruit and vegetable consumption, diabetes, and hypertension. After the interactive voice response survey, study staff called participants back and used a semistructured interview to collect information on the participant's perceptions of interactive voice response call audibility, instruction clarity, interview pace, language courtesy and appropriateness, the validity of questions, and the lottery incentive. Descriptive statistics were used for the interactive voice response survey, while a framework analysis was used to analyze qualitative data.

Results: Key findings that favored interactive voice response survey participation or completion included preference for brief surveys of 10 minutes or shorter, preference for evening calls between 6 PM and 10 PM, preference for courteous language, and favorable perceptions of the lottery-type incentive. While key findings curtailing participation were suspicion about the caller's identity, unclear voice, confusing skip patterns, difficulty with the phone interface such as for selecting inappropriate digits for both ordinary and smartphones, and poor network connectivity for remote and rural participants.

Conclusions: Interactive voice response surveys should be as brief as possible and considerate of local preferences to increase completion rates. Caller credibility needs to be enhanced through either masking the caller or prior community mobilization. There is need to evaluate the preferred timing of interactive voice response calls, as the finding of evening call preference is inconclusive and might be contextual.

(JMIR Form Res 2020;4(12):e21671) doi:10.2196/21671

KEYWORDS

interactive voice response; noncommunicable diseases; qualitative; Uganda

Introduction

Low- and middle-income country populations suffer approximately 75% of all noncommunicable disease deaths annually (approximately 32 million deaths) [1,2]. Moreover, over 15 million of the noncommunicable disease deaths occurring in the low- and middle- income countries are premature (affecting people aged 30 to 69 years), accounting for about 85% of the global premature deaths from noncommunicable diseases [2].

The use of mobile phone surveys to collect data is expected to increase, leveraging the growing ownership of mobile phones in low- and middle- income countries, although the evidence of this utility is still limited [3,4]. Mobile phone surveys could complement existing noncommunicable disease risk factor surveys such as the World Health Organization (WHO) recommended stepwise approach for surveillance of noncommunicable diseases [5].

Interactive voice response (IVR) surveys are one type of mobile phone survey that could be used. IVR surveys use prerecorded audio files that ask participants to use the keypad on their mobile phone to answer questions. IVRs are mainly known for their use in customer service and public health work and have been used in the United States since the 1970s [6] but are also increasingly being used for continuity of patient health care beyond the hospital setting [7-9]. More recently in low- and middle- income countries, international development work has used IVRs, alongside other multimedia such as radio, with the former offering advantages for interactive reach to their audiences to stimulate behavior change [6].

IVR technology has been piloted and used mainly in health care settings in high-income countries dating back to the early 2000s [7,8], but IVR use is still limited in low- and middle- income countries [10,11]. Within high-income countries, IVR use is generally limited to exploring aspects of self-care [12,13], follow-up of patient care [7,14,15], and evaluating patient-provider interactions in clinical settings [9] but is rarely used for research or surveillance purposes [8,16,17]. Since mobile phone surveys are a relatively new methodology, particularly for low- and middle- income countries, evidence from community respondents on their perceptions on mobile phone surveys and possible reasons for taking the survey and nonresponse can contribute to better future mobile phone survey design and programming efforts.

A qualitative study [10] in Ghana that used focus groups to evaluate the experience of caregiver's health care seeking for their sick child, based on receiving health information through an IVR, reported that all the 37 participants were naïve to IVR but held favorable perceptions about its use for symptom screening and providing guidance for care seeking. Negative perceptions included the fear for nonhuman interaction in using the IVR, a lack of familiarity with IVR, and the related cost [9,18]. Small-scale studies [19-22] have reported IVR use in low- and middle- income countries mainly for monitoring medication adherence, such as for tuberculosis and HIV. Within sub-Saharan Africa, sectors other than health, such as agriculture and social development have successfully used IVR for surveillance and community engagement [6,23].

It is unclear why some respondents complete surveillance questions using IVR and why some do not. We sought to explore user perceptions and experiences of receiving and responding to an interactive voice response mobile phone survey for noncommunicable disease risk factors, to inform the design and delivery of future surveys delivered using mobile phones.

Methods

Researcher Reflexivity

RT the first author was the interviewer for all interviews of this study, collected field notes, transcribed, and led the coding and data analysis. RT is a public health physician, a native of the country of this study. About a third of the participants were known to the researcher, while the rest were obtained through his networks.

Study Design

A qualitative study design [24] was used to elicit the experiences of participants who had completed a structured interview on noncommunicable disease risk factors using an IVR survey delivered to respondents who owned or had access to a mobile phone [16]. This entailed call-backs to all the phone numbers of respondents to the initial automated IVR survey, irrespective of their response status. Those who answered and consented to being interviewed through follow-up calls delivered by a human caller were administered an in-depth interview over the phone to explore reasons for the initial response or nonresponse.

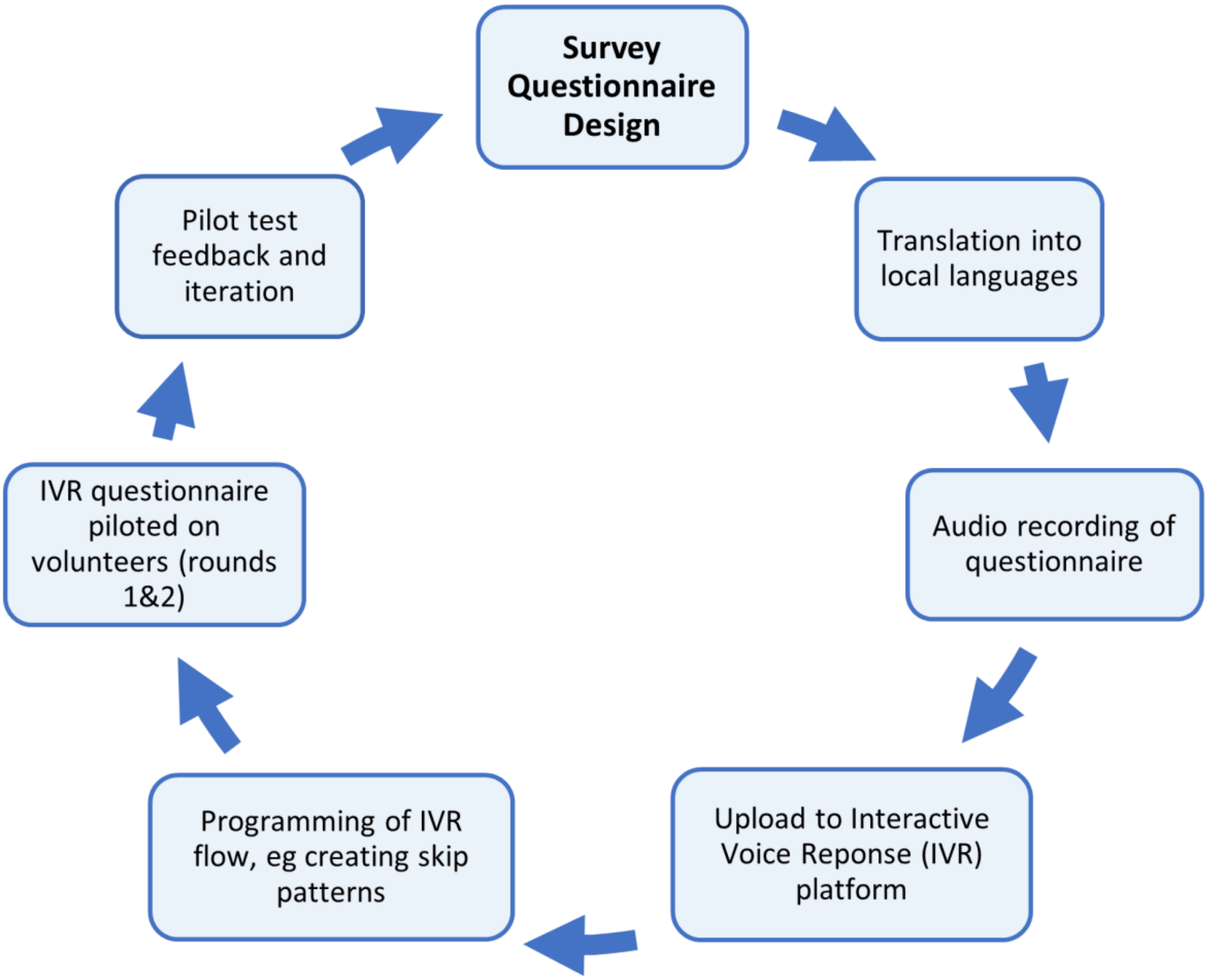
Development and Adaptation of Survey Tools

The study deployed an adapted questionnaire based on an English-language version of questions selected by a joint team from the Johns Hopkins University, WHO, and the United States Centers for Disease Control and Prevention [16]. These questions had been derived from the WHO stepwise approach for surveillance of noncommunicable diseases survey [5], behavioral risk factor surveys [25,26], and the Tobacco Questions for Surveys [27]. The questions were adapted to the local Ugandan context and included local examples of fruits and vegetables, questions on smoking and tobacco use, alcohol consumption, physical activity, and history of checking for high blood pressure or blood glucose level. Although English is one of the official languages in Uganda, a significant portion of the population does not speak English, and in order to increase the reach of the survey, the adapted questionnaire was translated and back-translated into 3 of the 6 other major languages spoken in various regions of the country: Luganda, Runyakitara, and Luo. The 4 language versions of the questionnaire (including English) were digitally audiorecorded and loaded onto an IVR platform. The audiorecorded questionnaire had 69 items and was delivered to all participants via an IVR platform (Viamo).

The IVR platform was a software interface developed by a global social enterprise. The platform used connectivity through the local mobile network operators registered and active in Uganda and pregenerated random-digit dialing codes (such as 077XXXXXXX, or 070XXXXXXX) to dial across different mobile network operators, with a prerecorded voice IVR. The platform used a call-in number to randomly call across the network operators to access survey participants.

This platform delivery mechanism required pretesting to assess feasibility and aspects of acceptability. The survey also informed participants of a chance to win an airtime incentive (in a lottery) after completion of the IVR, as described elsewhere [28], specifically the possibility of winning the equivalent of US $0, $1.35, or $2.70. Figure 1 summarizes the process of conducting the pilot study prior to the main mobile phone survey.

Figure 1. Process of IVR development and testing.



Study Population and Sample Size

The study population comprised adults who could understand or potentially speak any of the 4 languages in which the survey was deployed. Through contacts with communities in and around Kampala, Uganda. A purposive sample of 60 volunteers, was recruited to target 15 participants for each of the 4 languages of the questionnaire. The phone contact information of each of the 60 volunteers was uploaded onto the IVR platform, which then delivered an IVR call (up to 3 call attempts per testing round, if there was no answer). For example, surveys were programmed to call out at 4 PM, then for all unanswered calls, 2 hours and 4 hours later. Any incomplete IVR surveys, following the 3 attempts were not repeated. The IVR call was followed by a human caller to all the 60 volunteering participants, irrespective of their IVR response or completion status. The purpose of the human caller was to explore participants' feedback on their experiences with the IVR encounter and perceptions about the survey. Responses were recorded from 59 participants, representing a response rate of 98.3%.

Data Collection

The IVR survey and interview guide were pilot tested on 3 researchers in English prior to study deployment. Thereafter, as depicted in Figure 2, IVR survey testing and qualitative interviews were conducted iteratively in 5 rounds guided by data saturation. All rounds occurred in April and May 2018. There were a minimum of 2 native speakers taking the survey in each language in each of the first 3 rounds of the survey, designed to validate any differences in opinion for the same reported finding/ query from the survey. However, after 3 rounds of piloting, participants were selected purposively from within the network of the study coordinator and were not necessarily informed that they would receive a survey call, to mimic the real-life context in which prior survey booking may be impractical.

Following survey delivery, a research assistant called each IVR call recipient using the same language in which the IVR had been delivered and asked about their perception on whether or not the IVR was audible, if the subject in the questions was clear, if the pace was right, if the language was polite and courteous, and if the questions were understandable and appeared to be relevant based on the information provided at the beginning of the survey. Respondents were also asked about the difficulties they faced in receiving and navigating the survey, for example, if the instructions for responses such as pressing phone digits were comprehended, if they felt they were in control of the survey, and if they had any other feedback for the survey team. Survey testing rounds 1 to 5 were conducted in all languages, as depicted in Table 1.

Figure 2. Data collection process.

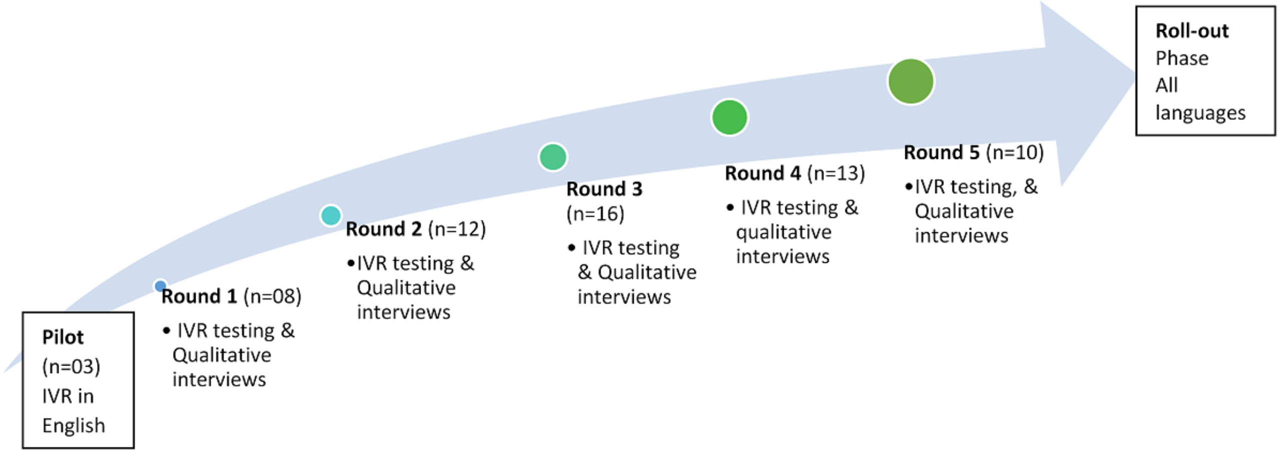


Table 1. Characteristics of IVR survey participants (n=59).

Respondents	Interactive voice response round					Survey status		
	1	2	3	4	5	Completed survey	Failed[a]	Received (incomplete)[b]
n (%)	8 (13.6)	12 (20.3)	16 (27.1)	13 (22.0)	10 (16.9)	18 (30.5)	10 (17.0)	31 (52.5)
Language								
English, n		●●	●●●	●		6	0	0
Luganda, n	●●	●●●	●●●●	●●●●	●●●	3	3	10
Runyakitara, n	●●●	●●●	●●●●	●●●●	●●●●	5	3	10
Luo, n	●●●	●●●●	●●●●●	●●●●	●●●	4	4	11

[a] Cancelled or no answer.

[b] Wrong language selected: 7/31, 22.5% (English 0; Luganda 2; Runyakitara 1; Luo 4).

Data and Theoretical Analysis Approach

A framework analysis, as first described by Ritchie and Spencer [29], was used to explore the themes [30] in the study related to audibility, question clarity, pacing of the study, language courtesy, and validity of questions. Framework analysis is advantageous in that it is purposive in nature (is not bounded to a specific epistemological position) guiding a researcher to identify themes that speak to specific objectives within a study, while exploring experiences within the narratives of participants [29-31]. In essence, both a priori coding from the objectives, and in vivo coding from emergent data are pursued in framework analysis [29,31]. Further to identifying themes within the study, we then explored for variability and the meaning of such divergent views using Janus-face theoretical constructs [32], thereby introducing postante codes to complement some apriori codes.

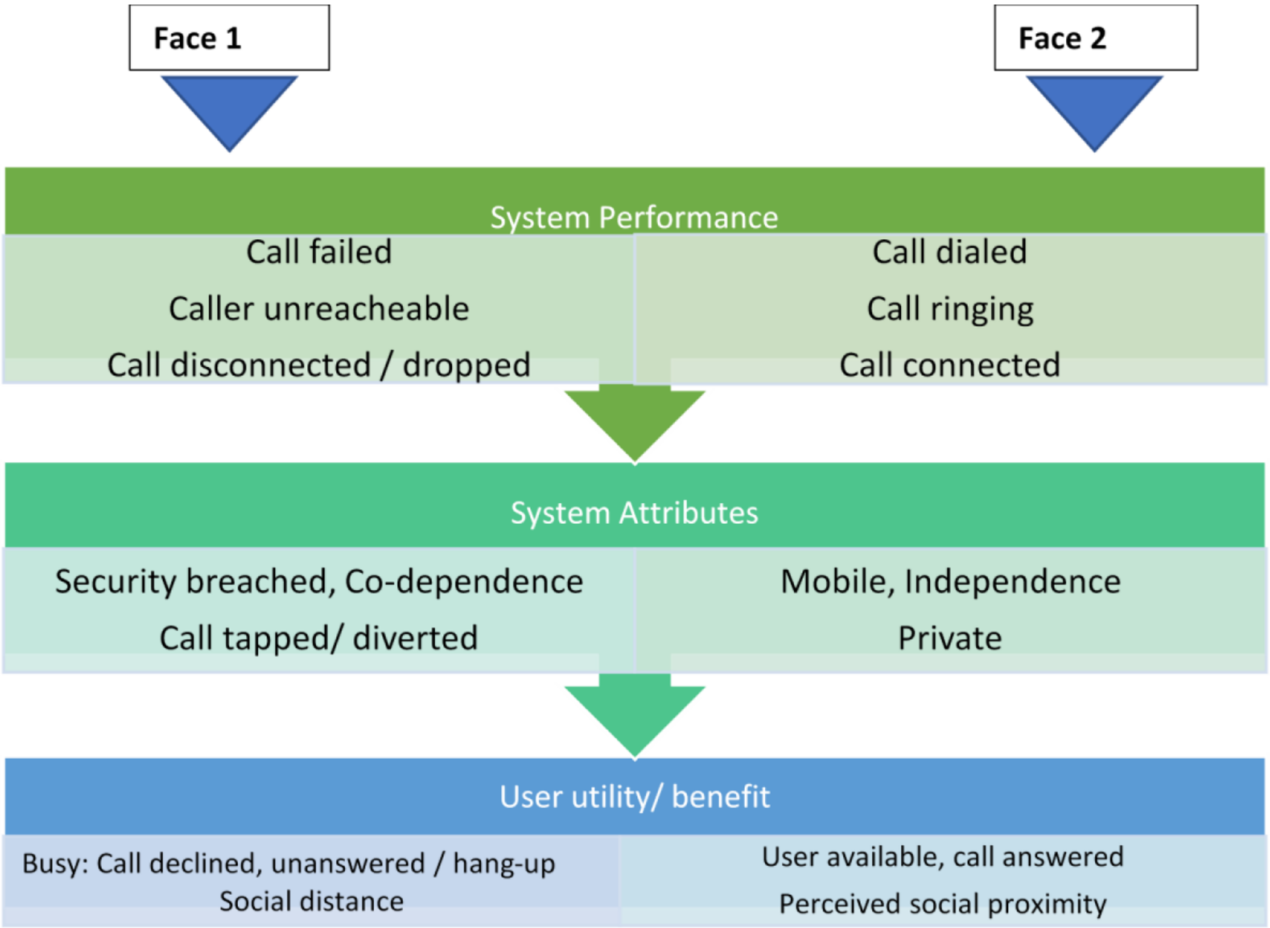
Janus-faced theory [32] (metaphorical perspective) on mobile phones was used to understand the interaction of participants with the mobile phone survey. The theory was chosen based on its simplicity for exploring distinct characteristics along the continuum from high- or low- interest regarding a naïve individual’s behavioral response while engaging with a technology [32]. We conceptualize the encounter of a naïve IVR user as likely to elicit a multiplicity of reactions, which may take the form of either acceptance or rejection of the IVR technology-interface. There were explanatory limitations for the use of potential alternative behavior change theories specifically, the Theory of Reasoned Action, The Theory of Planned Behavior, and the Social Exchange Theory [33-35]. Notably, the trio were limited in their assumption of and individual's prior positive behavioral exposure, thus the choice of the Janus-faced theory [33-35].

The Janus-faces model proposed by Arnold [32] is derived from the metaphor of the Roman deity Janus who was cursed and blessed with 2 faces—each facing a different direction (backward and forward at the same time) [32,36]. The significance is that, while mobile phones and other technologies are designed and built to direct a specific purpose, in reality, a growing evidence base reports sociotechnical system of interaction findings that people's reactions to technology, its use, and adoption can be ironic and paradoxical rather than unified and purposeful [36,37].

An example of the theory's application to mobile phone utility and performance is presented in Figure 3. System performance criteria includes on the one-hand, issues such as call dialed, call ringing (reached), call connected, while on the other hand, it includes things such as—call failed, caller unreachable, call disconnected or dropped, which could be perceived by a user as either advantageous or not.

Ethics approvals were obtained from the Makerere University School of Public Health and the Uganda National Council for Science and Technology, while participant informed consent was embedded within the IVR. The process is published elsewhere [38].

Figure 3. Janus-faced mobile phone utility.



Results

Survey Calls

Table 1 below shows the characteristics of respondents who took the IVR survey (including the language selected) and who later were interviewed qualitatively; 59 of the respondents provided feedback to the IVR and participated in the qualitative interviews. Nearly a half of the participants were female 44% (26/59), and the age ranged between 23 to 47 years (median age was 31 years and 35 years for female and male respondents, respectively).

Qualitative Interview Findings

For the qualitative interviews, the majority of the respondents did not complete their surveys on the first attempt but rather on the second or third call attempt. However, in summary, and as depicted in Table 2, reasons that were stated for survey completion were related to the perceived credibility of the institution providing the survey (Makerere University); the fact that the survey was health-related; the clarity of questions, language, and instructions; being of short duration (10 minutes or less); and the possibility of winning an airtime incentive. On the other hand, the reasons reported for noncompletion were related to being busy, poor network connectivity, and suspicion because of the unknown identity of the caller.

The summary findings in Table 2 are synthesized according to 6 emergent themes related to the process of IVR survey delivery: timing of the survey, call quality, language-related issues, phone type used, survey duration, network connectivity, and perceptions on the incentives. Within each of these themes, we explore the concepts of overall experience with the survey, audibility, question and language clarity, courtesy and question validity. Furthermore, we contrast successful and unsuccessful encounters, from the provider standpoint of intended survey delivery.

Table 2. Summary of themes and issues related to survey completion or noncompletion.

Theme/issue	Observation/comment	Possible impact or implication
Timing of survey	Evening times between 6 PM and 10 PM were preferred. Wrong timing (day times), had lower completion rates.	Timing of IVR[a] surveys should be evaluated for increasing reach and completion rate.
Call quality and credibility of the caller	Inaudibility and challenges with skip patterns compromised receiving calls. The credibility of the caller was paramount for motivating participation.	Platform programming should be tested and piloted for robustness prior to rolling out IVR surveys. Caller credibility should be ascertained in the IVR introduction.
Language-related issues	Language courtesy provided favorable IVR experiences, while wrong language selection affected validity of responses.	A double prompt for the selection of appropriate language is essential. Piloting should ensure courteous translations.
Phone type used	Both ordinary and smart phone users encountered similar challenges with IVR instructions, such as pressing the wrong digits.	User-technology interface is a barrier to the validity of IVR surveys, whose impact requires continuous evaluation.
Survey duration	Shorter realistic survey duration such as 10 minutes is preferred.	IVR surveys need not last longer than 10 minutes.
Mobile network connectivity	Rural respondents, and those located geographically distant from the Capital City had mobile connectivity challenges.	IVR surveys require additional strategies for reaching rural and remote populations, such oversampling.
Perceptions about the incentive	The lottery-type incentive for airtime was perceived favorably by participants	There is need to evaluate various IVR incentive thresholds to find which one increases survey completion.

[a] IVR: interactive voice response.

Timing of the Survey

Several volunteer participants preferred their surveys in the evening (between 6 PM and 10 PM local time), because they were less busy then, off their day’s commitments.

> *It is very difficult for me to receive any call for a survey during the day, because I am not in charge of my schedule at work* [P4 English]
>
> *I saw the call, it actually came in twice, but I was in the field, and I merely ignored it. Maybe next time, if you call me in the evening, I may pick the call* [P13 Runyakitara]

In general, calls that were sent in the evening within the first 2 rounds of delivery had higher response rates than those sent out during the day (9/15 compared to 4/15 of the total 20 participants in the first 2 rounds including failed calls).

Once the call got received, some encounters (participant-phone interaction) were either successful—resulting in a completed call or were unsuccessful—resulting in early termination or nonacceptance. Details of call receipt are described in the subsequent phrases.

Call Quality and Credibility of Caller

Successful Encounters

Successful call recipients were mostly those who received their calls in the evenings between 6 PM and 10 PM local time. Also, repeat calls had higher chances of acceptance compared to initial calls. Informing participants to expect the calls beforehand—although requested by some—did not seem to increase successful encounters. Also, from the final 2 rounds of the pilot, the majority of the recipients reported that a health-related survey from the Ministry of Health and Makerere University School of Public Health interested them and motivated their participation.

However, 2 participants who experienced an early termination of the survey, for unknown reasons, voiced dissatisfaction with the follow-up call because it commenced the interview afresh.

> *My survey stopped abruptly, but when I received the next call introducing the same survey, it just begun afresh, and this was really disappointing. Hmm, because it meant that I had to spend more time on your survey* [P15 Luganda]

Unsuccessful Encounters

Within the first round of pilot interviews, there were 4 main reported reasons for call failure, including missing calls due to wrong timing, inaudibility, challenges with the skip pattern, and also suspicion and skepticism as to the identity of the caller for some.

Several of the volunteer participants who missed their calls, requested to receive call backs later to take the survey. At least 17/39 participants made this request. A majority of the participants took their call during the second call attempt. Several participants explained that this was because they did not have the phone with them all the time. In fact, some participants shared their phones with other members of their household.

> *I am at home focusing on some other chores, after work, so I had forgotten that I had an important in-coming call, I am sorry* [P7 Luganda]
>
> *I saw the calls, they came in twice, but I did not know that I had to take the call. The caller number looked strange, it was not a usual call, so I thought it might be a conman from [location named], or another country* [P26 Runyakitara]

Eight call recipients (2 for each different language) reported inaudible calls during the first round of the pilot survey. They all struggled to listen in to make sense of what the survey was about, irrespective of the survey language. When we scrutinized these participants, some held ordinary mobile phones, while others used smart phones. However, all of them were in a rural setting, although based on routine phone calls, the network connectivity was fine. An alteration within the platform, improved audibility with the second and subsequent rounds of the pilot survey.

Two participants discussed the difficulties they encountered with skip patterns, which either altered the flow of the survey or curtailed their ability to complete the survey.

> *The tobacco screening question kept repeating itself, whether I pressed that I was a smoker, or not. It did not allow me to proceed to the next set of questions. Each time I punched in 1 or 3, it merely repeated the question until I was fed up and ended the survey* [P9 English]
>
> *The survey kept asking me what my age was, and as an example to enter 18 on the phone's keyboard if my age was eighteen. However, each time I entered my age, it kept on asking me the same question* [P16 English]

Similarly, the initial calls for other languages other than Luo (English, Luganda, and Runyakitara) had challenges with skip patterns, such as automatically moving to the next question irrespective of the selection of a prompted answer option. These required altering from the programming side within the IVR platform, following which the challenge of skip patterns was resolved, for subsequent survey testing rounds. There were also varied experiences of IVR-mobile phone survey participation related to the language of delivery itself.

Language-Related Issues

Successful Encounters

Within the first round of the pilot survey, 2 Luo speakers reported that their survey went very well. It was audible, it was clear, the timing was appropriate, the skip patterns worked very well, and the survey was easy to comprehend.

However, none of the participants who took the survey in any of the other 3 languages of the IVR described it as courteous. On the whole, after providing feedback to the recording studio and rerecording, the second round and subsequent rounds attained the required benchmark for language courtesy, question clarity, appropriate pace, audibility, and validity as gauged from the participant's feedback.

Unsuccessful Encounters

At least 7 respondents reported taking the survey in the wrong language which they could not comprehend, although the introduction of the IVR provided for selecting an appropriate language option. The quote below exemplifies the challenge of selecting the wrong language.

> *I am a Swahili speaker, but I received the survey and took it in Luganda. I am not sure if my answer options were accurate or not* [P33 Luganda]

Mobile Network Connectivity

Limited clarity of survey questions was only consistently reported by participants who were in rural locations—either on the farm, in a University, or in homes that were more than 250 km distant from the Capital city. Their mobile phone survey was generally inaudible, and it self-terminated after a couple of attempts of replying. The survey team therefore interpreted this as due to poor mobile network connectivity.

Survey Duration

Three participants responding in English voiced strong opinions based on their experience with the IVR survey lasting about 10 minutes that the information about the survey's duration in the introduction need to be altered from 20 to the realistic 10 minutes, to manage a participant's expectations. When asked how long the survey took, the majority of the participants who had completed the survey responded:

> *...about ten minutes...* [P7 English]

Thus, resonating well with the experiences of those voicing the concern on survey duration.

Perceptions About Study Incentives

When asked about what they thought of the method of incentive, all the participants were pleased about the promised lottery-type incentive for receiving airtime. Some also reported that it encouraged them complete the survey, as they stood a chance of winning this incentive. Furthermore, it was reported by a few that this type of incentive in research was generally new to them, but it did not really matter.

Discussion

General

This pilot survey aimed to explore the perceptions of users of and nonresponses to the survey. The key findings that favored IVR survey participation or completion included preference for short surveys of 10 minutes or shorter, preference for evening calls between 6 PM and 10 PM, preference for courteous language, caller's credibility, and favorable perceptions of the lottery-type incentive. While key findings curtailing participation or survey completion included if the voice was unclear, skip patterns were confusing, difficulties in interfacing with the phone to complete the survey, such as erroneous selection of digits for response options on both the ordinary and smart phones, suspicion about the caller's identity, and poor network connectivity for remote and rural participants.

Most of the participants in this study preferred their IVR calls between 6 PM and 10 PM, suggesting a preference for calls outside normal working hours. Intuitively, late evening call preference is related to a period of limited interruption from the rest of the day's competing demands. As reported from other studies [3,39,40], from a cultural perspective, it seems interruptive to receive a call while at work, especially if conducting formal work requiring team-effort, such as teaching in class, working in an operation theatre among others. In rural places where phone-charging is rationed to locations where there is power. It could be that phones are charged during part of the day making them inaccessible to a user, while in the evening the user catches up with missed phone calls. Similarly, if a phone is shared between a couple or household members, the individual that did not have it during the day might only have access in the evening when the phone holding partner or family member returns [3,39,40].

While for the majority, audible calls that were clear were received favorably, thus offering a promise to the acceptability of the mobile phone survey, the inaudible IVR calls, and those where skip patterns had errors compromised call completion. Considering that an IVR recipient requires to first listen to the voice call then to accurately interact with the phone to complete provided instructions. A high voice quality call that has simple and clear instructions is likely to maintain a respondent's interest. Future IVR surveys require an extensive piloting phase to ensure the qualities of voice clarity, simplicity, nonambiguity, and respondent's motivation or captivation for guaranteeing a successful IVR survey—as evidenced from other studies in sub-Saharan Africa decrying the IVR interaction [10,18-20].

Relatedly, a major finding in this study is that the quality of call reception (both the audibility and skip patterns) was related to programming challenges. In this IVR-mobile phone survey, the audios did not require rerecording, rather, an adjustment within the platform to increase their audibility. Likewise, the skip pattern errors were rectified within the platform, rather than with the audiorecording. This goes to confirm that in a software-based interface, programming, testing, and verifying appropriateness is important before roll-out of a software mobile health program, in this case the IVR-mobile phone survey platform. Contrary to the Janus-faced theory, which anticipates varied responses for each scenario [32,36,37], for the case of errors in the platform development, the resulting unintended errors elicited laborious encounters with the IVR survey for participants. Essentially, irrespective of the participant or their phone type, it appeared that errors in IVR delivery elicited annoyance and a poor experience with the IVR. Therefore, as with all communication strategies, piloting of IVR platforms (the communication channel) is important prior to mobile phone survey delivery for ensuring the expected quality of the IVR for recipients and the appropriate delivery of the intended message.

We found that caller credibility was crucially important as a motivator for survey participation. About a third of the participants reported that this survey from the Makerere University, with the Ministry of Health motivated their interest and participation. On the other hand, there was a sense of skepticism for some regarding responding to IVR (automated voice) calls because of fears of privacy—related to capturing individual's identity, conmen—relating to potential fraud, and political interests that were unwelcome, as reported elsewhere [38]. This finding demands the prior sensitization of a community about planned research, including conducting community mobilization, an important prestep in routine house-to-house surveys such as the census. Regarding IVR-mobile phone surveys, considering the competing agents using the automated voice calls for information dissemination, or mobilization, an alternative of a prior SMS text message clarifying the intent of an IVR-mobile phone survey, the planned survey timelines, and clearly stating the authority sanctioning the survey will be useful for increasing survey participation. If resources are inadequate for prior community mobilization or SMS messaging, a viable alternative might be for caller masking, such as using a label "Health Survey from Organization X" instead of an identifiable caller phone number that is unfamiliar to recipients.

Perceiving the language as courteous motivated participation and survey completion. This finding appears related to the social connectivity with one's local language—an important aspect of communication. Since IVR survey delivery mimics a human interaction, the quality of experience is important for motivating participation and completion, as reported in a Ghanaian study [10,18]. Therefore, language courtesy should be an important attribute considered in IVR survey development, piloting, and testing, prior to roll-out. Relatedly, although infrequent in this survey, taking the survey in an inappropriate language compromised the quality of survey responses. It might be that in a multilingual society such as Uganda, participants might not readily locate their preferred language in the IVR instructions. However, it is expected that this limitation will improve with increasing familiarity with IVRs, because respondents are not required to read and write but rather listen and act accordingly. Nonetheless, it is important for future IVR-mobile phone survey developers for multilingual settings to explore the extent of this problem—selecting the wrong language option. Also, a double-checking prompt would be useful to confirm that a given language is the appropriate choice.

Network connectivity was responsible in some instances for dropped calls. To explore the magnitude of this, a stratified analysis (for network operators) was conducted to assess the dropped calls. Among the 4 mobile phone providers, one had higher prevalence in the rural compared to others, and their call drop rate was considerably lower. Relatedly, the limited clarity of survey questions, alongside dropped calls was consistently reported by volunteers that were in rural locations—either on the farm, in a University, or in homes that are more than 250 km distant from the capital city. Their mobile phone survey was generally not too audible, and it self-terminated after a couple of attempts at talking back. We concluded that this was due to limited network connectivity. Mobile phone network coverage in Uganda is best in urban locations, likely due to economic motives of capturing high-density communities—thus depicting economies of scale. Therefore, developers and implementers of IVR surveys require strategies that capture rural populations, when representativeness is critical for answering survey objectives—such as considering oversampling of the rural remote populations.

Our finding for the preference for surveys lasting 10 minutes or shorter suggests existing competing work or leisure demands, thereby requiring shorter and precise mobile phone survey, as evidenced from the main Uganda IVR-mobile phone survey which lasted an average of 13 minutes, yet with a low completion rate of 35.2% [28]. While routine face-to-face surveys such as the Demographic and Health Survey conducted every 5 years in low- and middle- income countries may last an average of 1 hour. The absence of physical human interaction in the IVR encounter tends to remove the normative desire of avoiding disappointing the interviewer, lest the participant be judged as rude—frontstage, back-stage acting. IVR participants are in control of their survey's continuation or termination, which might shield their fears for potential retribution. Also, the lack of human interaction negates the opportunity to negotiate the timing and duration of the IVR. Therefore, commencing the IVR might somewhat rely on the curiosity of a respondent wanting to discover what the survey is about. However, considering that the average survey duration was about 10 minutes in this pilot, and there were no complaints that it took very long, this finding strongly implies that IVRs require brevity to maintain the interest of participants.

Conclusions and Recommendations

Our findings show the willingness of participants to take an IVR survey. Key attributes of an IVR survey with promise for high uptake and completion within a multilingual context include: a preference for evening calls, of high voice quality and clear instructions, lasting 10 minutes or shorter, from a credible caller, and in a courteous language.

Findings emphasize the need for extensive platform development in the testing period to ensure stability, prior to roll-out of an IVR survey. There is need to further evaluate these attributes to increase IVR acceptability and completion rates in such settings.

It appears from our findings that both ordinary and smartphone users encounter interactive challenges with an IVR call, thus emphasizing a need for education of the community on use of IVRs. Suspicion as to the credibility of the survey authority suggests a need for caller masking. There is need for further research to explore reasons for low completion rates of IVRs compared to face-to-face surveys and whether language selection and education status affect the quality of surveys.

Study Limitations

While this study used a qualitative methodology, user perceptions on IVR and nonresponse were collected through a phone interview and were not validated physically, which within the context of the study could have introduced some sociodesirability bias; however, phone interviews are a widely accepted method in qualitative research [24].

The pilot was limited to 3 nationally representative languages in addition to English, although 4 languages would have been more representative. Being an explorative study, the nuances from the findings might apply to the rest of the country, given that the cultural context is similar.

Additionally, at least a third of the participants were known to the first author who did the recruitment. This could have positively affected participation in the IVR and the qualitative interviews. However, both procedures followed standard research ethical practice after obtaining informed consent.

Acknowledgments

RT, HS, DG, GWP, and ER conceptualized this study, RT conducted field work, and coded the data with ER. ER provided study oversight. RT drafted the first manuscript, all authors participated in data analysis, read all versions of the manuscript and approved the final version. Funding for this study was provided by the Bloomberg philanthropies.

We thank the staff at Viamo (a global social enterprise), who supported the development and deployment of the IVR platform.

Conflicts of Interest

None declared.

References

1. GBD 2013 DALYsHALE Collaborators, Murray CJL, Barber RM, Foreman KJ, Abbasoglu Ozgoren A, Abd-Allah F, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990-2013: quantifying the epidemiological transition. Lancet 2015 Nov 28;386(10009):2145-2191 [FREE Full text] [doi: 10.1016/S0140-6736(15)61340-X] [Medline: 26321261]

2. WHO. Noncommunicable diseases, key facts. World Health Organization. Geneva, Switzerland; 2018 Jun 01. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2019-02-03]
3. Bloomfield GS, Vedanthan R, Vasudevan L, Kithei A, Were M, Velazquez EJ. Mobile health for non-communicable diseases in Sub-Saharan Africa: a systematic review of the literature and strategic framework for research. Global Health 2014 Jun 13;10:49 [FREE Full text] [doi: 10.1186/1744-8603-10-49] [Medline: 24927745]
4. Lee S, Cho Y, Kim S. Mapping mHealth (mobile health) and mobile penetrations in sub-Saharan Africa for strategic regional collaboration in mHealth scale-up: an application of exploratory spatial data analysis. Global Health 2017 Aug 22;13(1):63 [FREE Full text] [doi: 10.1186/s12992-017-0286-9] [Medline: 28830540]
5. WHO. Noncommunicable diseases and their risk factors STEPwise approach to surveillance (STEPS). World Health Organization. Geneva, Switzerland URL: https://www.who.int/ncds/surveillance/steps/en/ [accessed 2018-07-28]
6. Marchant E. Interactive voice response and radio for peacebuilding: a macro view of the literature and experiences from the field. Oxford University Research Archive. Pennsylvania, USA: Center for Global Communication Studies; 2016 Feb 29. URL: https://ora.ox.ac.uk/objects/uuid:ecc1668e-a814-48e8-93e6-39bcfd8e7c1d [accessed 2018-08-13]
7. Lee H, Friedman ME, Cukor P, Ahern D. Interactive voice response system (IVRS) in health care services. Nursing Outlook 2003 Dec;51(6):277-283. [doi: 10.1016/s0029-6554(03)00161-1]
8. Stritzke WGK, Dandy J, Durkin K, Houghton S. Use of interactive voice response (IVR) technology in health research with children. Behavior Research Methods 2005 Feb 1;37(1):119-126. [doi: 10.3758/bf03206405]
9. Falconi M, Johnston S, Hogg W. A scoping review to explore the suitability of interactive voice response to conduct automated performance measurement of the patient's experience in primary care. Prim Health Care Res Dev 2015 Aug 5;17(03):209-225. [doi: 10.1017/s1463423615000407]
10. Brinkel J, Dako-Gyeke P, Krämer A, May J, Fobil J. An investigation of users' attitudes, requirements and willingness to use mobile phone-based interactive voice response systems for seeking healthcare in Ghana: a qualitative study. Public Health 2017 Mar;144:125-133. [doi: 10.1016/j.puhe.2016.11.017] [Medline: 28274374]
11. Gibson DG, Pereira A, Farrenkopf BA, Labrique AB, Pariyo GW, Hyder AA. Mobile Phone Surveys for Collecting Population-Level Estimates in Low- and Middle-Income Countries: A Literature Review. J Med Internet Res 2017 May 05;19(5):e139 [FREE Full text] [doi: 10.2196/jmir.7428] [Medline: 28476725]
12. Duarte AC, Thomas SA. The Use of Phone Technology in Outpatient Populations: A Systematic Review. Open Nurs J 2016 Apr 30;10(1):45-58 [FREE Full text] [doi: 10.2174/1874434601610010045] [Medline: 27347255]
13. Heisler M, Halasyamani L, Resnicow K, Neaton M, Shanahan J, Brown S, et al. "I am not alone": the feasibility and acceptability of interactive voice response-facilitated telephone peer support among older adults with heart failure. Congest Heart Fail 2007 May;13(3):149-157 [FREE Full text] [doi: 10.1111/j.1527-5299.2007.06412.x] [Medline: 17541307]
14. Forster AJ, van Walraven C. Using an interactive voice response system to improve patient safety following hospital discharge. J Eval Clin Pract 2007 Jun;13(3):346-351. [doi: 10.1111/j.1365-2753.2006.00702.x]
15. Hettema JE, Hosseinbor S, Ingersoll KS. Feasibility and reliability of interactive voice response assessment of HIV medication adherence: research and clinical implications. HIV Clin Trials 2012 Jan 06;13(5):271-277 [FREE Full text] [doi: 10.1310/hct1305-271] [Medline: 23134627]
16. Gibson DG, Pariyo GW, Wosu AC, Greenleaf AR, Ali J, Ahmed S, et al. Evaluation of Mechanisms to Improve Performance of Mobile Phone Surveys in Low- and Middle-Income Countries: Research Protocol. JMIR Res Protoc 2017 May 05;6(5):e81 [FREE Full text] [doi: 10.2196/resprot.7534] [Medline: 28476729]
17. Pariyo GW, Wosu AC, Gibson DG, Labrique AB, Ali J, Hyder AA. Moving the Agenda on Noncommunicable Diseases: Policy Implications of Mobile Phone Surveys in Low and Middle-Income Countries. J Med Internet Res 2017 May 05;19(5):e115 [FREE Full text] [doi: 10.2196/jmir.7302] [Medline: 28476720]
18. Brinkel J, May J, Krumkamp R, Lamshöft M, Kreuels B, Owusu-Dabo E, et al. Mobile phone-based interactive voice response as a tool for improving access to healthcare in remote areas in Ghana - an evaluation of user experiences. Trop Med Int Health 2017 May 04;22(5):622-630 [FREE Full text] [doi: 10.1111/tmi.12864] [Medline: 28278352]
19. Daftary A, Hirsch-Moverman Y, Kassie GM, Melaku Z, Gadisa T, Saito S, et al. A Qualitative Evaluation of the Acceptability of an Interactive Voice Response System to Enhance Adherence to Isoniazid Preventive Therapy Among People Living with HIV in Ethiopia. AIDS Behav 2017 Nov 24;21(11):3057-3067 [FREE Full text] [doi: 10.1007/s10461-016-1432-8] [Medline: 27221743]
20. Howard AA, Hirsch-Moverman Y, Saito S, Gadisa T, Daftary A, Melaku Z. The ENRICH Study to evaluate the effectiveness of a combination intervention package to improve isoniazid preventive therapy initiation, adherence and completion among people living with HIV in Ethiopia: rationale and design of a mixed methods cluster randomized trial. Contemp Clin Trials Commun 2017 Jun;6:46-54 [FREE Full text] [doi: 10.1016/j.conctc.2017.03.001] [Medline: 28626811]
21. Bauermeister J, Giguere R, Leu C, Febo I, Cranston R, Mayer K, et al. Interactive Voice Response System: Data Considerations and Lessons Learned During a Rectal Microbicide Placebo Adherence Trial for Young Men Who Have Sex With Men. J Med Internet Res 2017 Jun 09;19(6):e207 [FREE Full text] [doi: 10.2196/jmir.7682] [Medline: 28600275]
22. Swendeman D, Jana S, Ray P, Mindry D, Das M, Bhakta B. Development and Pilot Testing of Daily Interactive Voice Response (IVR) Calls to Support Antiretroviral Adherence in India: A Mixed-Methods Pilot Study. AIDS Behav 2015 Jun 1;19 Suppl 2(S2):142-155 [FREE Full text] [doi: 10.1007/s10461-014-0983-9] [Medline: 25638037]

23. Chancellor TC, Priebe JS, Mkenda PA. Crowdsourcing Field Observations from Smallholder Farmers in Tanzania Using Interactive Voice Response. outlook pest man 2019 Jun 01;30(3):104-110. [doi: 10.1564/v30_jun_02]
24. Creswell JW. Research Design (International Student Edition): Qualitative, Quantitative and Mixed Methods Approaches. Thousand Oaks, California: SAGE Publications; Mar 15, 2013:1-304.
25. Pierannunzi C, Hu SS, Balluz L. A systematic review of publications assessing reliability and validity of the Behavioral Risk Factor Surveillance System (BRFSS), 2004-2011. BMC Med Res Methodol 2013 Mar 24;13:49. [doi: 10.1186/1471-2288-13-49] [Medline: 23522349]
26. Schneider KL, Clark MA, Rakowski W, Lapane KL. Evaluating the impact of non-response bias in the Behavioral Risk Factor Surveillance System (BRFSS). J Epidemiol Community Health 2012 Apr;66(4):290-295. [doi: 10.1136/jech.2009.103861] [Medline: 20961872]
27. WHO. Regional strategy for utilization of tobacco questions for surveys (TQS). WHO Regional Office for South-East Asia, New Delhi, India: World Health Organization; 2012. URL: https://apps.who.int/iris/handle/10665/205850 [accessed 2018-03-02]
28. Gibson DG, Wosu AC, Pariyo GW, Ahmed S, Ali J, Labrique AB, et al. Effect of airtime incentives on response and cooperation rates in non-communicable disease interactive voice response surveys: randomised controlled trials in Bangladesh and Uganda. BMJ Glob Health 2019;4(5):e001604 [FREE Full text] [doi: 10.1136/bmjgh-2019-001604] [Medline: 31565406]
29. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Michael Hubermann MBM, editor. The qualitative researcher's companion. Thousand Oaks, California: SAGE Publications, Inc; Feb 19, 2002:305-329.
30. Srivastava A, Thomson SB. Framework analysis: a qualitative methodology for applied policy research. Journal of Governance and Administration 2009;72 [FREE Full text]
31. Parkinson S, Eatough V, Holmes J, Stapley E, Midgley N. Framework analysis: a worked example of a study exploring young people’s experiences of depression. Qualitative Research in Psychology 2015 Nov 16;13(2):109-129. [doi: 10.1080/14780887.2015.1119228]
32. Arnold M. On the phenomenology of technology: the “Janus-faces” of mobile phones. Information and Organization 2003 Oct;13(4):231-256. [doi: 10.1016/s1471-7727(03)00013-7]
33. Ajzen I. The theory of planned behaviour is alive and well, and not ready to retire: a commentary on Sniehotta, Presseau, and Araújo-Soares. Health Psychol Rev 2015 Feb 12;9(2):131-137. [doi: 10.1080/17437199.2014.883474] [Medline: 26209198]
34. Conner M, Sparks P. Theory of planned behavior and health behavior. In: Conner M, Norman P, editors. Predicting Health Behavior. London, UK: Open University Press; 2005:170-221.
35. Glanz K, Rimer BK, Viswanath K. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behaviour and Health Education: Theory, research and Practice. San Francisco, California: John Wiley & Sons, Inc; 2008:1-516.
36. Wilken R. An exploratory comparative analysis of the use of metaphors in writing on the Internet and mobile phones. Social Semiotics 2013 Nov;23(5):632-647. [doi: 10.1080/10350330.2012.738999]
37. Allen DFN. Confronting the Janus-face. The armed forces and African transitional politics, PhD Thesis. Washington DC: Johns Hopkins University School of Advanced International Studies, Johns Hopkins University; 2018. URL: https://jscholarship.library.jhu.edu/bitstream/handle/1774.2/60030/ALLEN-DISSERTATION-2018.pdf?sequence=1&isAllowed=y [accessed 2019-12-14]
38. Mwaka E, Nakigudde J, Ali J, Ochieng J, Hallez K, Tweheyo R, et al. Consent for mobile phone surveys of non-communicable disease risk factors in low-resource settings: an exploratory qualitative study in Uganda. Mhealth 2019 Aug;5:26-26 [FREE Full text] [doi: 10.21037/mhealth.2019.07.05] [Medline: 31559271]
39. Zurovac D, Otieno G, Kigen S, Mbithi AM, Muturi A, Snow RW, et al. Ownership and use of mobile phones among health workers, caregivers of sick children and adult patients in Kenya: cross-sectional national survey. Global Health 2013 May 14;9(1):20 [FREE Full text] [doi: 10.1186/1744-8603-9-20] [Medline: 23672301]
40. Porter G. Mobile phones, mobility practices, and Transport organization in sub-Saharan Africa. Mobility in History 2015;6(1):1-13 DOI: https://doi.org/10.3167/mih.2015.060109 [FREE Full text] [doi: 10.3167/mih.2015.060109]

Abbreviations

HIV: human immunodeficiency virus
IVR: interactive voice response
SMS: short message service
WHO: World Health Organization

Edited by G Eysenbach; submitted 21.06.20; peer-reviewed by N Muinga, A Mahnke; comments to author 28.08.20; revised version received 11.09.20; accepted 28.10.20; published 03.12.20.

Please cite as:
Tweheyo R, Selig H, Gibson DG, Pariyo GW, Rutebemberwa E
User Perceptions and Experiences of an Interactive Voice Response Mobile Phone Survey Pilot in Uganda: Qualitative Study
JMIR Form Res 2020;4(12):e21671
URL: https://formative.jmir.org/2020/12/e21671
doi:10.2196/21671
PMID:33270037

Original Paper

Effects of Goal Type and Reinforcement Type on Self-Reported Domain-Specific Walking Among Inactive Adults: 2×2 Factorial Randomized Controlled Trial

Mindy L McEntee[1], PhD; Alison Cantley[1], MPH; Emily Foreman[1], BS; Vincent Berardi[2], PhD; Christine B. Phillips[3], PhD; Jane C. Hurley[1], PhD; Melbourne F. Hovell[4], PhD, MPH; Steven Hooker[4], PhD; Marc A Adams[1], PhD, MPH

[1]College of Health Solutions, Arizona State University, Phoenix, AZ, United States
[2]Department of Psychology, Crean School of Health and Behavioral Sciences, Chapman University, Orange, CA, United States
[3]Atria Senior Living, Louisville, KY, United States
[4]College of Health and Human Services, San Diego State University, San Diego, CA, United States

Corresponding Author:
Mindy L McEntee, PhD
College of Health Solutions
Arizona State University
425 N 5th Street
Phoenix, AZ
United States
Phone: 1 (602) 496 0860
Email: mindy.mcentee@asu.edu

Abstract

Background: WalkIT Arizona was a 2×2 factorial trial examining the effects of goal type (adaptive versus static) and reinforcement type (immediate versus delayed) to increase moderate to vigorous physical activity (MVPA) among insufficiently active adults. The 12-month intervention combined mobile health (mHealth) technology with behavioral strategies to test scalable population-health approaches to increasing MVPA. Self-reported physical activity provided domain-specific information to help contextualize the intervention effects.

Objective: The aim of this study was to report on the secondary outcomes of self-reported walking for transportation and leisure over the course of the 12-month WalkIT intervention.

Methods: A total of 512 participants aged 19 to 60 years (n=330 [64.5%] women; n=425 [83%] Caucasian/white, n=96 [18.8%] Hispanic/Latinx) were randomized into interventions based on type of goals and reinforcements. The International Physical Activity Questionnaire-long form assessed walking for transportation and leisure at baseline, and at 6 months and 12 months of the intervention. Negative binomial hurdle models were used to examine the effects of goal and reinforcement type on (1) odds of reporting any (versus no) walking/week and (2) total reported minutes of walking/week, adjusted for neighborhood walkability and socioeconomic status. Separate analyses were conducted for transportation and leisure walking, using complete cases and multiple imputation.

Results: All intervention groups reported increased walking at 12 months relative to baseline. Effects of the intervention differed by domain: a significant three-way goal by reinforcement by time interaction was observed for total minutes of leisure walking/week, whereas time was the only significant factor that contributed to transportation walking. A sensitivity analysis indicated minimal differences between complete case analysis and multiple imputation.

Conclusions: This study is the first to report differential effects of adaptive versus static goals for self-reported walking by domain. Results support the premise that individual-level PA interventions are domain- and context-specific and may be helpful in guiding further intervention refinement.

Trial Registration: Preregistered at clinicaltrials.gov: (NCT02717663) https://clinicaltrials.gov/ct2/show/NCT02717663

International Registered Report Identifier (IRRID): RR2-10.1016/j.cct.2019.05.001

(JMIR Form Res 2020;4(12):e19863) doi:10.2196/19863

KEYWORDS

exercise; population health; goals; reward; walking; mHealth; health promotion; health behavior; adaptive intervention; behavioral intervention

Introduction

Few adults meet the recommended physical activity (PA) guidelines, despite evidence of a strong dose-response relationship between PA and a range of health benefits, including decreased mortality [1], improved fitness, improved physical and mental functioning, and enhanced quality of life [2,3]. To date, interventions to increase PA have primarily focused on individual behavior change and had limited impact on population health [4]. Accordingly, there remains a need to develop potent interventions capable of producing an impact on a broader scale.

Mobile health (mHealth) technology provides a platform for increasing intervention reach and quickly tailoring content in response to an individual's behavior and preferences, but it requires evidence-based and scalable interventions to improve population health. Evidence suggests that behavioral strategies (eg, goal setting, financial reinforcement, feedback on performance) tend to be more effective than cognitive strategies (eg, education, motivation enhancement, self-belief) to increase PA in adults, yet no single strategy has consistently outperformed the rest [1]. It remains unclear how behavioral components may interact with each other in multicomponent interventions.

The WalkIT Arizona trial was designed to address this gap in the literature by studying the effects of an mHealth intervention that combined two evidence-based behavioral strategies—goal setting and positive reinforcement through the use of financial incentives—on objectively measured PA [5]. This paper reports on the study's secondary outcome: self-reported PA comprised of walking for transportation and leisure, and biking for transportation. Observational research has identified differences in the prevalence and correlates of domain-specific physical activities not otherwise captured via objective measures [6-9]. Thus, we examined intervention effects separately for transportation and leisure and for activity type (walking versus biking). Although studies suggest that participants' reported PA is less accurate than objective measures [10,11], self-reported data provide additional context that may be useful for understanding participant behavior and guiding the refinement of interventions.

We hypothesized that there would be significant main effects of goal type and reinforcement timing consistent with previous studies: those with adaptive goals would report more PA than those with static goals [12,13], and those receiving immediate reinforcement would report more PA than those receiving delayed reinforcement [13]. We further hypothesized that we would see a combined effect of the two intervention parameters, such that participants receiving adaptive goals and immediate reinforcement would report more walking than all other intervention groups. As the intervention did not specifically target any particular domain or type of PA, we hypothesized that there would be similar effects for transportation and leisure walking, and for transportation biking.

Methods

Study Design

WalkIT Arizona was a 2×2 factorial randomized trial evaluating the effects of goal setting (adaptive versus static goals) combined with financial incentives (immediate versus delayed reinforcement) to increase moderate-to-vigorous PA (MVPA) among insufficiently active adults. Participant selection was balanced across geographic information system–measured neighborhood walkability (high/low) and socioeconomic status (high/low) at the census block group level, with recruitment balanced across calendar months to adjust for seasonal effects. These design factors were important to the broader study testing multilevel interactions between individual-level intervention components and neighborhood design factors for PA maintenance during the follow-up period. The study was powered to detect a 2.1 minute/day difference in main effects and 4.2 minute/day difference in interaction effects between groups using accelerometer-measured MVPA, with a sample size of 120 participants per group. Participants completed a 12-month intervention followed by a 12-month observational follow-up period. Self-reported data were collected at baseline and at 6, 12, 18, and 24 months. Analyses presented here were conducted following completion of the 12-month intervention. This study was approved by the local institutional review board; further study details are published elsewhere [5].

Participants

Insufficiently active adults aged 19 to 60 years (N=512) were randomized for participation between May 2016 and May 2018. Participants were screened online and via phone interview prior to attending an office visit. Inactive status was verified following a 10-day baseline period in which participants were asked to wear a wrist-worn accelerometer during their normal activities. Baseline was extended beyond the scheduled 10 days for some participants due to issues with the accelerometer, problems with the mobile app, nonadherence to accelerometer wear protocol, or illness. Participants were told they would receive one of four different PA interventions. Baseline participant characteristics are displayed in Table 1.

Table 1. Baseline participant characteristics.

Characteristics	Total (N=512)	Adaptive goal + immediate reinforcement (n=128)	Static goal + immediate reinforcement (n=128)	Adaptive goal + delayed reinforcement (n=128)	Static goal + delayed reinforcement (n=128)
Age, mean (SD)	45.5 (9.1)	45.6 (9.5)	46.0 (8.9)	46.7 (8.6)	43.5 (9.3)
BMI, mean (SD)	33.9 (7.1)	33.7 (7.3)	33.8 (7.3)	33.6 (7.0)	34.5 (7.6)
Female, n (%)	330 (64.5)	82 (64.1)	80 (62.5)	81 (63.3)	87 (68.0)
Race and ethnicity[a], n (%)					
White	425 (84.0)	108 (84.4)	106 (82.8)	105 (82.0)	106 (82.8)
Black	31 (6.1)	5 (3.9)	9 (7.0)	9 (7.0)	8 (6.3)
American Indian or Alaskan Native	14 (2.7)	4 (3.1)	3 (2.3)	2 (1.6)	5 (3.9)
Asian	12 (2.3)	4 (3.1)	3 (2.3)	3 (2.3)	2 (1.6)
Native Hawaiian or other Pacific Islander	7 (1.4)	3 (2.3)	1 (0.8)	2 (1.6)	1 (0.8)
Hispanic or Latinx	96 (18.8)	22 (17.2)	26 (20.3)	24 (18.8)	24 (18.8)
Prefer not to answer	32 (6.3)	5 (3.9)	8 (6.3)	10 (7.8)	9 (7.0)
Current tobacco smoker, n (%)	26 (5.1)	3 (2.4)	10 (7.8)	5 (3.9)	8 (6.3)
Current e-smoker, n (%)	10 (2.0)	2 (1.6)	3 (2.4)	1 (0.8)	4 (3.2)
Married/living with partner, n (%)	346 (67.6)	82 (64.1)	85 (66.4)	93 (72.7)	86 (67.2)
Residence type, n (%)					
Single family house	392 (76.6)	94 (73.4)	93 (72.7)	104 (81.3)	101 (78.9)
Apartment	66 (12.9)	16 (12.5)	18 (14.1)	15 (11.7)	17 (13.3)
Years at current residence, mean (SD)	7.3 (7.4)	7.5 (7.9)	7.3 (7.5)	8.2 (7.0)	6.4 (7.1)
Children residing in household, n (%)	251 (49.0)	61 (47.7)	61 (47.5)	64 (50.0)	65 (50.8)
Number of children in household, mean (SD)	1.0 (1.2)	1.0 (1.2)	1.0 (1.3)	1.0 (1.3)	1.0 (1.1)
Household income, median	$60,000-$79,999	$80,000-$99,999	$60,000-$79,999	$60,000-$79,999	$80,000-$99,999
Education, median	College graduate	College graduate	College graduate	College graduate	College graduate
Employed full time, n (%)	390 (76.2)	98 (76.6)	97 (75.8)	94 (73.4)	101 (78.9)
Distance from home to work (meters), median	16,316	15,368	16,718	15,597	16,926

[a] Race/ethnicity cumulative is greater than 100%, as participants were asked to select all that applied.

Intervention Components

WalkIT Arizona intervention components have been described in detail elsewhere [5]. Briefly, participants in all four groups were provided with an activity monitor, which they were asked to wear on their wrist for 1 year, and a set of educational materials on the first intervention day. Throughout the intervention phase, participants could receive feedback via text message at any time on their accumulated MVPA minutes once they synced their activity monitor to the automated mHealth servers. All groups also received daily antecedent prompts using a pool of messages from our preliminary studies to evoke motivation, overcome barriers, remind about benefits, and provide other general health advice based on previous research. Text messaging was the primary communication channel between the mHealth system and participants. All feedback, goals, and reinforcement were communicated via this channel. In addition, participants were randomized to receive one of two types of daily goals (adaptive or static) and one of two types of financial incentives (immediate or delayed reinforcement).

Goal Setting

Participants allocated to the static goal group were asked to accumulate 30 minutes or more of MVPA daily throughout the 1-year intervention phase (eg, “Goal for 4/1 is 30 minutes”). A static goal of 30 minutes daily on at least 5 days per week aligns with current PA guidelines to obtain 150 minutes/week. Participants allocated to the adaptive goal group were assigned a goal daily based on a previously tested percentile-rank algorithm [12-14]. Unlike a static goal, adaptive goals had the potential to adjust up, down, or stay the same, depending on each participant's unique performance over the previous 9 accelerometer-measured observations (eg, “Goal for 4/1 is 7 minutes”). Each new adaptive goal was valid for the single day only. Regardless of goal type or financial reinforcement timing, participants were praised for meeting their goals (eg, “Well done! Goal met! 32 minutes today…goal for 4/2 is 15 minutes”).

Reinforcement Timing

Participants allocated to the delayed, noncontingent reinforcement group received escalating financial reinforcement on a 60-day schedule (ie, US $15 in month 2, US $30 in month 4, US $50 in month 6, US $75 in month 8, and US $95 in month 10) for participating and syncing their accelerometer. Participants allocated to the immediate reinforcement group earned points for meeting PA goals as described elsewhere [5]. For example, “Cheers, James! Goal met! 63 minutes yesterday. Reward points=100! Balance is 400 points. Goal for 7/1 is 35 minutes.” Points were worth US $0.01, and participants in the immediate reinforcement group were sent e-gift cards each time they accumulated US $5.00, since this was the minimum denomination for most gift cards. Participants in both the immediate and delayed reinforcement groups could select from a catalog of available e-gift card retailers (eg, Amazon, Target, Sephora, Home Depot, Walmart, Starbucks, etc) available from Tango Inc and change their selection at any time. E-gift cards were sent to participants using our automated mHealth system that was online 24 hours/day, 365 days per year.

Measures

Self-reported PA was assessed using sections 2, 4, and 5 of the International Physical Activity Questionnaire (IPAQ)-long form. The IPAQ was part of a larger battery of self-reported measures given at baseline, 6 months, and 12 months. Total self-reported minutes of walking per week were calculated separately for transportation and leisure domains; total self-reported minutes of biking per week were collected for transportation only using IPAQ scoring guidelines. The IPAQ has demonstrated comparable reliability and validity with other self-reported measures of PA [15] and provided an opportunity to examine changes to domain-specific physical activities not otherwise captured via objective measures.

Demographic information was collected via self-report at screening and baseline. Categories for assessing gender, race, and ethnicity were based on National Institutes of Health guidelines. Participants also reported their date of birth, education, marital status, residence type (single family house, apartment), number of adults and children in the household, and years residing at the current address.

Statistical Analyses

A generalized linear mixed model approach was used to examine intervention effects across treatment groups. Model selection was guided by distributional properties of outcome data. The outcome was total minutes of self-reported PA: time spent walking in the last week was computed separately for transportation and leisure, while time spent biking was limited to the transportation domain. All data distributions were positively skewed, with a relatively large number of zero values at each time point. We determined that hurdle models provided a better conceptual fit than zero-inflation models, as zero values could only result from remaining inactive during the intervention. Hurdle models contain two parts: a binary logit model, or hurdle, which estimated the likelihood of participants reporting 0 minutes/week of activity, and a truncated count regression model, which estimated the total number of reported minutes of activity in the last week (for those reporting values greater than 0). All count models were truncated at zero and used a negative binomial distribution to address overdispersion of data. Separate analyses were used to examine PA by activity type (walking versus biking) and domain (transportation versus leisure).

Negative binomial hurdle (NBH) models provided a nuanced examination of differences across intervention groups by activity type and domain. NBH models tested main effects and interactions among intervention parameters (goal type, reinforcement timing, time) with a random intercept allowed to vary by participant. Models 1 and 2 examined two-way interactions and included the third intervention parameter (ie, goal type or reinforcement timing) as a covariate. Model 3 examined a three-way goal by reinforcement by time interaction but had less power due to the additional interaction term. Time was specified as an ordered factor to allow for examination of linear and quadratic trends over the course of the intervention, as intervention components were assumed to have nonlinear effects at the individual level. All models were adjusted for census block-level socioeconomic status and neighborhood walkability since these factors were part of the broader research design. Predictor variables were kept consistent for hurdle and count models. All models were estimated using the generalized linear mixed models using template model builder (glmmTMB) package in R [16] and utilized the truncated_nbinom2 family for error distribution.

As glmmTMB models are estimated using only complete cases, we conducted a sensitivity analysis to examine the impact of missing data. Original analyses were compared to models estimated using (1) baseline values carried forward and (2) multiple imputation. Multiple imputation was performed using the multivariate imputation by chained equations (MICE) package [17] with 12 iterations, corresponding to the percentage of missing data [18]. Effects were pooled separately for count and hurdle models. The effects package [19,20] was used to visualize interactions.

Results

Participant flow is depicted in Figure 1. Participants reported a similar pattern of unadjusted walking activity across intervention groups: both leisure and transportation walking time increased from baseline to 6 months, then decreased from 6 months to 12 months. Across groups, reported walking time at 12 months remained greater than baseline for both leisure ($F_{1.9,764.5}$=35.79; $P<.001$; partial eta-square=0.08) and transportation ($F_{1.6,653.7}$=15.63; $P<.001$; partial eta-square=0.04). Biking for transportation was reported by a small subset of participants at each time point. Mean reported transportation biking for the entire sample showed a nonsignificant increase from baseline to 12 months; the trajectory of unadjusted transportation biking varied by intervention group. Participants with at least 1 missing self-reported PA data point (121/512, 23.6%) were more likely to report living with a partner but did not significantly differ from those with complete data on any other demographic characteristic. Split violin plots comparing the distribution of self-reported walking time by goal and reinforcement type are depicted in Figure 2. Mean self-reported PA times by intervention group, activity type, and domain are shown in Table 2.

Figure 1. Participant flow.

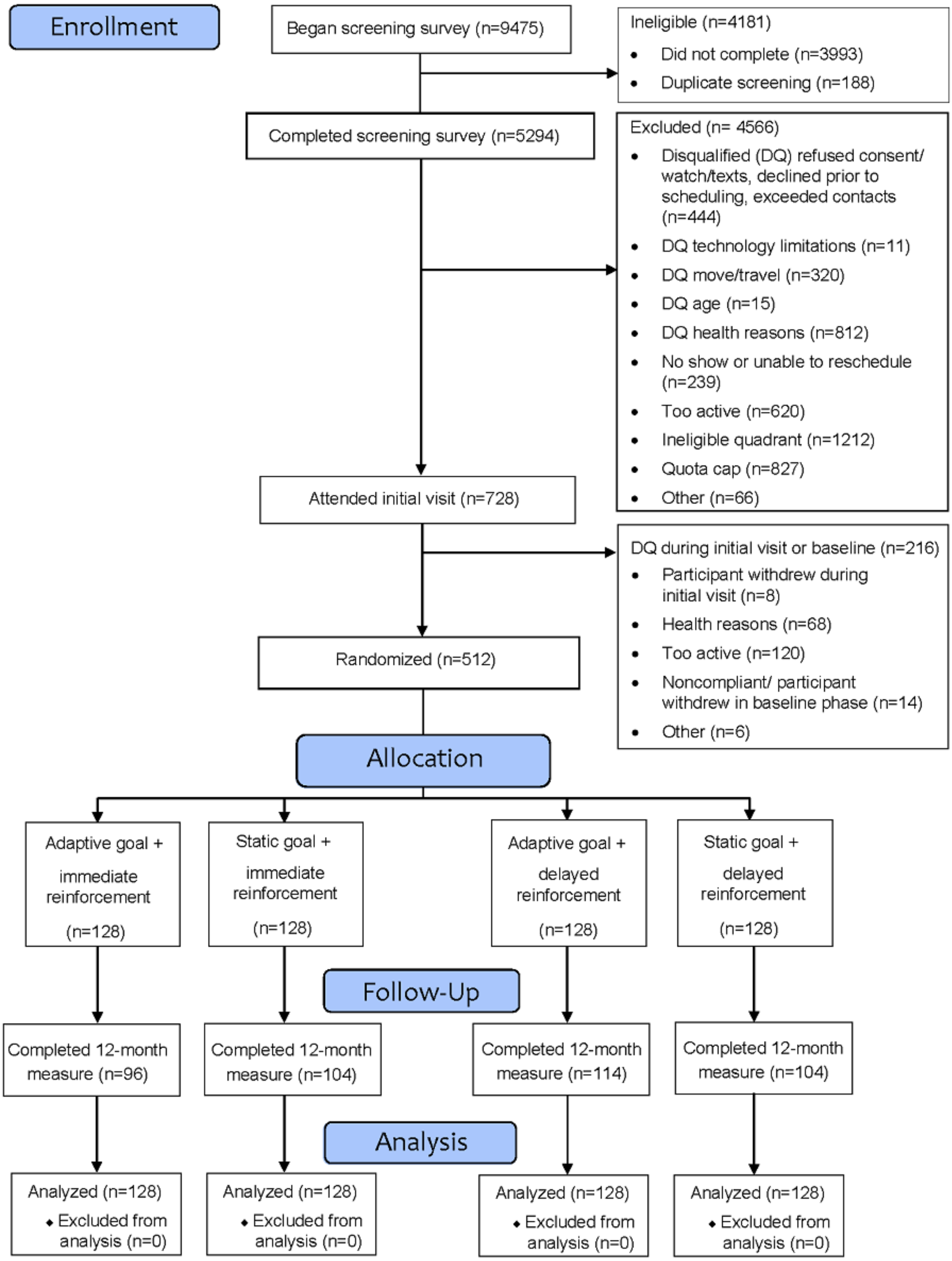


Figure 2. Split violin plots showing distribution of self-reported walking at baseline (BL), 6 months (6M), and 12 months (12M). Horizontal lines indicate 25th, 50th, and 75th percentiles computed from density estimates.

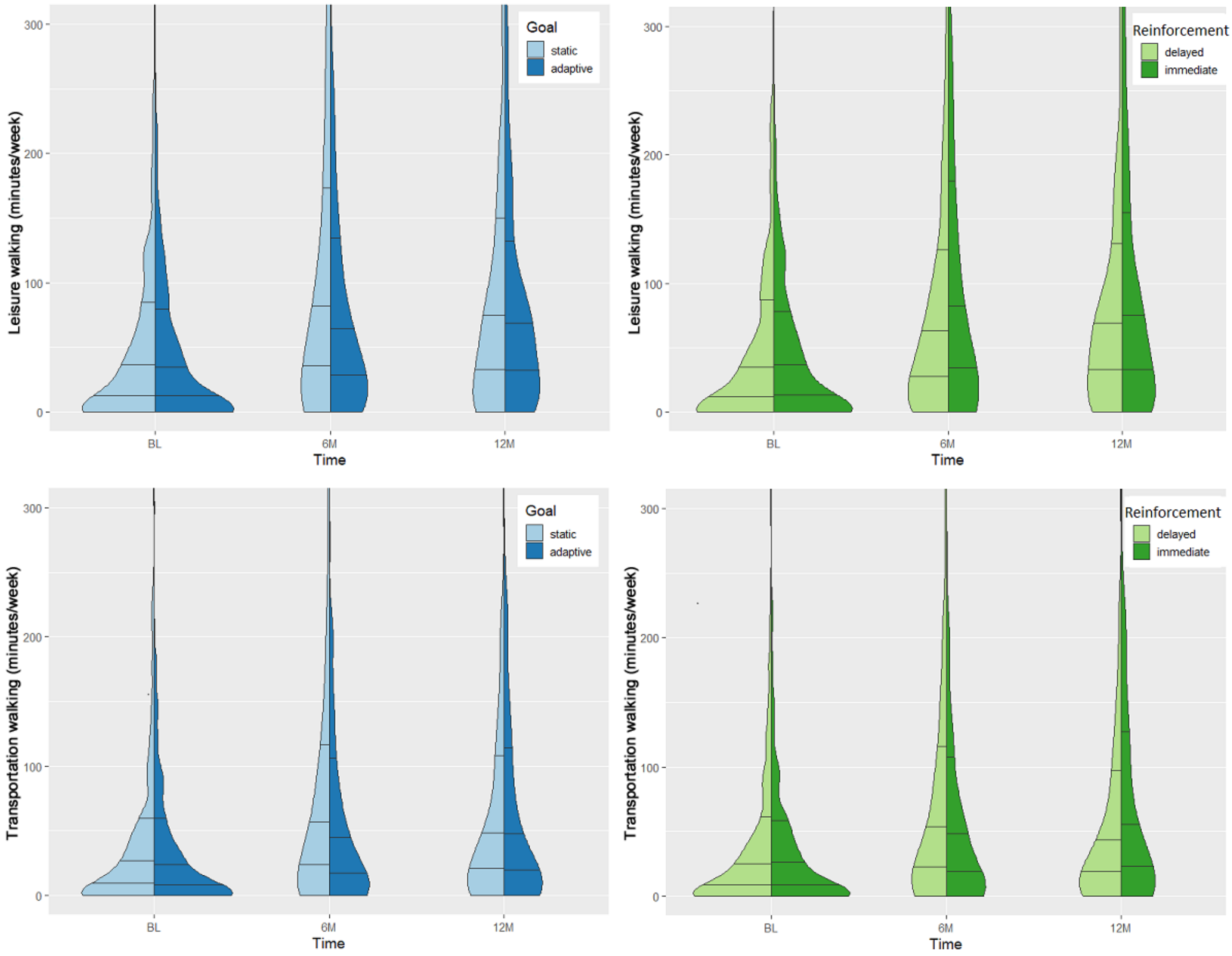


Table 2. Self-reported leisure walking, transportation walking, and transportation biking (minutes/week).

Self-reported physical activity	Total (N=512)	Adaptive goal + immediate reinforcement (n=128)	Static goal + immediate reinforcement (n=128)	Adaptive goal + delayed reinforcement (n=128)	Static goal + delayed reinforcement (n=128)
Leisure walking					
Baseline					
Mean[a] (SD)	45.3 (77.0)	41.1 (93.5)	36.6 (61.7)	37.9 (87.6)	53.3 (115.0)
Reported >0 min[b], n (%)[c]	296 (57.8)	73 (57.0)	75 (58.6)	70 (54.7)	78 (60.9)
6 months					
Mean (SD)	97.3 (128.2)	78.7 (155.2)	102.6 (196.2)	80.1 (170.6)	80.1 (100.8)
Reported >0 min, n (%)	331 (76.6)	80 (20.0)	79 (71.2)	91 (80.5)	81 (75.0)
12 months					
Mean (SD)	89.8 (110.3)	78.0 (113.0)	65.7 (87.9)	58.5 (92.4)	58.0 (78.7)
Reported >0 min, n (%)	329 (78.7)	76 (79.2)	79 (76.0)	95 (83.3)	79 (76.0)
Transportation walking					
Baseline					
Mean (SD)	42.2 (91.4)	40.2 (51.5)	38.0 (48.9)	39.4 (62.1)	63.4 (120.6)
Reported >0 min, n (%)	297 (58.0)	76 (59.4)	73 (57.0)	67 (52.3)	81 (63.3)
6 months					
Mean (SD)	85.5 (160.0)	93.2 (113.5)	117.6 (155.7)	92.9 (126.2)	84.9 (110.0)
Reported >0 min, n (%)	305 (70.0)	67 (66.3)	80 (71.4)	77 (67.0)	81 (75.0)
12 months					
Mean (SD)	64.6 (93.4)	83.8 (95.6)	102.2 (133.1)	86.8 (102.2)	86.4 (106.9)
Reported >0 min, n (%)	296 (70.6)	67 (69.8)	74 (71.2)	78 (67.8)	77 (74.0)
Transportation biking					
Baseline					
Mean (SD)	4.6 (25.8)	9.3 (43.1)	4.9 (23.2)	2.9 (11.1)	1.0 (9.4)
Reported >0 min, n (%)	33 (6.4)	12 (9.1)	9 (6.9)	9 (7.0)	3 (2.3)
6 months					
Mean (SD)	5.4 (29.2)	5.5 (22.8)	4.2 (22.9)	4.5 (18.6)	7.7 (45.5)
Reported >0 min, n (%)	33 (7.5)	11 (10.6)	6 (5.3)	10 (8.7)	6 (5.6)
12 months					
Mean (SD)	5.6 (25.7)	3.9 (17.0)	2.6 (10.8)	9.9 (35.8)	5.4 (29.5)
Reported >0 min, n (%)	34 (8.0)	6 (6.1)	9 (8.5)	13 (11.3)	6 (5.8)

[a] Calculated means include respondents reporting no physical activity (ie, 0 minutes).

[b] min: minutes.

[c] Percentages are based on total number of participant responses received for each time point.

The subheadings below indicate the interaction specified in NBH models, with other remaining parameters entered as covariates. As sensitivity analysis revealed little impact of missing data, the results discussed below are for complete cases; model parameters using multiple imputation are presented in Multimedia Appendices 1 to 6. Any differences between complete case analysis and multiple imputation results are noted below.

Leisure Walking

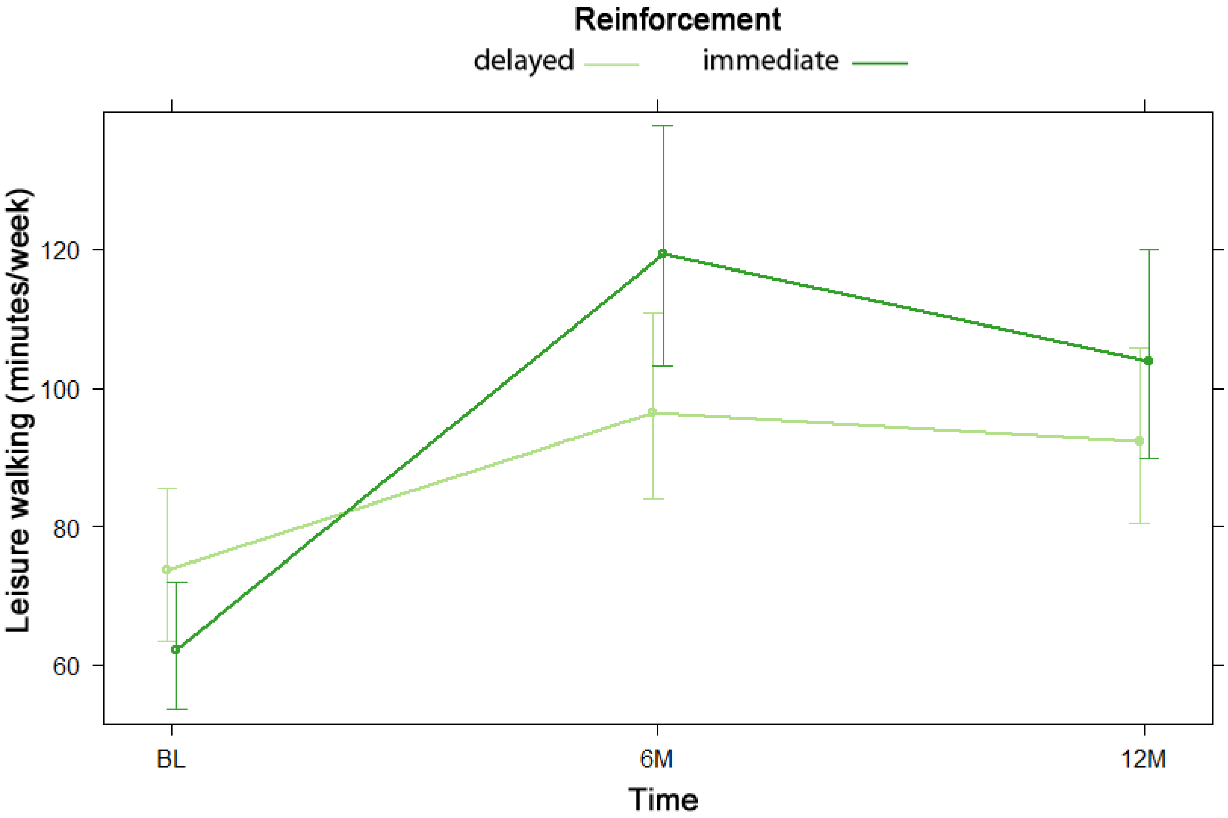
Model 1: Goal by Time

The overall proportion of participants reporting any (versus no) leisure walking increased from 57.8% (296/512) at baseline to 78.7% (329/418) at 12 months, as shown in Table 2. NBH model results are displayed as odds and risk ratios in Multimedia Appendix 7. For the hurdle model, participants with static goals had greater odds of reporting any (versus no) leisure walking at baseline, whereas those with adaptive goals had greater odds of reporting any (versus no) leisure walking at 6 and 12 months, although this interaction was nonsignificant (*P*=.08). For the count model, consisting only of participants who reported leisure walking, those who received adaptive goals reported 17% fewer minutes of leisure walking compared to those with static goals (95% CI 0.72-0.95). Lack of a significant goal by time interaction indicated that leisure walking did not differ by goal type across time points.

Model 2: Reinforcement by Time

NBH model results displaying odds and risk ratios are shown in Multimedia Appendix 8. For the hurdle model, there was no difference in the odds of reporting any (versus no) leisure walking based on reinforcement type, nor was there any significant reinforcement by time interaction effect. For the count model, consisting only of participants who reported leisure walking, there was a significant reinforcement by time interaction, indicating both linear (risk ratio [RR]=1.22, 95% CI 1.02-1.46) and quadratic (RR=0.79, 95% CI 0.67-0.94) effects. As shown in Figure 3, while all participants reported increased leisure walking time from baseline to 6 months, this increase was greater among participants who received immediate compared with delayed reinforcement. From 6 to 12 months, those with immediate reinforcement reported a decrease in time spent walking for leisure, whereas those with static reinforcement reported a very slight increase.

Figure 3. Reinforcement by time interaction in negative binomial count model 2 for leisure walking at baseline (BL), 6 months (6M), and 12 months (12M).



Model 3: Goal by Reinforcement by Time

Conditional estimates for an NBH model testing a three-way goal by reinforcement by time interaction and nested two-way interactions for leisure walking are displayed in Table 3. For the hurdle model, the results mirrored model 1: participants with static goals had greater odds of reporting any leisure walking at baseline, whereas those with adaptive goals had greater odds of reporting any leisure walking at 6 and 12 months, although this goal by time interaction was nonsignificant (*P*=.07). For the count model, which consisted only of participants who endorsed leisure walking, the mean reported time was 96.84 minutes/week (95% CI 81.65-114.84). Figure 4 shows a significant three-way interaction in which the reinforcement timing by linear time interaction effect varied significantly by goal type (RR=0.68, 95% CI 0.48-0.97).

Table 3. Negative binomial hurdle model examining goal by reinforcement by time (model 3) for leisure walking.

	Zero hurdle model		Count model	
Parameter[a]	OR[b,c] (95% CI)	*P* value	RR[d] (95% CI)	*P* value
Intercept	3.45 (2.30-5.18)	<.001****	96.84 (81.65-114.84)	<.001****
Socioeconomic status block (high)	0.82 (0.60-1.12)	.165	0.89 (0.77-1.02)	.131
Walkability block (high)	0.94 (0.69-1.29)	.641	1.03 (0.90-1.18)	.637
Goal (adaptive)	1.15 (0.74-1.80)	.540	0.84 (0.69-1.01)	.067*
Reinforcement (immediate)	0.88 (0.57-1.37)	.512	1.07 (0.88-1.30)	.580
Time: linear	1.80 (1.15-2.81)	.009***	1.13 (0.94-1.34)	.184
Time: quadratic	0.70 (0.43-1.11)	.129	0.92 (0.78-1.10)	.368
Goal by time: linear	1.79 (0.95-3.39)	.073*	1.09 (0.85-1.40)	.488
Goal by time: quadratic	0.90 (0.46-1.79)	.771	1.00 (0.79-1.28)	.987
Reinforcement by time: linear	1.07 (0.57-2.02)	.875	1.47 (1.15-1.89)	.002***
Reinforcement by time: quadratic	1.14 (0.60-2.20)	.678	0.70 (0.55-0.90)	.005***
Goal by reinforcement	1.07 (0.57-2.00)	.746	1.01 (0.76-1.32)	.842
Goal by reinforcement by time: linear	0.69 (0.28-1.71)	.465	0.68 (0.48-0.97)	.033**
Goal by reinforcement by time: quadratic	0.80 (0.31-2.08)	.711	1.28 (0.90-1.81)	.156

[a] Referent groups for parameters are listed in parentheses.

[b] Odds ratio (OR) reflects the odds of reporting any leisure walking (versus none).

[c] OR, risk ratio (RR), and 95% CI are exponentiated coefficients of conditional estimates.

[d] RR reflects the proportional increase (values >1) or decrease (values <1) in non-zero leisure walking minutes/week associated with a one unit change in the predictor.

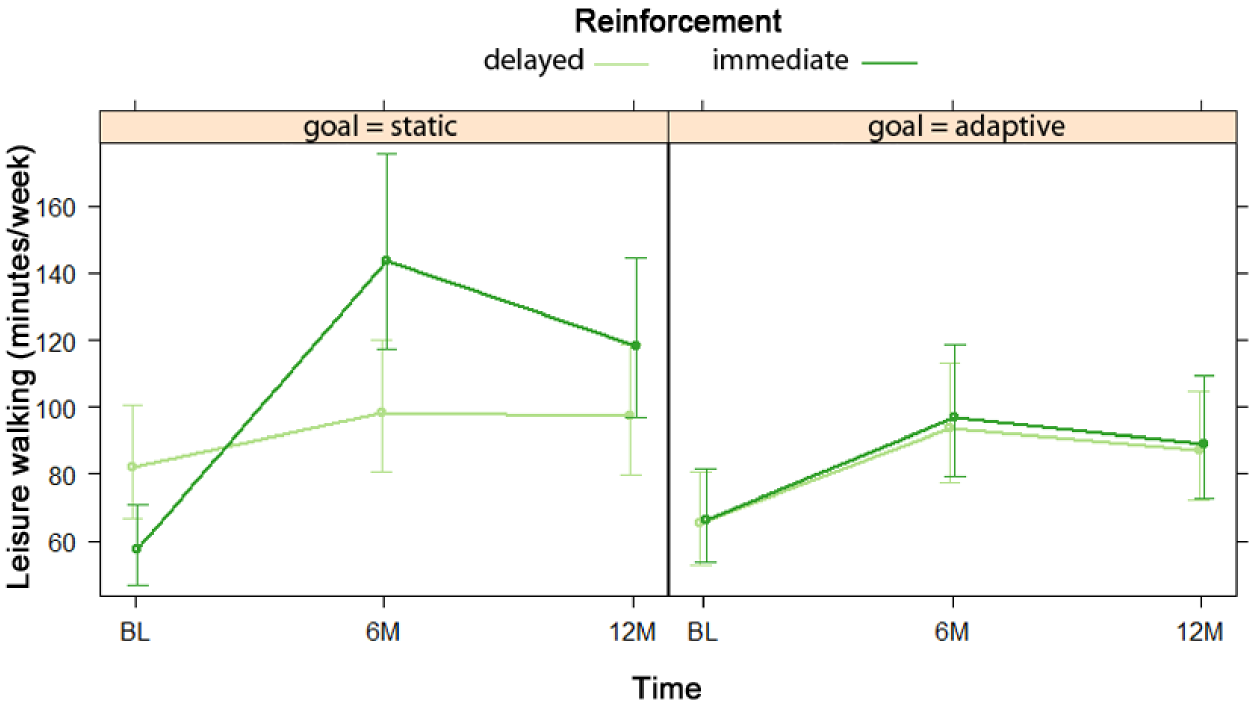
*P<.1

**P<.05,

***P<.01,

****P<.001.

Figure 4. Goal by reinforcement by time interaction in negative binomial count model 3 for leisure walking at baseline (BL), 6 months (6M), and 12 months (12M).



Transportation Walking

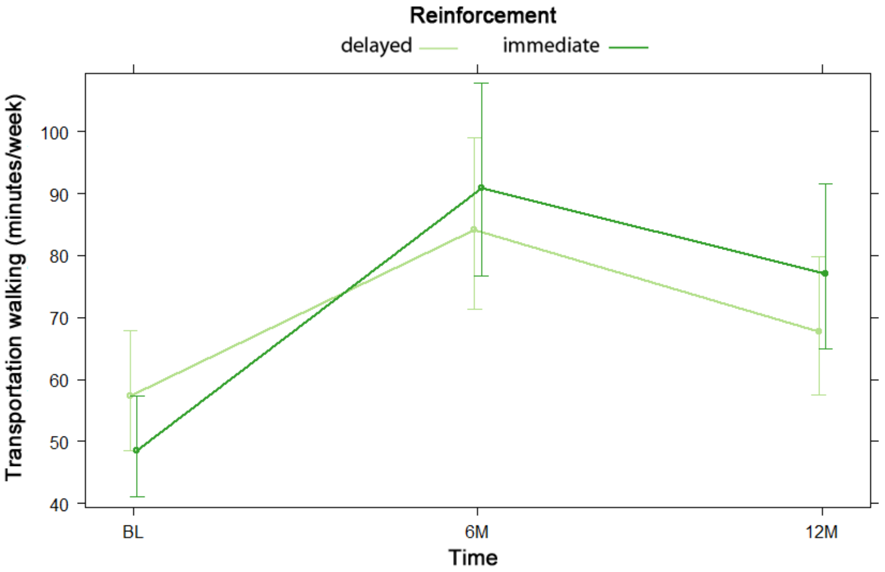
Model 1: Goal by Time

The overall proportion of participants reporting any (versus no) transportation walking increased from 58.0% (297/512) at baseline to 70.6% (296/419) at 12 months, as shown in Table 2. NBH model results are displayed as odds and risk ratios in Multimedia Appendix 9. For the hurdle model, there was no significant difference in odds of reporting any transportation walking by goal type, nor were there any significant goal by time interaction effects. For the count model, which consisted only of participants who reported transportation walking, there were no significant differences in reported transportation walking time by goal type, nor were there any significant goal by time interaction effects.

Model 2: Reinforcement by Time

NBH model results are shown as odds and risk ratios in Multimedia Appendix 10. For the hurdle model, there was no significant difference in odds of reporting any transportation walking by reinforcement type, nor was there any significant reinforcement by time interaction effect. For the count model, which consisted only of participants who endorsed transportation walking, time spent walking for transportation did not differ by reinforcement type. As shown in Figure 5, there was a significant reinforcement by time interaction such that participants with immediate reinforcements outperformed those with delayed reinforcements, with a greater increase in transportation walking from baseline to 6 months, and a smaller decrease from 6 to 12 months. However, this effect was no longer significant in multiple imputation analysis (*P*=.07).

Figure 5. Reinforcement by time interaction in negative binomial count model 2 for transportation walking at baseline (BL), 6 months (6M), and 12 months (12M).



Model 3: Goal by Reinforcement by Time

Conditional estimates for an NBH model testing a three-way goal by reinforcement by time interaction and nested two-way interactions for transportation walking are shown in Table 4. For the hurdle model, there were no independent or interaction effects for goal type or reinforcement timing on odds of reporting any (versus no) transportation walking. Among participants who endorsed transportation walking in the count model, mean reported time was 81.98 minutes/week (95% CI 66.04-101.77). There were no significant independent or interaction effects for goal type or reinforcement timing on reported transportation walking minutes/week.

Table 4. Negative binomial hurdle model examining goal by reinforcement by time (model 3) for transportation walking.

Parameter[a]	Zero hurdle model OR[b,c] (95% CI)	 *P* value	Count model RR[d] (95% CI)	 *P* value
Intercept	3.00 (1.80-4.98)	<.001***	81.98 (66.04-101.77)	<.001***
Socioeconomic status block (high)	0.79 (0.54-1.17)	.240	0.70 (0.59-0.83)	<.001***
Walkability block (high)	1.83 (1.23-2.71)	.003**	1.03 (0.86-1.22)	.711
Goal (adaptive)	0.65 (0.37-1.13)	.132	0.99 (0.77-1.26)	.912
Reinforcement (immediate)	0.81 (0.46-1.43)	.417	1.00 (0.79-1.28)	.956
Time: linear	1.67 (1.03-2.72)	.038*	1.07 (0.89-1.29)	.450
Time: quadratic	0.75 (0.45-1.25)	.270	0.75 (0.63-0.91)	.003**
Goal by time: linear	1.24 (0.64-2.41)	.526	1.10 (0.84-1.44)	.489
Goal by time: quadratic	0.81 (0.40-1.64)	.557	1.12 (0.86-1.47)	.405
Reinforcement by time: linear	1.13 (0.58-2.22)	.745	1.22 (0.93-1.59)	.120
Reinforcement by time: quadratic	0.79 (0.39-1.63)	.545	0.95 (0.72-1.24)	.687
Goal by reinforcement	1.39 (0.63-3.06)	.299	1.03 (0.69-1.38)	.891
Goal by reinforcement by time: linear	0.67 (0.26-1.75)	.416	1.01 (0.69-1.49)	.881
Goal by reinforcement by time: quadratic	1.72 (0.63-4.70)	.292	1.07 (0.73-1.57)	.735

[a] Referent groups for parameters are listed in parentheses.

[b] Odds ratio (OR) reflects the odds of reporting any leisure walking (versus none).

[c] OR, risk ratio (RR), and 95% CI are exponentiated coefficients of conditional estimates.

[d] RR reflects the proportional increase (values >1) or decrease (values <1) in non-zero transportation walking minutes/week associated with a one unit change in the predictor.

*$P<.05$.

**$P<.01$.

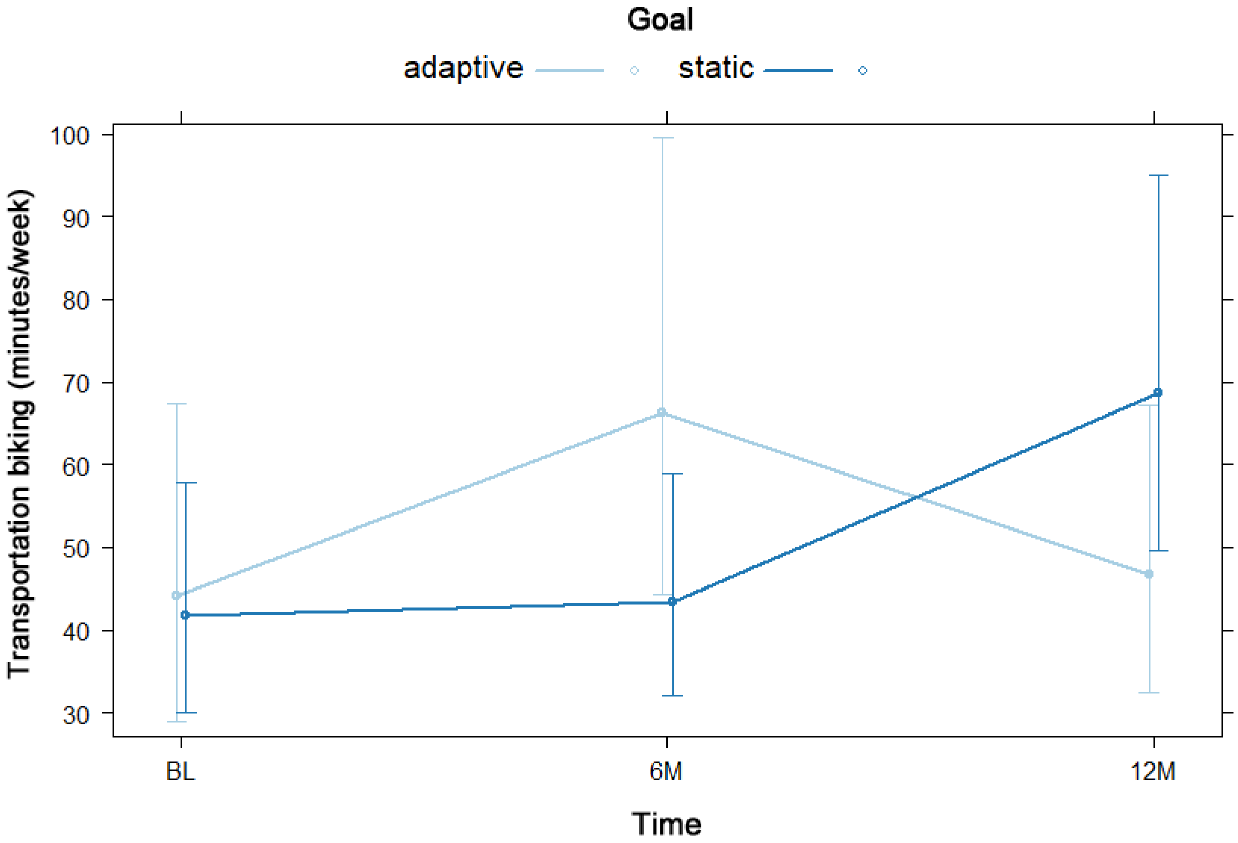
***$P<.001$.

Transportation Biking

Model 1: Goal by Time

The overall proportion of participants reporting any (versus no) transportation biking increased from 6.4% (33/512) at baseline to 8.0% (34/425) at 12 months, as shown in Table 2. NBH model results are displayed as odds and risk ratios in Multimedia Appendix 11. For the hurdle model, there was no significant difference in odds of reporting any transportation biking by goal type, nor was there any significant goal by time interaction effect. For the count model, which consisted only of participants who reported transportation biking, there was no significant difference in reported transportation biking time by goal type. There was a significant goal by time interaction such that participants with static goals reported more transportation biking than those with adaptive goals at baseline and 6 months, whereas participants with adaptive goals reported more transportation biking than those with static goals at 12 months, as shown in Figure 6.

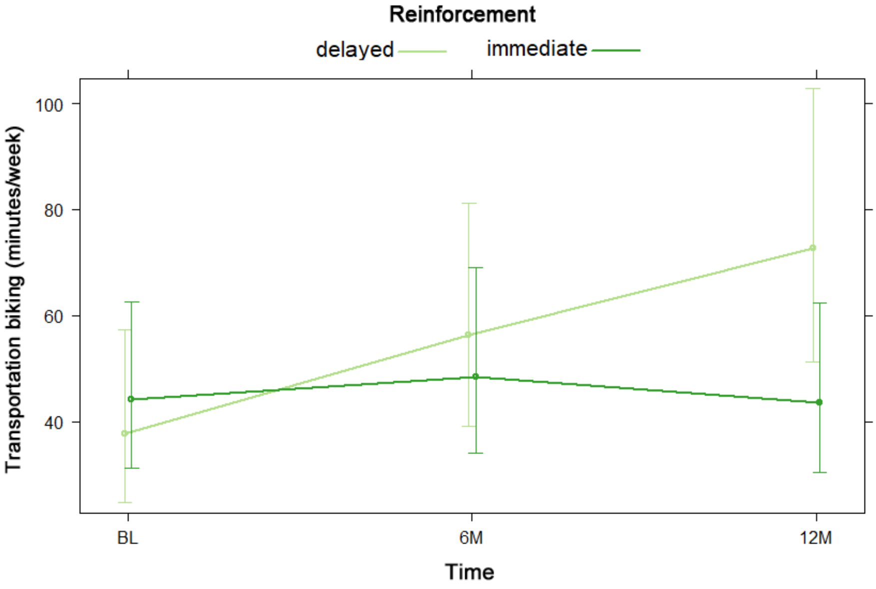
Figure 6. Goal by time interaction in negative binomial count model 1 for transportation biking at baseline (BL), 6 months (6M), and 12 months (12M).



Model 2: Reinforcement by Time

NBH model results are shown as odds and risk ratios in Multimedia Appendix 12. For the hurdle model, there was no significant difference in odds of reporting any transportation biking by reinforcement type, nor was there any significant reinforcement by time interaction effect. For the count model, which consisted only of participants who endorsed transportation biking, there was no significant difference in transportation biking by reinforcement type. There was a significant reinforcement by time interaction such that participants with immediate reinforcement reported more transportation biking than those with delayed reinforcement at baseline, whereas participants with delayed reinforcement reported more transportation biking than those with immediate reinforcement at 6 and 12 months, as shown in Figure 7.

Figure 7. Reinforcement by time interaction in negative binomial count model 2 for transportation biking at baseline (BL), 6 months (6M), and 12 months (12M).



Model 3: Goal by Reinforcement by Time

Conditional estimates for an NBH model testing a three-way goal by reinforcement by time interaction and nested two-way interactions for transportation biking are shown in Table 5. For the hurdle model, there was no independent effect of goal or reinforcement type on odds of reporting any (versus no) transportation biking. There was a significant three-way interaction for the hurdle model such that likelihood of reporting any biking varied over time by reinforcement and goal type, as shown in Figure 8. Among participants who endorsed transportation biking in the count model, the mean reported time was 53.31 minutes/week (95% CI 30.19-94.14). There were no significant independent effects of goal or reinforcement type on transportation biking minutes/week. There was a significant three-way interaction for the count model, such that the effect of reinforcement differed by goal and time (RR=3.79, 95% CI 1.71-8.40). As shown in Figure 9, participants with delayed reinforcement and static goals reported the greatest increase in transportation biking from baseline to 12 months, whereas those with immediate reinforcement and static goals were the only group to report a decrease in transportation biking time. Immediate reinforcement had the opposite effect from 6 to 12 months among participants with adaptive goals.

Figure 8. Goal by reinforcement by time interaction for negative binomial hurdle model 3 for transportation biking at baseline (BL), 6 months (6M), and 12 months (12M).

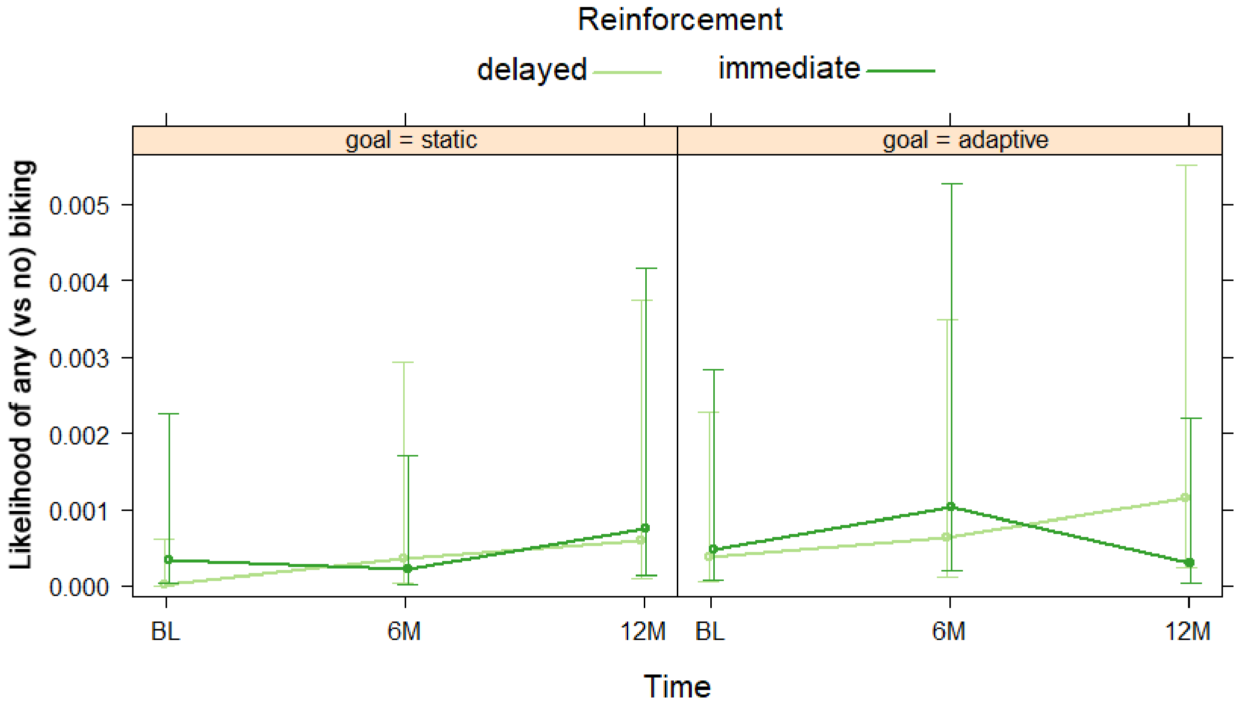


Figure 9. Goal by reinforcement by time interaction in negative binomial count model 3 for transportation biking at baseline (BL), 6 months (6M), and 12 months (12M).

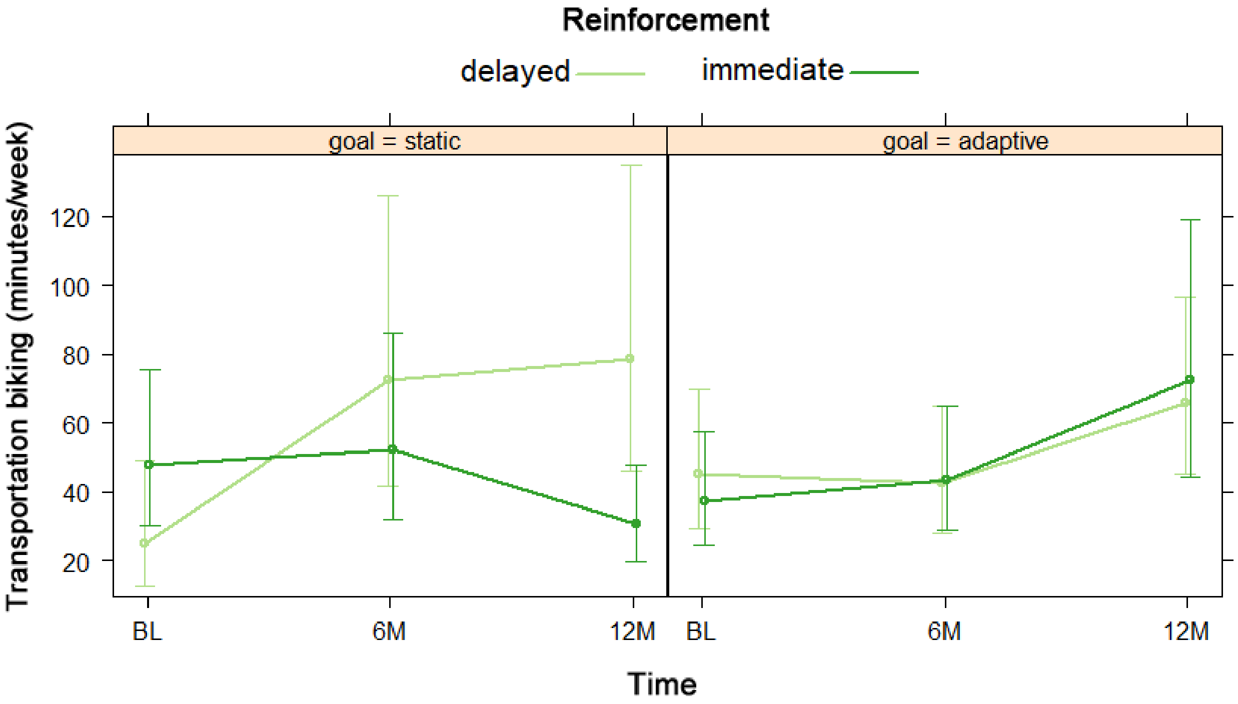


Table 5. Negative binomial hurdle model examining goal by reinforcement by time (model 3) for transportation biking.

Parameter[a]	Zero hurdle model OR[b,c] (95% CI)	*P* value	Count model RR[d] (95% CI)	*P* value
Intercept	0.016 (0.00002-0.001)	<.001****	53.31 (30.19-94.14)	<.001****
Socioeconomic status block (high)	1.32 (0.40-4.36)	.651	0.74 (0.52-1.06)	.104
Walkability block (high)	0.85 (0.26-2.79)	.794	1.30 (0.91-1.85)	.151
Goal (adaptive)	3.87 (0.57-26.16)	.165	0.96 (0.56-1.66)	.895
Reinforcement (immediate)	2.25 (0.30-16.93)	.432	0.82 (0.45-1.47)	.496
Time: linear	10.22 (1.27-82.43)	.029**	2.26 (1.38-3.69)	.001***
Time: quadratic	0.39 (0.08-1.95)	.250	0.67 (0.44-1.01)	.055*
Goal by time: linear	0.21 (0.02-2.05)	.182	0.58 (0.32-1.06)	.076*
Goal by time: quadratic	2.68 (0.42-17.12)	.297	1.84 (1.05-3.23)	.034**
Reinforcement by time: linear	0.17 (0.02-1.85)	.147	0.32 (0.18-0.59)	<.001****
Reinforcement by time: quadratic	5.10 (0.67-39.07)	.117	1.16 (0.68-1.99)	.590
Goal by reinforcement	0.36 (0.03-4.69)	.436	1.19 (0.57-2.49)	.637
Goal by reinforcement by time: linear	1.91 (0.12-29.97)	.645	3.79 (1.71-8.40)	.001***
Goal by reinforcement by time: Quadratic	0.08 (0.01-0.99)	.049**	0.82 (0.39-1.70)	.588

[a] Referent groups for parameters are listed in parentheses.

[b] Odds ratio (OR) reflects the odds of reporting any leisure walking (versus none).

[c] OR, risk ratio (RR), and 95% CI are exponentiated coefficients of conditional estimates.

[d] RR reflects the proportional increase (values >1) or decrease (values <1) in non-zero transportation biking minutes/week associated with a one unit change in the predictor.

*$P<.1$.

*$P<.05$.

**$P<.01$.

***$P<.001$.

Discussion

Principal Findings

This study reported the secondary outcomes of the WalkIT Arizona trial, a 12-month mHealth intervention combining goal setting (adaptive versus static goals) with financial reinforcement (immediate versus delayed) to increase PA among insufficiently active adults. Analyses examined differences in self-reported PA by intervention group, activity type (walking, biking), and activity domain (transportation, leisure). It is notable that only a small subset of participants (no more than 8% of the sample at any time point) reported biking, which was limited to the transportation domain. While walking was an activity readily accessible to all participants, biking required additional equipment and skills, including perceived comfort and safety while riding. Biking for transportation likely also required additional planning with regard to route, weather, and storage. For these reasons, we have focused our discussion on walking, as these findings represent a larger proportion of the study sample and are more likely to be generalizable.

All intervention groups reported greater time walking at 12 months relative to baseline for both leisure and transportation, with differences in the trajectory of walking time observed by group and domain. Closer examination of effects using NBH count models indicated differences in duration of walking time by domain: the independent effect of goal (model 1), a reinforcement by time interaction (models 2 and 3), and a three-way goal by reinforcement by time interaction (model 3) were only significant for leisure walking. Time was the only significant independent factor contributing to the count model for reported transportation walking in both complete case and multiple imputation analyses. Prior studies that have shown this effect for goal type [12,13] have utilized objective data to capture MVPA. This study is the first to report differential effects of adaptive versus static goals for self-reported walking domains.

Although the WalkIT Arizona intervention did not target any specific domain of activity, differential effects were observed for transportation and leisure walking, and our hypotheses regarding similar intervention effects across leisure and transportation domains were not supported. However, these findings show that adaptive goals alone were similarly effective to static goals at increasing reported leisure and transportation activities over time. Immediate reinforcement alone or combined with goal setting were more effective than delayed reinforcement at increasing leisure walking at 12 months but not transportation-related walking. It is possible that immediate financial reinforcement is a stronger intervention stimulus to promote leisure walking than delayed reinforcement, but not a strong enough stimulus to overcome barriers (eg, low walkability) to adopting transportation walking. While we adjusted models for block-level walkability, we did not account for distance between participants' home and work, or walkability surrounding their workplace. These factors may further explain some of our findings, as this study occurred in a large, sprawling metropolitan area. These results support the premise that individual-level PA interventions are domain- and context specific and could be helpful in guiding further multilevel intervention refinement.

It is interesting to compare these findings with primary study outcomes that utilized accelerometer-measured MVPA. In primary analyses, a main effect of goal type was significant such that those with adaptive goals had a greater probability of initiating any MVPA bout minutes/day (versus none), whereas a main effect of reinforcement was significant such that immediate reinforcement was more successful at increasing total MVPA bout minutes/day. Interactions between goal type and reinforcement timing on MVPA bout minutes/day indicated that the group with adaptive goals combined with immediate reinforcement outperformed other groups, except for the group with static goals combined with immediate reinforcement. Analyses with self-reported data also indicated differences between hurdle and count models, although these findings were less robust. For hurdle models, a goal by time effect was observed only for leisure walking and was nonsignificant. Count models showed a main effect for goal type favoring static goals and a main effect for reinforcement timing favoring immediate reinforcement, as well as a goal by reinforcement by time interaction supporting static or adaptive goals combined with immediate reinforcement. These results were also only significant for leisure walking. There were no significant effects of intervention parameters for reported minutes of transportation walking; an observed reinforcement by time interaction was no longer significant in multiple imputation analyses.

Our registered secondary aim referred to self-reported PA as measured by the IPAQ but was not specific to walking or cycling. Although self-reported PA may be less accurate than objective data, the examination of domain-specific PA (eg, transportation versus leisure walking) provides a better conceptual alignment and allows for a more comprehensive understanding of participant behavior within a walking intervention, which may be useful in guiding intervention refinement. The WalkIT Arizona intervention maintained a broad focus on increasing ambulatory activities at a moderate intensity or greater. While there was little difference in the proportion of participants who endorsed any walking for leisure versus transportation, reported duration of walking increased significantly more for leisure walking than transportation walking at 12 months. These findings are consistent with prior studies indicating the prevalence and correlates of leisure walking differ from those of transportation walking [7,8,21,22]. Some studies have suggested that gender may impact the domain of activity, with women being more likely to report walking for leisure and men to report walking for transportation, despite no gender difference observed in total walking for any purpose [6]. The fact that 64.5% (330/512) of the current sample were women may have contributed to a greater increase in leisure walking compared with transportation walking. Other research has noted that leisure walking tends to increase with age [22], while attributes of the built environment (eg, walkability) have consistently been shown to impact transportation walking more than leisure walking. These latter points may suggest greater flexibility with leisure walking, suggesting this domain may be more receptive to change with these individual-level intervention components. As such, an mHealth intervention utilizing goals and reinforcement, without intervening on walking destinations and access to transit, was less effective at increasing transportation PA.

Study Limitations

The reported findings should be considered in the context of several limitations. As this study reported on secondary outcomes, the WalkIT Arizona trial was powered to detect effects using accelerometer data and not self-reported PA, which has greater variability. We elected to examine main effects of parameter by time using two-way interactions in models 1 and 2, as these models had greater power than model 3, which included a third interaction term. Lack of power may have also contributed to differences in findings between these analyses and the primary outcomes. Concerns have also been raised regarding the accuracy and sensitivity of self-reported PA, as correlations with objective measures tend to be low to moderate [11]. However, self-reported PA offered a unique opportunity to capture the context of participants walking activity, an aspect not readily accessible through objective measures in this study. Notably, self-reported PA was collected at only three time points that inquired about behavior over the previous week and may not necessarily have reflected more nuanced variability in PA over the course of the intervention. We used an intent-to-treat approach to preserve randomization and performed a sensitivity analysis comparing complete cases to multiple imputation. Results from complete case analysis and multiple imputation analyses were more consistent for leisure walking than for transportation walking.

Conclusion

This is the first study to report differential effects of adaptive versus static goals and immediate versus delayed reinforcement for self-reported walking by domain. Despite limited power for these secondary analyses, the results support the premise that individual-level PA interventions are domain- and context-specific. This information may be helpful in guiding intervention refinement and increasing generalizability to other populations.

Acknowledgments

This work was supported by the National Cancer Institute at the National Institutes of Health (R01CA198915). The funding agency was not involved in any aspect of this study or manuscript. The authors acknowledge the support of many undergraduate and graduate research assistants helping with data collection.

The authors also acknowledge JCH’s contributions, leading to her posthumous authorship.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Multiple imputation negative binomial hurdle model examining goal x time interaction (model 1) for leisure walking.
[DOCX File , 15 KB - formative_v4i12e19863_app1.docx]

Multimedia Appendix 2
Multiple imputation negative binomial hurdle model examining reinforcement x time interaction (model 2) for leisure walking.
[DOCX File , 14 KB - formative_v4i12e19863_app2.docx]

Multimedia Appendix 3
Multiple imputation negative binomial hurdle model examining goal x reinforcement x time interaction (model 3) for leisure walking.
[DOCX File , 15 KB - formative_v4i12e19863_app3.docx]

Multimedia Appendix 4
Multiple imputation negative binomial hurdle model examining goal x time Interaction (model 1) for transportation walking.
[DOCX File , 14 KB - formative_v4i12e19863_app4.docx]

Multimedia Appendix 5
Multiple imputation negative binomial hurdle model examining reinforcement x time interaction (model 2) for transportation walking.
[DOCX File , 15 KB - formative_v4i12e19863_app5.docx]

Multimedia Appendix 6
Multiple imputation negative binomial hurdle model examining goal x reinforcement x time interaction (model 3) for transportation walking.

[DOCX File , 15 KB - formative_v4i12e19863_app6.docx]

Multimedia Appendix 7
Negative binomial hurdle model examining goal x time interaction (model 1) for leisure walking.
[DOCX File , 14 KB - formative_v4i12e19863_app7.docx]

Multimedia Appendix 8
Negative binomial hurdle model examining reinforcement x time interaction (model 2) for leisure walking.
[DOCX File , 15 KB - formative_v4i12e19863_app8.docx]

Multimedia Appendix 9
Negative binomial hurdle model examining goal x time interaction (model 1) for transportation walking.
[DOCX File , 14 KB - formative_v4i12e19863_app9.docx]

Multimedia Appendix 10
Negative binomial hurdle model examining reinforcement x time interaction (model 2) for transportation walking.
[DOCX File , 14 KB - formative_v4i12e19863_app10.docx]

Multimedia Appendix 11
Negative binomial hurdle model examining goal x time interaction (model 1) for transportation biking.
[DOCX File , 14 KB - formative_v4i12e19863_app11.docx]

Multimedia Appendix 12
Negative binomial hurdle model examining reinforcement x time interaction (model 2) for transportation biking.
[DOCX File , 15 KB - formative_v4i12e19863_app12.docx]

Multimedia Appendix 13
CONSORT-EHEALTH checklist v1.6.1.
[PDF File (Adobe PDF File), 98 KB - formative_v4i12e19863_app13.pdf]

References

1. Rhodes RE, Janssen I, Bredin SSD, Warburton DER, Bauman A. Physical activity: Health impact, prevalence, correlates and interventions. Psychol Health 2017 Aug;32(8):942-975. [doi: 10.1080/08870446.2017.1325486] [Medline: 28554222]
2. Bize R, Johnson JA, Plotnikoff RC. Physical activity level and health-related quality of life in the general adult population: a systematic review. Prev Med 2007 Dec;45(6):401-415. [doi: 10.1016/j.ypmed.2007.07.017] [Medline: 17707498]
3. Brown DR, Carroll DD, Workman LM, Carlson SA, Brown DW. Physical activity and health-related quality of life: US adults with and without limitations. Qual Life Res 2014 Dec;23(10):2673-2680 [FREE Full text] [doi: 10.1007/s11136-014-0739-z] [Medline: 24952110]
4. Conn VS, Hafdahl AR, Mehr DR. Interventions to increase physical activity among healthy adults: meta-analysis of outcomes. Am J Public Health 2011 Apr;101(4):751-758 [FREE Full text] [doi: 10.2105/AJPH.2010.194381] [Medline: 21330590]
5. Adams MA, Hurley JC, Phillips CB, Todd M, Angadi SS, Berardi V, et al. Rationale, design, and baseline characteristics of WalkIT Arizona: A factorial randomized trial testing adaptive goals and financial reinforcement to increase walking across higher and lower walkable neighborhoods. Contemp Clin Trials 2019 Jun;81:87-101. [doi: 10.1016/j.cct.2019.05.001] [Medline: 31063868]
6. Pollard TM, Wagnild JM. Gender differences in walking (for leisure, transport and in total) across adult life: a systematic review. BMC Public Health 2017 Dec 20;17(1):341 [FREE Full text] [doi: 10.1186/s12889-017-4253-4] [Medline: 28427376]
7. Paul P, Carlson SA, Carroll DD, Berrigan D, Fulton JE. Walking for Transportation and Leisure Among U.S. Adults--National Health Interview Survey 2010. J Phys Act Health 2015 Jun 16;12 Suppl 1:S62-S69 [FREE Full text] [doi: 10.1123/jpah.2013-0519] [Medline: 25133651]
8. Ussery EN, Carlson SA, Whitfield GP, Watson KB, Berrigan D, Fulton JE. Walking for Transportation or Leisure Among U.S. Women and Men - National Health Interview Survey, 2005-2015. MMWR Morb Mortal Wkly Rep 2017 Jun 30;66(25):657-662 [FREE Full text] [doi: 10.15585/mmwr.mm6625a1] [Medline: 28662018]
9. Menai M, Charreire H, Feuillet T, Salze P, Weber C, Enaux C, et al. Walking and cycling for commuting, leisure and errands: relations with individual characteristics and leisure-time physical activity in a cross-sectional survey (the ACTI-Cités project). Int J Behav Nutr Phys Act 2015 Dec 09;12:150 [FREE Full text] [doi: 10.1186/s12966-015-0310-5] [Medline: 26646510]

10. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011;8:115 [FREE Full text] [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
11. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor GS, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. Int J Behav Nutr Phys Act 2008;5:56 [FREE Full text] [doi: 10.1186/1479-5868-5-56] [Medline: 18990237]
12. Adams MA, Sallis JF, Norman GJ, Hovell MF, Hekler EB, Perata E. An adaptive physical activity intervention for overweight adults: a randomized controlled trial. PLoS One 2013;8(12):e82901 [FREE Full text] [doi: 10.1371/journal.pone.0082901] [Medline: 24349392]
13. Adams MA, Hurley JC, Todd M, Bhuiyan N, Jarrett CL, Tucker WJ, et al. Adaptive goal setting and financial incentives: a 2 × 2 factorial randomized controlled trial to increase adults' physical activity. BMC Public Health 2017 Mar 29;17(1):286 [FREE Full text] [doi: 10.1186/s12889-017-4197-8] [Medline: 28356097]
14. Adams MA. A pedometer-based intervention to increase physical activity: Applying frequent, adaptive goals and a percentile schedule of reinforcement.: UC San Diego; 2009. URL: https://escholarship.org/uc/item/1gc062sh [accessed 2019-12-29]
15. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
16. Brooks M, Kristensen K, Benthem K, Magnusson A, Berg C, Nielsen A, et al. glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. The R Journal 2017;9(2):378. [doi: 10.32614/rj-2017-066]
17. van BS, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. Journal of Statistical Software 2011;45:1-67.
18. Bodner TE. What Improves with Increased Missing Data Imputations? Structural Equation Modeling: A Multidisciplinary Journal 2008 Oct 22;15(4):651-675. [doi: 10.1080/10705510802339072]
19. Fox J. Effect Displays in for Generalised Linear Models. J. Stat. Soft 2003;8(15). [doi: 10.18637/jss.v008.i15]
20. Fox J, Weisberg S. An R companion to applied regression, 3rd Edition. Thousand Oaks, CA: Sage Publications; 2019.
21. Carlson SA, Whitfield GP, Peterson EL, Ussery EN, Watson KB, Berrigan D, et al. Geographic and Urban-Rural Differences in Walking for Leisure and Transportation. Am J Prev Med 2018 Dec;55(6):887-895. [doi: 10.1016/j.amepre.2018.07.008] [Medline: 30344032]
22. Paul P, Carlson SA, Fulton JE. Walking and the Perception of Neighborhood Attributes Among U.S. Adults-2012. J Phys Act Health 2017 Jan;14(1):36-44. [doi: 10.1123/jpah.2015-0685] [Medline: 27775464]

Abbreviations

glmmTMB: generalized linear mixed models using template model builder
IPAQ: International Physical Activity Questionnaire
mHealth: mobile health
MVPA: moderate-to-vigorous physical activity
NBH: negative binomial hurdle
PA: physical activity
RR: risk ratio

Edited by G Eysenbach; submitted 05.05.20; peer-reviewed by D Bowen, J Wilroy; comments to author 15.06.20; revised version received 29.06.20; accepted 30.10.20; published 04.12.20.

Please cite as:
McEntee ML, Cantley A, Foreman E, Berardi V, Phillips CB, Hurley JC, Hovell MF, Hooker S, Adams MA
Effects of Goal Type and Reinforcement Type on Self-Reported Domain-Specific Walking Among Inactive Adults: 2×2 Factorial Randomized Controlled Trial
JMIR Form Res 2020;4(12):e19863
URL: https://formative.jmir.org/2020/12/e19863
doi:10.2196/19863
PMID:33275107

Original Paper

Internet-Delivered Tobacco Treatment for People Using Cannabis: A Randomized Trial in Two Australian Cannabis Clinics

Josephine Hindson[1], BA, MPsycClin; Tanya Hanstock[2], BA, DPsych; Adrian Dunlop[1], MBBS, PhD; Frances Kay-Lambkin[2], BSc, PhD

[1]Hunter New England Health District, Newcastle, Australia
[2]University of Newcastle, Callaghan, Australia

Corresponding Author:
Frances Kay-Lambkin, BSc, PhD
University of Newcastle
University Dr
Callaghan
Australia
Phone: 61 416204031
Email: frances.kaylambkin@newcastle.edu.au

Abstract

Background: Tobacco use is disproportionately higher in people who smoke cannabis than in the general population, increasing the severity of dependence for cannabis use, decreasing the likelihood of successful quit attempts for both cannabis and tobacco, and increasing the risk of relapse for both substances. Opportunities to address tobacco use in people using cannabis are being missed.

Objective: This study aims to investigate the feasibility of engaging tobacco smokers who were accessing treatment for cannabis, with a tobacco-focused *i*nternet-based *He*althy *L*ifestyle *P*rogram (iHeLP; 4 modules). It was hypothesized that iHeLP completion would be associated with decreases in tobacco use and improved quality of life (QoL) and psychological health. It was also hypothesized that iHeLP completion would be higher in those who additionally received telephone support. Given that iHeLP aimed to improve healthy lifestyle behaviors, it was also hypothesized that there would be reductions in cannabis use.

Methods: A total of 13 smokers seeking treatment for cannabis use were randomly allocated to iHeLP alone or iHeLP plus telephone support. Participants were engaged in iHeLP over 8 weeks and completed a 12-week follow-up assessment.

Results: Results from 10 participants who completed the follow-up indicated that the acceptability of iHeLP was high-very high in terms of general satisfaction, appropriateness of services, effectiveness, and met need. Additional telephone support increased modal module completion rates for iHeLP from 0 to 2 but did not provide any other significant advantages over iHeLP alone in terms of cannabis use, tobacco use, QoL, or psychological health. Participants in the iHeLP-alone condition (n=4) reported a mean reduction of 5.5 (SD 9.00) tobacco cigarettes per day between baseline and follow-up, with a concomitant mean reduction in expired carbon monoxide (CO) of 5.5 parts per million (ppm, SD 6.91). The iHeLP plus telephone support group (n=6) reported a mean reduction of 1.13 (SD 4.88) tobacco cigarettes per day and a mean reduction of 9.337 ppm of expired CO (SD 5.65). A urinalysis indicated that abstinence from cannabis was achieved by 2 participants in the iHeLP-alone group and three participants in the iHeLP plus telephone support group. Between baseline and follow-up assessments, iHeLP-alone participants reported a mean reduction in days of use of cannabis in the prior month of 6.17 days (SD 13.30). The average reduction in the number of days of cannabis use for the iHeLP plus telephone support group was also 6.17 days (SD 13.59).

Conclusions: Despite the small sample size, this study provides preliminary support for the use of internet-delivered, tobacco-focused interventions in tobacco smokers seeking treatment for cannabis use.

(JMIR Form Res 2020;4(12):e14344) doi:10.2196/14344

KEYWORDS

tobacco; cannabis; help-seeking behavior; internet-based intervention

Introduction

Cannabis is the most widely used illicit drug in the world [1], with up to 227 million people worldwide reporting the use of cannabis [2]. Australia has one of the highest rates of cannabis use in the world, with 10% of Australians older than 14 years reporting the use of cannabis in the previous 12 months [3].

Tobacco is another commonly used substance worldwide. It is estimated that in 2015, 21% of the adult population smoked tobacco in that year [4]. In Australia, 12.8% of people aged older than 14 years smoke tobacco, a reduction from 15.1% in 2010 [3]. Despite this, a number of subgroups in the Australian population still report relatively high rates of tobacco use. One such subgroup is the substance-using treatment-seeking population, with concurrent tobacco use estimates ranging from 74% to 100% [5].

There are generally high rates of tobacco smoking in people who smoke cannabis, including both concurrent and simultaneous uses. It is estimated that 50% of adults with cannabis use disorders are currently smoking tobacco, increasing a number of risks. Rates of dependence on cannabis when smoking tobacco are higher than when smoking cannabis only [6-8]. Combined cannabis and tobacco use also increases the dependence on tobacco [9], decreases successful quit attempts for both cannabis and tobacco [10-12], and increases the chance of relapse for both substances [13]. Combined cannabis and tobacco use also increases the risk of respiratory issues compared with smoking cannabis alone [8,14].

Only a small percentage of cannabis users access treatment [15,16]. Barriers to accessing treatment include the perception that cannabis use is not problematic enough to warrant treatment; concerns that quitting may exacerbate mental health and sleep- or pain-related issues [17,18]; reluctance to engage in traditional alcohol/other drug treatment services; notion that there are no specific treatment services for cannabis [19]; and avoidance of the stigma associated with being a *drug user* [18]. In contrast, cannabis users perceive tobacco as toxic and addictive [20] and rank it more harmful than alcohol or cannabis [21]. The majority of cannabis users report a clear intent to stop tobacco use [20], providing an opportunity for engaging them in discussions about their current lifestyle.

A large body of evidence exists for effective interventions that reduce and cease tobacco use, including pharmacotherapy and psychosocial therapies. For example, in a review of tobacco treatment research, Prochaska and Benowitz [22] reported that psychotherapies, delivered in both individual and group counseling settings, demonstrated effectiveness for tobacco cessation compared with self-help treatment. The appeal of psychosocial treatment is that strategies can be applied to both cannabis use and tobacco use simultaneously, providing a promising approach to addressing the co-occurrence of these behaviors. For example, Hoffman et al [23] conducted a review of 106 representative meta-analyses of cognitive behavioral therapy (CBT) for numerous disorders, including substance dependence. They found that similar CBT strategies were effective in the treatment of cannabis and nicotine dependence. In the only study of its kind to date, Lee et al [24] investigated concurrent cannabis and tobacco computer-delivered treatment. They compared 32 cannabis-dependent adults who received treatment focused on both cannabis and tobacco to 54 participants, from a previous study, who had received either therapist- (n=28) or computer-delivered (n=26) treatment focusing only on cannabis. The treatment involved CBT, contingency management, and nicotine replacement therapy (NRT). The results indicated that the majority of participants were interested in receiving a tobacco intervention and that concurrent treatment for tobacco use did not compromise cannabis use outcomes. The tobacco intervention did not yield a high rate of abstinence at the end of treatment; however, it did motivate more than half the smokers to at least attempt to quit using tobacco.

A number of reviews have assessed the efficacy of internet-based treatment for tobacco use. Rooke et al [25] conducted a meta-analysis of 34 studies of 10,632 participants, and found that computer-delivered interventions significantly reduced tobacco and alcohol use overall, and that the magnitude of this change was comparable to individual counseling provided for tobacco use. In a Cochrane review of tobacco cessation programs, Civljak et al [26] analyzed 28 trials with over 45,000 participants involving any internet intervention. The comparison groups could include both internet and noninternet interventions. The results revealed that internet interventions that were tailored to the individual and interactive were more likely to aid in smoking cessation at 6 months posttreatment than those that did not. The authors suggested that younger people and women may be specifically interested in internet treatment; however, this requires more research.

A large study of adult tobacco smokers in the United States compared an internet-only intervention, where access to the internet-based intervention was given for 6 months, with an internet plus telephone support intervention for which the internet intervention was accompanied by five phone calls. The internet plus telephone support group reported significantly better outcomes up to 12 months of follow-up, suggesting that the addition of telephone support helped participants quit at an earlier time point that in turn had related health benefits [27].

Our team has recently developed an *i*nternet-delivered *He*althy *L*ifestyle *P*rogram (iHeLP) that targets tobacco use, in addition to other lifestyle factors of physical activity and healthy eating. The internet is increasingly becoming an option to deliver psychological treatment, especially CBT and motivational interviewing (MI), which are psychosocial treatment strategies with the strongest evidence for reducing both cannabis and tobacco use [28]. This study aimed to assess the feasibility of using iHeLP among tobacco smokers seeking treatment for cannabis use disorder by engaging with Specialist Cannabis Clinics in New South Wales (NSW) Health, Australia. Another aim of this study was to examine the role of telephone support in encouraging the uptake of iHeLP and to determine whether the use of the iHeLP program is associated with reductions in both cannabis and tobacco use.

Methods

Ethics

This study was designed as a feasibility study. It was approved by the Hunter New England Research Ethics and Governance Unit (HNEHREC Ref No: 15/05/20/4.06).

Participants

Participants included 13 adults (7/13, 54% men; 6/13, 46% women) recruited from 2 Cannabis Clinics, one based within the Hunter New England Local Health District (HNELHD) and the other in the Central Coast Local Health District in NSW, Australia. Participants were aged between 20 and 50 years (mean age 34.15, SD 8.48 years). The eligibility criteria were patients older than 18 years, who had access to the internet via a personal device (computer, laptop, or mobile phone), who smoked at least five tobacco cigarettes per day, and who were not experiencing active psychosis at the time of recruitment. Intention to quit tobacco or cannabis was not a requirement for study participation. Participants could be at any stage of treatment for their cannabis use disorder, as long as they were currently attending appointments at the Specialist Cannabis Clinics [29].

Procedure

Treating clinicians at both services initially screened patients for eligibility for the study. Those who met the eligibility criteria and were interested in participating were contacted within a week by the researcher (JH) to schedule an initial appointment and answer any questions regarding the study.

After obtaining informed consent, participants completed a 1-hour baseline face-to-face assessment. They were then randomized into one of the following 2 conditions: access to iHeLP (an internet-based healthy lifestyle program) with no support or access to iHeLP with telephone support. These interventions are described later. Research clinicians were blinded to treatment allocation until the conclusion of the baseline assessment was obtained.

Participants were provided with individual log-in details following randomization and had access to the iHeLP program for a period of 8 weeks. The participants were asked to complete the tobacco section of the program at a minimum and could choose to complete as many other sections as relevant. The participants were given Aus $20 (US $13.75) gift card for participation in the baseline assessment. Access to iHeLP and phone-based services were provided free of charge.

A follow-up face-to-face assessment was conducted at 12 weeks postbaseline. An Aus $20 (US $13.75) gift card was also offered to the participants who attended the follow-up appointment.

Treatment

Participants were randomized to iHeLP alone versus iHeLP plus telephone support. iHeLP targets tobacco use (2 modules), diet quality (1 module), and physical activity (including sedentary behavior; 1 module). Internet sessions are based on MI techniques and CBT strategies, as recommended for maximizing compliance. Within iHeLP, participants were asked to at least complete the tobacco modules as a starting point and then work their way through the remaining modules (physical activity and diet quality) in a sequence of their own, choosing over the 8-week treatment period. Automated emails were programmed to prompt module completion after a period of nonactivity. Participants accessed the treatment package from their home computer (or a preferred internet access port). The iHeLP intervention was not tailored for people using cannabis who also smoke tobacco, rather it incorporates standard tobacco cessation strategies and approaches that have demonstrated benefits in the general population and in people with mental health problems [30]. Please see Multimedia Appendix 1 for screenshots of the iHeLP program.

Telephone Support

Telephone support was provided to half of the study participants and involved 8 weekly phone calls on a set day and time to support access to and completion of iHeLP. Each call was limited to a duration of 10 min and was carried out manually. The content of such a phone call included reviewing the previous week, reinforcing the importance of making healthy lifestyle decisions, increasing adherence to the iHeLP modules, and discussing tobacco cessation attempts, including the use of NRT to manage tobacco cravings and withdrawal.

Measures

Demographic information was collected at the baseline assessment. This included age, gender, education level, indigenous status, living arrangements, employment status, and source of income. The frequency of internet access was also measured. In addition, the following measurements were taken.

Depression Anxiety Stress Scale 21

The Depression Anxiety Stress Scale 21 (DASS 21) [31] was used at baseline and follow-up to assess current symptoms of depression (eg, "I felt I had nothing to look forward to"), anxiety (eg, "I felt I was close to panic"), and stress (eg, "?"). The validity and reliability of DASS 21 is sound, and a Cronbach of .90 was reported in existing literature for the depression subscale, and a Cronbach of .82 was reported for the anxiety subscale [32].

The Fagerstrom Test for Nicotine Dependence

The Fagerstrom Test for Nicotine Dependence (FTND) [33] was used for both assessments. FTND is a six-item scale, with scores ranging from 0 to 10, with higher scores indicating greater dependence. The reliability of FTND is good, with an acceptable Cronbach of .72 and a test-retest correlation of 0.82 [34].

Eurohis-QoL 8 Item

The Eurohis-QoL 8-item index [35] is a shortened version of the World Health Organization Quality of Life Instrument–Abbreviated Version and was used in both assessments. It measures the psychological, physical, social, and environmental domains on a 5-point Likert scale (0-5), with a total score between 8 and 40. In previous research, this instrument was found to have an acceptable Cronbach of .78 [35].

Australian Treatment Outcome Profile

The Australian Treatment Outcome Profile (ATOP) [36] is a standardized self-report measure used within the HNELHD Drug and Alcohol Clinical Services in NSW. It measures drug and alcohol use in the prior 4 weeks in terms of days used (frequency) and the average amount used (quantity). Self-reported frequency and quantity of cannabis and tobacco use in the prior 4 weeks were measured for all participants at baseline and follow-up. It should be noted that for this study, a *cone* (ie, *standard unit*) was used to estimate the quantity of cannabis use smoked on a use occasion. Cannabis can be shaped into a *cone* and smoked using a pipe or *bong*. Using this method, typically, one *cone* of cannabis is used per bong/pipe. Cannabis can also be rolled into a cigarette shape (a *joint*) and smoked like tobacco. For this study, it was estimated that the equivalent of 3 cones of cannabis was used per *joint*. Participants were asked to estimate the number of *cones* of cannabis they used per use occasion across both *bongs*/*pipes* and *joints*. A *use occasion* was defined as a period of using cannabis from the time a person started smoking until the time they finished smoking. A person could report more than one *use occasion* in a day. ATOP also contains questions regarding days worked and studied in the previous 4 weeks. It has yes/no questions regarding housing situations, children in care, violence, and arrests in the past 4 weeks. It uses a 10-point Likert scale (0-10) to measure the psychological health, physical health, and overall quality of life (QoL). The scale has an acceptable concurrent validity and interrater reliability in the established literature [36].

Credibility/Expectancy Questionnaire

The Credibility/Expectancy Questionnaire [37] was administered at baseline only to determine the extent to which participants perceived internet-based treatments as a credible form of treatment for substance use and to ask them to estimate the expected outcomes associated with the completed treatment for substance use internet-based. Responses were rated on a 9-point Likert scale (eg, 1=not at all logical/confident/useful to 9=very logical/confident/useful), and estimated improvements were nominated from 0% to 100%. The psychometric properties have been previously reported, and it was found that the expectancy factor had a standardized alpha of between .79 and .90, the credibility factor had a Cronbach between .81 and .86, and the whole scale had a standardized alpha of between .84 and .85 [37].

Client Satisfaction Questionnaire

The Client Service Questionnaire (CSQ) [38] was administered at follow-up only. This is an eight-item measure that includes both numerical questions and the possibility of providing feedback on the treatment received. All response options are based on a four-point scale. The psychometric properties have been previously assessed [38] with high internal consistency (Cronbach =.93) and construct validity (ranging from 0.6 to 0.8) reported.

Expired Carbon Monoxide

The Bedfont piCO+Smokerlyzer (Bedfont Scientific) was used to assess expired carbon monoxide (CO) at baseline and follow-up assessments and is an objective measure of the current tobacco use status. Previous research has shown that expired CO is a reliable measure for validating self-reported tobacco use [39]. A reading of 10 parts per million (ppm) or higher is indicative of continued tobacco use.

Urinalysis

A urinalysis was used to confirm self-reported abstinence from illicit substance use at both assessments. DipScan measures amphetamines, benzodiazepines, cocaine metabolites, methamphetamines, opiates, and cannabis metabolites at cutoff levels defined by the Australian Standard AS/NZS 4380:2008.

iHeLP Website Analytics

The iHeLP program tracked the participant's progress through each module to obtain an objective measure of the extent of engagement (ie, program completion) associated with the program.

Statistical Analysis

The main focus of the statistical analysis was descriptive and was conducted using SPSS Statistics Software (version 22.0, IBM Corporation). To address engagement and completion of iHeLP, module completion rates were calculated for each condition (iHeLP vs iHeLP plus telephone support), based on the number of modules completed (not attempted or started=0).

To address changes in cannabis use (ATOP), tobacco (ATOP), QoL scores, DASS 21, and DASS 21 anxiety scores, paired sample *t* tests were used to analyze baseline to follow-up changes in each variable for the sample as a whole. Statistical tests for differential changes in these variables as a function of treatment allocation were not performed due to the small sample size. Instead, descriptive statistics were reported for each of these variables for baseline and follow-up completers according to treatment allocation.

Results

Demographics

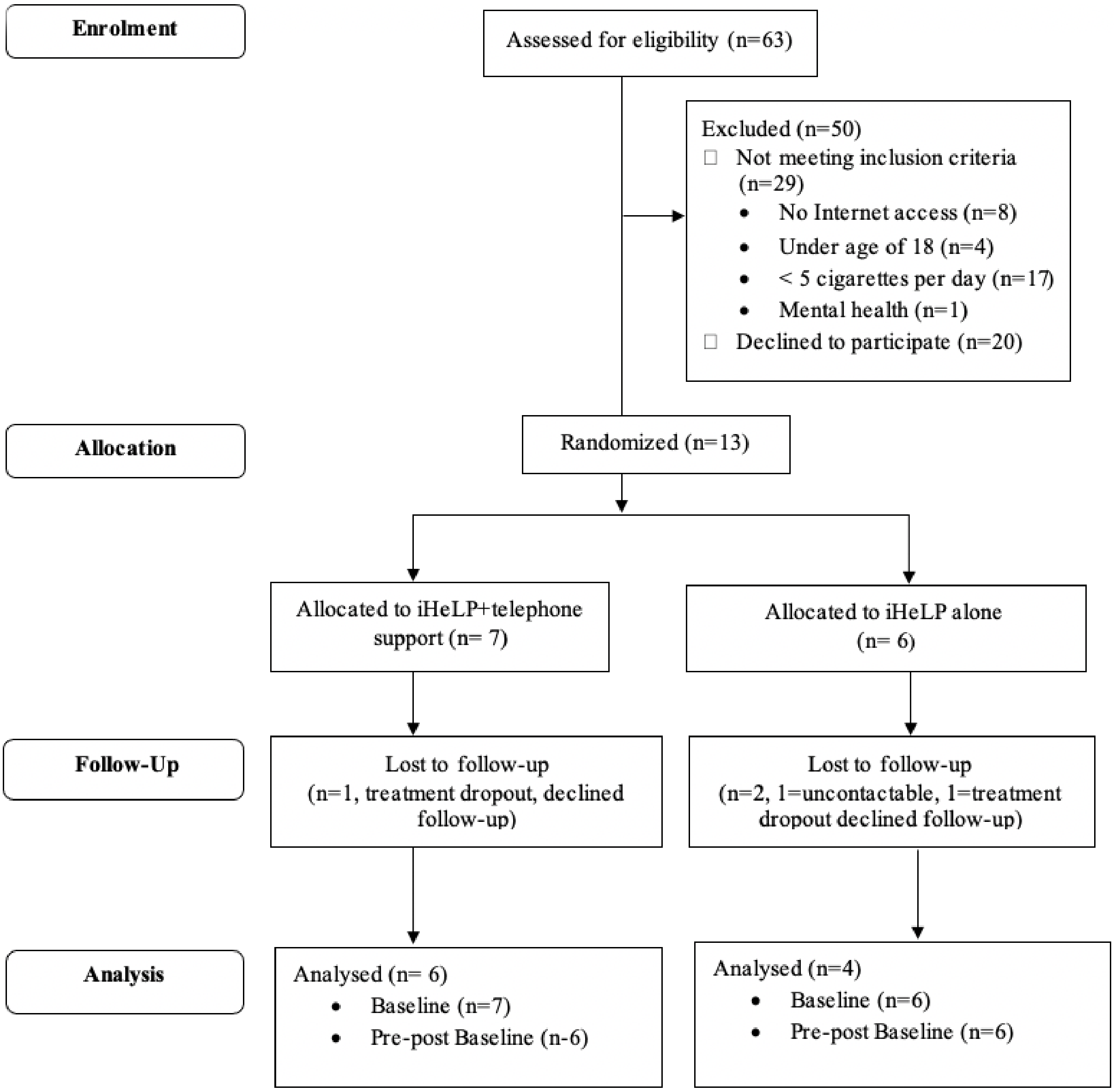
A total of 13 participants were enrolled in the program over a period of 3 months and completed the baseline assessment. The mean age of the sample was 34.15 years (SD 8.484 years; range 20-50 years), and there were 7 men and 6 women. As indicated in Table 1, 3 (3/13, 23%) participants were identified as being Aboriginal or Torres Strait Islander. The majority of participants were unemployed, 2 had full-time jobs, and 2 had part-time jobs (Table 1). The mean number of days worked in the last 4 weeks was 6.15 (SD 8.69 days; range 0-20 days), and the average number of days studied in the previous 4 weeks was 2.00 (SD 5.65 days; range 0-20 days). At baseline, internet access ranged from several times a day, every day, and several times a week to once a month or less. See Table 1 for demographic details of the participants.

Table 1. Baseline participant characteristics (n=13).

Demographic characteristics	Values
Age (years), mean (SD)	34.15 (8.484)
Gender (male), n	7
Aboriginal or Torres Strait Islander, n	3
Highest level of education, n	
Technical College (technical diploma/certificate)	9
High school	2
Did not complete high school	2
Employment status, n	
Full-time work	2
Part-time work	2
Household duties	1
Unemployed	8
Source of income, n	
Wage/salary	3
Own business	1
Government benefits	7
Family payments (government)	2
Internet access, n	
Several times a day	6
Every day	2
Several times a week	4
Once a month or less	1

Of the 13 people recruited to the study, 6 were randomized to iHeLP alone, while 7 were randomized to the iHeLP plus telephone support group. Of those who entered the study, 4 from the iHeLP alone and 6 from the iHeLP plus telephone support group completed the study. The recruitment and study enrollment details are shown in Figure 1.

Figure 1. Flow of participants through the study.



Substance Use

At baseline, the mean value for FTND was 5.15 (SD 2.30; range 0-9; medium dependence), the average expired CO was 25.00 (SD 12.65; range 4-53), and the average tobacco use for the past month was 16.77 (SD 7.88; range 5-30) cigarettes per day. The average number of days of cannabis use in the past month was 17.54 (SD 13.14 days), and the average amount of cannabis consumed in each use occasion was 11.08 (SD 11.94 cones). At baseline, the average depression score on DASS-21 was 12.92 (SD 10.60; *mild*), and the average DASS-21 anxiety score was 12.46 (SD 9.28; *moderate*).

Participants' Engagement and Acceptance of iHeLP

Participants in the iHeLP-alone condition attended a mean of 6.75 sessions (SD 5.25) at their respective cannabis clinics. Participants in the iHeLP plus telephone support condition attended a mean of 6.83 sessions (SD 3.37). The mean number of modules completed by the iHeLP program for the iHeLP-alone group was 1.17 modules (SD 1.33; range 0-3) and 1.85 modules (SD 1.46; range 0-4) for the iHeLP plus telephone support group. Participants in the iHeLP plus telephone support group received an average of 6.83 (SD 0.98) supportive phone calls over the 8-week treatment period, lasting for an average of 5.50 (SD 1.76) min each.

People in the iHeLP-alone condition reported a modal session completion of 0 (median 1.00; SD 1.17). In the iHeLP plus telephone support group, 2 modal modules were completed (median 2.00; mean 1.85, SD 1.12; Table 2).

Table 2. Frequency of participants completing each module of the iHeLP program as a function of study condition (iHeLP alone versus iHeLP enhanced).

	iHeLP alone	iHeLP enhanced
Module number completed		
0	3	2
1	0	0
2	2	3
3	1	1
4	0	1

Overall, the participants reported being satisfied with iHeLP. As can be seen in Table 3, all mean satisfaction scores on the CSQ were between 3 and 4 for both the iHeLP-alone and iHeLP plus telephone support conditions. This indicates that there was a high to very high satisfaction with the program. Average *general satisfaction* was rated by iHeLP-alone participants as 4.00 (*very satisfied*) and iHeLP plus telephone support as 3.83 (SD 0.41; *mostly satisfied to very satisfied*; n=6). Average *appropriateness of services* was rated by iHeLP-alone participants as 4.00 (*highly appropriate*) and as 3.83 for the iHeLP plus telephone support group (SD 0.41; *generally appropriate to highly appropriate;* n=6). Table 3 displays the satisfaction ratings across all domains measured according to the treatment allocation.

Table 3. Mean ratings on the Client Service Questionnaire provided by participants, as a function of study condition (iHeLP alone versus iHeLP enhanced)[a]

Program	Promptness	Focused on helping versus procedures	Quality of service	Service aligned with wants	Service met needs	Recommend service to a friend?	Satisfaction with help received	Services helped deal with problems more effectively	Overall	Would you return?	Service appropriate	Problems now?	Kept info private?
iHeLP alone	4.00	4.00	4.00	4.00	3.25	4.00	4.00	3.50	3.75	4.00	3.75	3.50	4.00
iHeLP enhanced	3.67	3.50	3.83	3.50	3.33	3.83	3.83	3.50	3.83	3.83	3.83	3.33	4.00

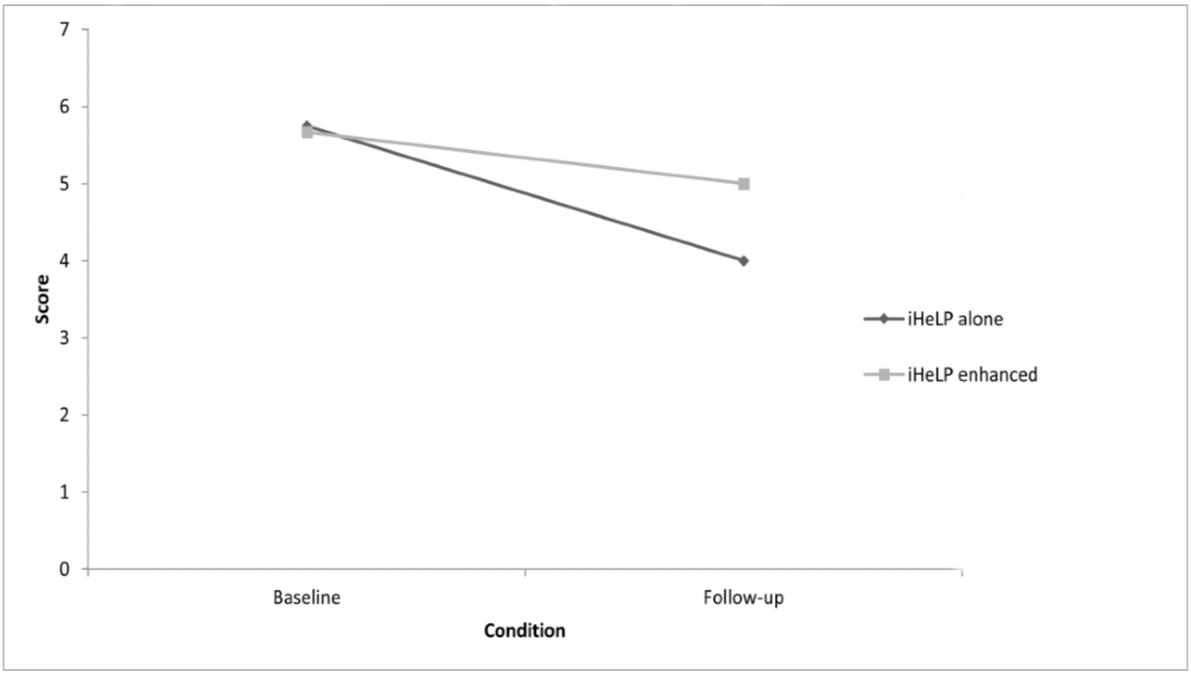
[a] response categories range from 1 to 4, with higher scores indicating higher satisfaction.

Changes in Tobacco Use Over Time

Paired sample *t* tests for the overall sample (n=10) revealed that FTND scores decreased from 5.7 (SD 1.95) at baseline to 4.6 (SD 2.55) at follow-up. This difference was statistically significant (t_9=2.283; *P*=.048).

The 4 participants in the iHeLP-alone group reported an FTND baseline mean of 5.750 (SD 1.03), and at follow-up, this had decreased to a mean of 4.00 (SD 1.32). The iHeLP plus telephone support group reported an FTND baseline mean of 5.67 (SD 0.84), which decreased to 5.00 (SD 1.08; Figure 2).

Figure 2. Changes in scores on the Fagerstrom test for nicotine dependence between baseline and follow-up assessment, according to treatment allocation.

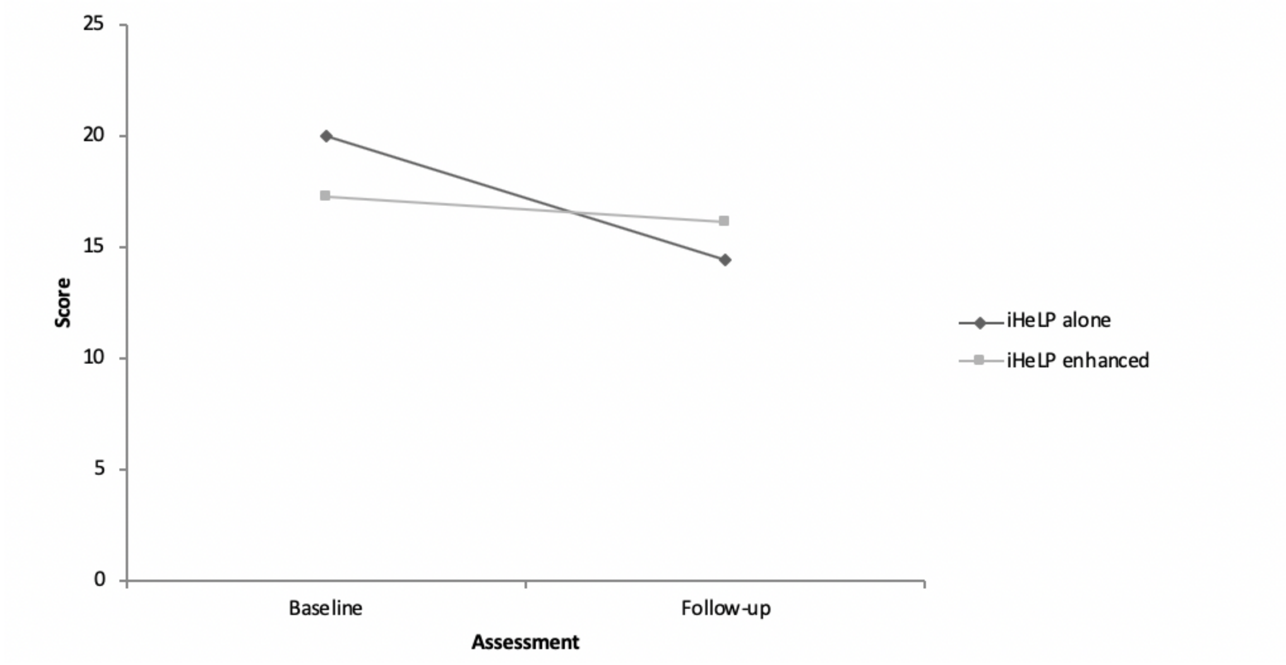


For the 10 participants completing follow-up assessments, the number of tobacco cigarettes smoked per day decreased from a mean of 18.4 (SD 7.53) at baseline to 15.5 (SD 6.40) at follow-up. Participants also reported a reduction in days of tobacco use in the month before assessment, decreasing from a mean at baseline of 28 (SD 0.00) days to 26.40 (SD 5.06) days at 12-week follow-up. Objective measures of expired CO supported these self-reported reductions in tobacco, with the mean expired CO for the 10 participants decreasing from 25.60 ppm (SD 14.546) at baseline to 17.80 ppm (SD 13.054) at follow-up.

The 4 iHeLP-alone participants who completed follow-up assessments reported a baseline mean expired CO reading of 22.50 ppm (SD 7.58), which reduced to 17.00 ppm at follow-up (SD 6.91). The iHeLP plus telephone support group reduced from 27.667 ppm expired CO (SD 6.19) at baseline to 18.33 ppm (SD 5.65) at follow-up.

At baseline, the 4 iHeLP-alone participants who completed both assessments reported a mean daily cigarette consumption in the prior month of 20.00 (SD 12.25). This reduced to a mean of 14.50 (SD 9.00) at follow-up assessment. The 6 iHeLP plus telephone support participants who completed both assessments reported a reduction in past month cigarettes per day from a mean of 17.33 (SD 2.94) at baseline to a mean of 16.17 (SD 4.88) at follow-up (Figure 3).

Figure 3. Changes in self-reported tobacco use (cigarettes per day) between baseline and follow-up assessments, according to treatment allocation.



Changes in Cannabis Use Over Time

The 10 participants providing follow-up data reported a median of 10.00 cannabis cones (mean 11.08, SD 11.94) per day at baseline in the month before the assessment, and this reduced to a median of 3 (mean 7.0, SD 11.56) per day at follow-up. Participants also reported a reduction in cannabis use days in the month before the assessment with a median at baseline of 28 days (mean 17.54, SD 13.14), which reduced to a median of 4 days (mean 10.90, SD 12.88). The urinalysis confirmed self-reported abstinence from cannabis use on all occasions. Of the 10 follow-up participants, 8 provided follow-up urine samples for analysis, of which 5 self-reported abstinence from cannabis, and the urinalysis was negative. Of the remaining patients, 3 self-reported continued use of cannabis, and all returned a positive urinalysis, indicating continued use. The 2 patients who declined to provide a urine sample self-reported continued use of cannabis at follow-up.

For the 4 participants in iHeLP alone who provided follow-up data, cannabis use in the prior month remained reasonably constant, with a baseline mean of 5 cones per day (mean 5.25, SD 6.58) and a mean of 5.25 cones per day (SD 6.08) at follow-up. Two participants in the iHeLP-alone group returned a negative urine sample, indicating abstinence from cannabis. The iHeLP plus telephone support group (n=6) reported reductions from a mean of 11 cones per day (mean 11.33, SD 6.58) at baseline to a mean of 8 cones per day (mean 8.18, SD 4.96) at follow-up. A negative urinalysis was returned for 3 of these participants, indicating abstinence from cannabis use.

The number of days cannabis was used in the iHeLP-alone group (n=4) reduced from a median of 16 days at baseline (mean 14.67, SD 13.125) to 3 days (mean 8.50, SD 13.30) at follow-up. The iHeLP plus telephone support group (n=6) reported a reduction in days of cannabis use from a median of 28 (mean 18.67, SD 14.46) days at baseline to a median of 10 (mean 12.50, SD 13.59) days at follow-up.

Participant Quality of Life Over Time

The participants who provided follow-up data (n=10) reported a mean of 22.6 (SD 4.6) at baseline for their current QoL, and this significantly increased to 27.40 (SD 4.12) at follow-up. For the 4 participants in iHeLP alone, QoL ratings increased from 20.50 (SD 2.24) at baseline to 29.50 (SD 1.96) at follow-up. The iHeLP plus telephone support group also reported an increase in QoL scores from 24.00 (SD 1.83) at baseline to 26.00 (SD 1.6) at follow-up.

DASS-21 Scores Over Time

The 10 participants providing follow-up data reported mean DASS 21 scores of 12.60 (median 12.00, SD 11.74) at baseline, and this decreased to 7.40 (median 8.00, SD 6.33) at follow-up. For the 4 iHeLP-alone participants, the mean baseline DASS-21 depression scores were 21.00 (median 20, SD 4.91) and 10.00 (median 11, SD 3.14) at follow-up. The iHeLP plus telephone support group reported a mean DASS-21 depression score of 7.00 (median 4, SD 4.01) at baseline, and this decreased to 5.67 (median 6, SD 2.56) at follow-up.

For DASS-21–anxiety, the 10 participants providing follow-up assessments reported a mean baseline score of 11.40 (median 10, SD 10.16), which decreased to 5.60 (median 5, SD 4.40) at follow-up (t_9=1.99; *P*=.08). Within the iHeLP-alone group (n=4), the mean baseline DASS-21–anxiety scores were 18.00 (median 19, SD 4.47), which decreased to a mean of 5.50 (median 4, SD 2.33) at follow-up. The iHeLP plus telephone support group (n=6) reported at baseline a DASS-21–anxiety score of 7.00 (median 8, SD 3.65), which decreased at follow-up to 5.67 (median 6, SD 1.91).

Discussion

Principal Findings

This study is the first of its kind to explore the use of an internet lifestyle intervention in tobacco smokers seeking treatment for cannabis use disorder. It has shown that participants offered the program would engage with the program even if they are not seeking treatment for lifestyle factors such as tobacco use. It has also been shown to have benefits for tobacco use and QoL.

The primary aim of this study is to explore the feasibility of iHeLP among tobacco smokers who were accessing treatment for cannabis use. The secondary aim was to evaluate if additional telephone support provided by a psychologist increased the engagement and completion rates of iHeLP in this sample. The clinical outcomes of interest were cannabis and tobacco use as well as QoL and psychological health, hypothesizing that cannabis and tobacco use would decrease while QoL and psychological health would improve over the course of the study.

The main finding of this study was that iHeLP was highly acceptable as seen by the high levels of satisfaction reported by the participants. The study also found that the participants did access iHeLP, with half of the total number of participants completing at least one module, which comprised the tobacco-related component of the internet-based program. This is significant, given that the participants were accessing the cannabis treatment clinic for their cannabis use specifically and were not seeking treatment for their tobacco use. This provides some initial support for the value of offering cannabis users an opportunistic intervention for tobacco use when they present for treatment. Although half of the study participants reported use of the internet every day (and for some, several times a day), 5 out of 13 participants were less frequent users of the internet at the beginning of the study. Internet access at baseline was not associated with the uptake or use of the iHeLP program, demonstrating the potential of internet-delivered interventions even with infrequent users of this technology.

Comparison With Prior Work

When participants had the addition of telephone support from a psychologist to iHeLP (iHeLP plus telephone support), there were no apparent increases in the rates of module completion. However, when looking at the modal number of sessions completed in each treatment group, those in the iHeLP plus telephone support group were most likely to complete two modules, with the mode for the iHeLP-alone group being 0. Thus, there is some suggestion here that therapist assistance (even as little as 5 min per week as provided in this study) did seem to coincide with the greater uptake with iHeLP. This is consistent with the previous research in internet-based interventions, which indicates that the addition of therapist support is associated with higher quit rates and more active involvement in the program [40,41].

Reductions in both tobacco dependence scores and the amount of tobacco smoked per day were reported by participants across both iHeLP conditions. In the total sample providing follow-up data (n=10), this reduction corresponded to dependence scores of *medium* at baseline to *low* dependence at 12-week follow-up, a clinically significant change in levels of tobacco dependence. Previous research has indicated that internet interventions can be as effective as counseling in producing a reduction in tobacco use, and those that are interactive and tailored to the individual produce better results [42,43]. This study adds to this existing research by demonstrating that internet interventions for tobacco use are feasible and can be considered as another treatment option for people accessing treatment for cannabis.

Cannabis use did decrease between baseline and follow-up, in terms of actual cannabis consumption and days of use. This is expected, given that participants were actively engaged in treatment for their cannabis use throughout the study, and had attended an average of 6 sessions with the specialist cannabis clinic. These cannabis reduction rates are consistent with other studies and highlight the difficulties clinicians and individuals face in encouraging changes in cannabis use [44]. These difficulties include retention in treatment and engagement [45]. Together, this body of evidence suggests that a continued and major effort is required to better understand the nature of cannabis use and how to encourage change in cannabis use in those seeking treatment.

QoL ratings improved for the sample providing follow-up data, indicating the potentially broader benefits of a lifestyle intervention in tobacco smokers seeking treatment for cannabis use disorders. Important clinical improvements were also reported by the participants in terms of their depression and anxiety over time. For DASS-21–depression, the iHeLP-alone group in particular were in the *severe* range at baseline, and this dropped to a *mild* rating at follow-up. This pattern was also observed in the iHeLP-alone group for DASS-21–anxiety, who decreased from a *severe* rating at baseline to a *mild* rating at follow-up. As there is a large difference between baseline scores for depression in the 2 groups in this study, future studies should consider using depression as a stratification variable or control for depression in the analysis.

Limitations

This study had a number of methodological issues. The major limitation of this study was the small sample size; therefore, the results need to be interpreted with caution. Recruitment proved difficult for a number of reasons. A large percentage of clients were unable to access the internet, with 26.7% of eligible clients reporting no access. This is in line with previous concerns regarding the limitations of internet interventions for this particular patient population [46]. Although catering to the proportion of cannabis users without internet access is critically important, internet interventions can be accessed by a large proportion of cannabis users. It should also be noted that at this stage, the iHeLP program was not accessible on mobile devices and this may have prevented some potential and willing participants from engaging with the study. Of the eligible sample with internet access in this study, despite reporting significant social disadvantage, 82% (10/13) reported at least weekly use of the internet. This provides at least partial support for the further exploration of internet-based treatment programs in people seeking treatment for cannabis use. Over half (58.6%) of individuals screened for this study did not meet the criteria for tobacco use. This percentage was much higher than expected from previous research [47]. This may be indicative of the knowledge that participants in studies comparing the perceptions of cannabis and tobacco have found that tobacco is perceived as harmful and toxic, and accordingly, cannabis users have already modified their tobacco use [21,48].

It should be noted that the researcher of this study (JH) was also a treating clinician for a percentage of participants (30%) at one of the recruitment sites for the study. While steps were taken, in line with ethics approvals, to reduce the perceptions of the coercion, these results may be subject to bias and should be interpreted with caution.

Finally, this study was conducted with cannabis users who already sought treatment for their cannabis use. It is not known how an internet intervention for tobacco use might appeal to or engage in nontreatment seekers for cannabis use and similarly whether internet access exists for cannabis users not accessing treatment.

Clinical Implications

Notwithstanding the significant limitations of this study, there is some initial support for the clinical value of integrating internet tobacco programs into substance use services, particularly in encouraging tobacco reduction in cannabis users. Future studies could investigate the efficacy of NRT in conjunction with internet programs to evaluate the efficacy of these treatment options in encouraging abstinence from tobacco in people using cannabis. It is also important to investigate the role of additional support alongside the internet-based treatment to increase adherence to and completion of similar internet-based programs for cannabis users.

As shown in previous studies, those accessing treatment for cannabis usually present with a more complex profile than those cannabis users who do not present for treatment [49]. Thus, this study should be replicated with a non–treatment-seeking sample of cannabis users who are concerned about their tobacco consumption. It is well documented that tobacco is perceived as more harmful than cannabis by cannabis smokers [20,50] and thus may be a more acceptable treatment option for cannabis users not accessing treatment. Early success with tobacco cessation, particularly via a broader *healthy lifestyle* approach, may act as a gateway to considering changes in other *lifestyle* areas (eg, cannabis use), thus encouraging treatment seeking.

Conclusions

The results from this study support the acceptability of an internet-based iHeLP (lifestyle-focused) program for tobacco smokers seeking treatment for cannabis use disorders. The addition of telephone support by a psychologist did not produce any significant advantages, apart from a higher modal completion rate in those with additional therapist support. There was a statistically significant reduction in nicotine dependence over the study period, which has significant public health and physical health implications. The results of this study are encouraging regarding the acceptability of the iHeLP program on a subgroup of the population who are both difficult to engage and often excluded from research studies.

Acknowledgments

The authors would like to acknowledge and thank the participants in the study who gave their time and perspectives willingly and openly. This study was funded by a small research grant awarded to the team from the University of Newcastle, Psychology Department, as a contribution toward the completion of study for JH. FL conceived the study. JH, FL, and TH contributed to the study design and implementation. FL designed and developed the iHeLP intervention and provided access to study participants in-kind. AD provided input into the study design and provided approval for staff and clients of the service to participate in the project. JH, FL, and TH worked together to produce the first draft of the manuscript. AD provided input into subsequent drafts. All authors contributed to the final version of the manuscript and approved its submission.

Conflicts of Interest

This study was carried out to satisfy part of the requirements for a Master of Psychology (Clinical) degree for JH at the University of Newcastle, Australia. Both JH and AD are employed by the NSW Health Service. There are no further conflicts of interest to declare.

Multimedia Appendix 1
Screenshots of the iHeLP Program.
[PDF File (Adobe PDF File), 4879 KB - formative_v4i12e14344_app1.pdf]

References

1. Degenhardt L, Hall W. Extent of illicit drug use and dependence, and their contribution to the global burden of disease. Lancet 2012 Jan 7;379(9810):55-70. [doi: 10.1016/S0140-6736(11)61138-0] [Medline: 22225671]
2. UNODC Annual Report 2014. United Nations Office on Drugs and Crime (UNODC). 2014. URL: https://www.unodc.org/documents/AnnualReport2014/Annual_Report_2014_WEB.pdf [accessed 2020-06-27]
3. National Drug Strategy Household Survey. Australian Institute of Health and Welfare. 2013. URL: https://www.aihw.gov.au/about-our-data/our-data-collections/national-drug-strategy-household-survey/2013-national-drug-strategy-household-survey-maste [accessed 2020-06-27]

4. WHO Report on the Global Tobacco Epidemic 2019. World Health Organization. 2019. URL: https://www.who.int/tobacco/global_report/en/ [accessed 2020-06-27]
5. Baker A, Ivers R, Bowman J, Butler T, Kay-Lambkin F, Wye P, et al. Where there's smoke, there's fire: high prevalence of smoking among some sub-populations and recommendations for intervention. Drug Alcohol Rev 2006 Jan;25(1):85-96. [doi: 10.1080/09595230500459552] [Medline: 16492581]
6. Peters EN, Schwartz RP, Wang S, O'Grady KE, Blanco C. Psychiatric, psychosocial, and physical health correlates of co-occurring cannabis use disorders and nicotine dependence. Drug Alcohol Depend 2014 Jan 1;134:228-234 [FREE Full text] [doi: 10.1016/j.drugalcdep.2013.10.003] [Medline: 24183498]
7. Agrawal A, Lynskey MT, Madden PA, Pergadia ML, Bucholz KK, Heath AC. Simultaneous cannabis and tobacco use and cannabis-related outcomes in young women. Drug Alcohol Depend 2009 Apr 1;101(1-2):8-12 [FREE Full text] [doi: 10.1016/j.drugalcdep.2008.10.019] [Medline: 19081202]
8. Rooke SE, Norberg MM, Copeland J, Swift W. Health outcomes associated with long-term regular cannabis and tobacco smoking. Addict Behav 2013 Jun;38(6):2207-2213. [doi: 10.1016/j.addbeh.2013.01.013] [Medline: 23501136]
9. Agrawal A, Madden PA, Bucholz KK, Heath AC, Lynskey MT. Transitions to regular smoking and to nicotine dependence in women using cannabis. Drug Alcohol Depend 2008 May 1;95(1-2):107-114 [FREE Full text] [doi: 10.1016/j.drugalcdep.2007.12.017] [Medline: 18325694]
10. Amos A, Wiltshire S, Bostock Y, Haw S, McNeill A. 'You can't go without a fag...you need it for your hash'--a qualitative exploration of smoking, cannabis and young people. Addiction 2004 Jan;99(1):77-81. [doi: 10.1111/j.1360-0443.2004.00531.x] [Medline: 14678065]
11. Ford DE, Vu HT, Anthony JC. Marijuana use and cessation of tobacco smoking in adults from a community sample. Drug Alcohol Depend 2002 Aug 1;67(3):243-248. [doi: 10.1016/s0376-8716(02)00066-2] [Medline: 12127195]
12. Hindocha C, Shaban ND, Freeman TP, Das RK, Gale G, Schafer G, et al. Associations between cigarette smoking and cannabis dependence: a longitudinal study of young cannabis users in the United Kingdom. Drug Alcohol Depend 2015 Mar 1;148:165-171 [FREE Full text] [doi: 10.1016/j.drugalcdep.2015.01.004] [Medline: 25622777]
13. Peters E, Budney AJ, Carroll KM. Clinical correlates of co-occurring cannabis and tobacco use: a systematic review. Addiction 2012 Aug;107(8):1404-1417 [FREE Full text] [doi: 10.1111/j.1360-0443.2012.03843.x] [Medline: 22340422]
14. Tan WC, Lo C, Jong A, Xing L, Fitzgerald MJ, Vollmer WM, Vancouver Burden of Obstructive Lung Disease (BOLD) Research Group. Marijuana and chronic obstructive lung disease: a population-based study. Can Med Assoc J 2009 Apr 14;180(8):814-820 [FREE Full text] [doi: 10.1503/cmaj.081040] [Medline: 19364790]
15. Agosti V, Levin FR. Predictors of treatment contact among individuals with cannabis dependence. Am J Drug Alcohol Abuse 2004;30(1):121-127. [doi: 10.1081/ada-120029869] [Medline: 15083557]
16. Stephens R, Roffman RA, Fearer SA, Williams C, Burke RS. The marijuana check-up: promoting change in ambivalent marijuana users. Addiction 2007 Jun;102(6):947-957. [doi: 10.1111/j.1360-0443.2007.01821.x] [Medline: 17523990]
17. Gates P, Copeland J, Swift W, Martin G. Barriers and facilitators to cannabis treatment. Drug Alcohol Rev 2012 May;31(3):311-319. [doi: 10.1111/j.1465-3362.2011.00313.x] [Medline: 21521384]
18. Ellingstad TP, Sobell LC, Sobell MB, Eickleberry L, Golden CJ. Self-change: a pathway to cannabis abuse resolution. Addict Behav 2006 Mar;31(3):519-530. [doi: 10.1016/j.addbeh.2005.05.033] [Medline: 15967588]
19. Vendetti J, McRee B, Miller M, Christiansen K, Herrell J, Marjuana Treatment Project Research Group. Correlates of pre-treatment drop-out among persons with marijuana dependence. Addiction 2002 Dec;97(Suppl 1):125-134. [doi: 10.1046/j.1360-0443.97.s01.8.x] [Medline: 12460134]
20. Akre C, Michaud P, Berchtold A, Suris J. Cannabis and tobacco use: where are the boundaries? A qualitative study on cannabis consumption modes among adolescents. Health Educ Res 2010 Feb;25(1):74-82. [doi: 10.1093/her/cyp027] [Medline: 19515745]
21. Thornton LK, Baker AL, Johnson MP, Lewin T. Perceived risk associated with tobacco, alcohol and cannabis use among people with and without psychotic disorders. Addict Behav 2013 Jun;38(6):2246-2251. [doi: 10.1016/j.addbeh.2013.02.003] [Medline: 23507458]
22. Prochaska JJ, Benowitz NL. The past, present, and future of nicotine addiction therapy. Annu Rev Med 2016;67:467-486 [FREE Full text] [doi: 10.1146/annurev-med-111314-033712] [Medline: 26332005]
23. Hofmann SG, Asnaani A, Vonk IJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. Cognit Ther Res 2012 Oct 1;36(5):427-440 [FREE Full text] [doi: 10.1007/s10608-012-9476-1] [Medline: 23459093]
24. Lee DC, Budney AJ, Brunette MF, Hughes JR, Etter J, Stanger C. Treatment models for targeting tobacco use during treatment for cannabis use disorder: case series. Addict Behav 2014 Aug;39(8):1224-1230 [FREE Full text] [doi: 10.1016/j.addbeh.2014.04.010] [Medline: 24813547]
25. Rooke S, Thorsteinsson E, Karpin A, Copeland J, Allsop D. Computer-delivered interventions for alcohol and tobacco use: a meta-analysis. Addiction 2010 Aug;105(8):1381-1390. [doi: 10.1111/j.1360-0443.2010.02975.x] [Medline: 20528806]
26. Taylor G, Dalili MN, Semwal M, Civljak M, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2017 Sep 4;9:CD007078 [FREE Full text] [doi: 10.1002/14651858.CD007078.pub5] [Medline: 28869775]

27. Graham AL, Cobb NK, Papandonatos GD, Moreno JL, Kang H, Tinkelman DG, et al. A randomized trial of Internet and telephone treatment for smoking cessation. Arch Intern Med 2011 Jan 10;171(1):46-53 [FREE Full text] [doi: 10.1001/archinternmed.2010.451] [Medline: 21220660]
28. Davis ML, Powers MB, Handelsman P, Medina JL, Zvolensky M, Smits JA. Behavioral therapies for treatment-seeking cannabis users: a meta-analysis of randomized controlled trials. Eval Health Prof 2015 Mar;38(1):94-114 [FREE Full text] [doi: 10.1177/0163278714529970] [Medline: 24695072]
29. Drug and Alcohol Psychosocial Interventions Professional Practice Guidelines. NSW Health. 2008. URL: https://www1.health.nsw.gov.au/pds/ActivePDSDocuments/GL2008_009.pdf [accessed 2020-06-27]
30. Banham L, Gilbody S. Smoking cessation in severe mental illness: what works? Addiction 2010 Jul;105(7):1176-1189. [doi: 10.1111/j.1360-0443.2010.02946.x] [Medline: 20491721]
31. Lovibond P, Lovibond S. 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] [Medline: 7726811]
32. Henry J, Crawford JR. The short-form version of the depression anxiety stress scales (DASS-21): construct validity and normative data in a large non-clinical sample. Br J Clin Psychol 2005 Jun;44(Pt 2):227-239. [doi: 10.1348/014466505X29657] [Medline: 16004657]
33. Fagerstrom KO, Heatherton TF, Kozlowski LT. Nicotine addiction and its assessment. Ear Nose Throat J 1990 Nov;69(11):763-765. [Medline: 2276350]
34. Weinberger AH, Reutenauer EL, Allen TM, Termine A, Vessicchio JC, Sacco KA, et al. Reliability of the Fagerström test for nicotine dependence, Minnesota nicotine withdrawal scale, and Tiffany questionnaire for smoking urges in smokers with and without schizophrenia. Drug Alcohol Depend 2007 Jan 12;86(2-3):278-282. [doi: 10.1016/j.drugalcdep.2006.06.005] [Medline: 16876968]
35. da Rocha NS, Power MJ, Bushnell DM, Fleck MP. The EUROHIS-QOL 8-item index: comparative psychometric properties to its parent WHOQOL-BREF. Value Health 2012 May;15(3):449-457 [FREE Full text] [doi: 10.1016/j.jval.2011.11.035] [Medline: 22583455]
36. Ryan A, Holmes J, Hunt V, Dunlop A, Mammen K, Holland R, et al. Validation and implementation of the Australian treatment outcomes profile in specialist drug and alcohol settings. Drug Alcohol Rev 2014 Jan;33(1):33-42. [doi: 10.1111/dar.12083] [Medline: 24206571]
37. 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]
38. Attkisson C, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. Eval Program Plann 1982;5(3):233-237. [doi: 10.1016/0149-7189(82)90074-x] [Medline: 10259963]
39. Maclaren DJ, Conigrave KM, Robertson JA, Ivers RG, Eades S, Clough AR. Using breath carbon monoxide to validate self-reported tobacco smoking in remote Australian indigenous communities. Popul Health Metr 2010 Feb 20;8(1):2 [FREE Full text] [doi: 10.1186/1478-7954-8-2] [Medline: 20170528]
40. Shafik A. A concept of the anatomy of the anal sphincter mechanism and the physiology of defecation. Dis Colon Rectum 1987 Dec;30(12):970-982. [doi: 10.1007/BF02554289] [Medline: 3319451]
41. Zbikowski SM, Hapgood J, Barnwell SS, McAfee T. Phone and web-based tobacco cessation treatment: real-world utilization patterns and outcomes for 11,000 tobacco users. J Med Internet Res 2008 Nov 14;10(5):e41 [FREE Full text] [doi: 10.2196/jmir.999] [Medline: 19017583]
42. Myung S, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937. [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
43. Christensen H, Batterham PJ, Soubelet A, Mackinnon AJ. A test of the interpersonal theory of suicide in a large community-based cohort. J Affect Disord 2013 Jan 25;144(3):225-234. [doi: 10.1016/j.jad.2012.07.002] [Medline: 22862889]
44. Magill M, Ray LA. Cognitive-behavioral treatment with adult alcohol and illicit drug users: a meta-analysis of randomized controlled trials. J Stud Alcohol Drugs 2009 Jul;70(4):516-527 [FREE Full text] [doi: 10.15288/jsad.2009.70.516] [Medline: 19515291]
45. Healey A, Kay-Lambkin F, Bowman J, Childs S. Avoiding emotional bonds: an examination of the dimensions of therapeutic alliance among cannabis users. Front Psychiatry 2013;4:70 [FREE Full text] [doi: 10.3389/fpsyt.2013.00070] [Medline: 23885242]
46. Civljak M, Stead LF, Hartmann-Boyce J, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2013 Jul 10(7):CD007078. [doi: 10.1002/14651858.CD007078.pub4] [Medline: 23839868]
47. Norberg MM, Mackenzie J, Copeland J. Quantifying cannabis use with the timeline followback approach: a psychometric evaluation. Drug Alcohol Depend 2012 Mar 1;121(3):247-252. [doi: 10.1016/j.drugalcdep.2011.09.007] [Medline: 21955364]
48. Akre C, Michaud PA, Berchtold A, Suris JC. Cannabis and tobacco use: where are the boundaries? A qualitative study on cannabis consumption modes among adolescents. Health Educ Res 2010 Feb;25(1):74-82. [doi: 10.1093/her/cyp027] [Medline: 19515745]

49. Copeland J, Allsop D. Dedicated 'cannabis only' treatment clinics in New South Wales, Australia: client and treatment characteristics and associations with first-time treatment seeking. Drug Alcohol Rev 2014 Sep;33(5):565-571. [doi: 10.1111/dar.12166] [Medline: 24902861]
50. Banbury A, Zask A, Carter SM, van Beurden E, Tokley R, Passey M, et al. Smoking mull: a grounded theory model on the dynamics of combined tobacco and cannabis use among adult men. Health Promot J Austr 2013 Aug;24(2):143-150. [doi: 10.1071/HE13037] [Medline: 24168742]

Abbreviations

ATOP: Australian Treatment Outcome Profile
CBT: cognitive behavioral therapy
CO: carbon monoxide
CSQ: Client Service Questionnaire
DASS: Depression Anxiety Stress Scale
FTND: Fagerstrom Test for Nicotine Dependence
HNELHD: Hunter New England Local Health District
iHeLP: internet-delivered Healthy Lifestyles Program
MI: motivational interviewing
NRT: nicotine replacement therapy
NSW: New South Wales
QoL: quality of life

Edited by G Strudwick; submitted 11.04.19; peer-reviewed by JP Allem, J Thrul; comments to author 21.05.19; revised version received 03.05.20; accepted 02.06.20; published 07.12.20.

Please cite as:
Hindson J, Hanstock T, Dunlop A, Kay-Lambkin F
Internet-Delivered Tobacco Treatment for People Using Cannabis: A Randomized Trial in Two Australian Cannabis Clinics
JMIR Form Res 2020;4(12):e14344
URL: https://formative.jmir.org/2020/12/e14344
doi:10.2196/14344
PMID:33284121

Original Paper

Use of Different Food Image Recognition Platforms in Dietary Assessment: Comparison Study

Stephanie Van Asbroeck[1], MSc; Christophe Matthys[1], PhD

[1]Department of Chronic Diseases and Metabolism, Clinical and Experimental Endocrinology, KU Leuven, Leuven, Belgium

Corresponding Author:
Christophe Matthys, PhD
Department of Chronic Diseases and Metabolism
Clinical and Experimental Endocrinology
KU Leuven
Herestraat 49
DIEET
Leuven
Belgium
Phone: 32 016 34 26 55
Email: christophe.matthys@uzleuven.be

Abstract

Background: In the domain of dietary assessment, there has been an increasing amount of criticism of memory-based techniques such as food frequency questionnaires or 24 hour recalls. One alternative is logging pictures of consumed food followed by an automatic image recognition analysis that provides information on type and amount of food in the picture. However, it is currently unknown how well commercial image recognition platforms perform and whether they could indeed be used for dietary assessment.

Objective: This is a comparative performance study of commercial image recognition platforms.

Methods: A variety of foods and beverages were photographed in a range of standardized settings. All pictures (n=185) were uploaded to selected recognition platforms (n=7), and estimates were saved. Accuracy was determined along with totality of the estimate in the case of multiple component dishes.

Results: Top 1 accuracies ranged from 63% for the application programming interface (API) of the Calorie Mama app to 9% for the Google Vision API. None of the platforms were capable of estimating the amount of food. These results demonstrate that certain platforms perform poorly while others perform decently.

Conclusions: Important obstacles to the accurate estimation of food quantity need to be overcome before these commercial platforms can be used as a real alternative for traditional dietary assessment methods.

(JMIR Form Res 2020;4(12):e15602) doi:10.2196/15602

KEYWORDS

image recognition; dietary assessment; automated food recognition; accuracy

Introduction

For many years, research has been conducted in order to elucidate the relationship between diet and health outcomes [1], and associations between diet and noncommunicable diseases like type 2 diabetes, cancer, and obesity have been observed [2-4]. These findings have recently been questioned due to criticism of the dietary assessment methods used [5-7]. In epidemiological research, diet is most often assessed via food frequency questionnaires or 24-hour dietary recalls [8-10]. These assessment methods have long been known to be sensitive to both random and systematic error because they are based on self-reporting and memory of the participant [10,11]. However, better alternatives are not readily available. To assess diet in a research context, cost and time obviously play a role [11]. Additionally, the burden for participants needs to be reduced as much as possible [11,12]. Furthermore, cognition of the participant is another factor to take into account. Here, especially, the estimation of food amounts has proven to be a difficult task for most [8,11].

Therefore, the accurate assessment of diet currently remains a problem [13]. One suggested alternative is the automated analysis of pictures of the participant's food via image recognition [8,13,14]. The automated analysis would not require a high level of literacy or the recall of what was eaten in the past [15]. Additionally, it would reduce the systematic bias that is almost inherent to self-reporting of dietary intake toward what is socially preferred [11].

Nonetheless, automated image recognition of food along with the estimation of the amount of food shown in the image is not an easy task. Food items usually do not have a set shape or structure. Instead, they are deformable in nature. Additionally, similar foods can look very different (high intraclass variance) while different foods can look quite similar (interclass variance). Beverages are especially difficult to assess because only a limited amount of information can be acquired visually [16].

Multiple groups have tried to tackle the challenge that image recognition of food poses. Usually these groups also conduct tests on the performance of their final product. Often they report quite good results with accuracies ranging from 44% to 97.2% [8,13,16-21]. However, they often test their system in ideal conditions (eg, a well-lit room, food on a white plate nicely photographed in the frame) [13]. Another option is that they test their system on pictures from a food dataset, sometimes the same dataset that was used for training of the recognition system, so called cross-validation of machine learning [16,18,20,21]. Several food image recognition platforms have been created, but it is important to choose or create a dataset representative of food consumed in real-life situations [18]. The latter is key in the future development of mobile health apps that would assist patients in the context of dietary management of specific pathologies. A recent paper reviewed the possibilities of different food object recognition systems for use on a mobile phone. Based on their findings, they provided a categorization of mobile food recognition systems but reported that it is difficult to conduct an objective comparison between different systems as they all use different testing scenarios [22]. Therefore, a comparative study was set up to determine how well commercially available food image recognition systems currently perform in different settings and what the biggest challenges are when striving for accurate identifications.

Methods

Study Design

A comparative study on the accuracy of image recognition platforms for the recognition of food and beverages was conducted. Relevant application programming interfaces (APIs) and apps were searched for and all were submitted for pretesting. The tested platforms are described in Table 1.

Amazon Rekognition was not tested because this platform does not offer a demo version. Additionally, Snappy Meal was incompatible with our system and could therefore not be tested. All pictures used in this study were taken on the same device, a Galaxy Tab A6 (Samsung Electronics Co Ltd) tablet.

Table 1. Selected platforms identifying food images.

Platform	Version	Specifically developed for food
Google Vision API[a]	Unknown	No
IBM Watson Recognition	Unknown	No; employs a food module when estimating food in general model
Amazon Rekognition	Unknown	No
LogMeal	Unknown	Yes
FoodAI	Unknown	Yes
Clarifai	1	No; we used the included module developed specifically for food
Snappy Meal	0.0.1	Yes
Lose It	9.6.14	Yes
Bitesnap	1.5.6	Yes
Foodvisor	2.3	Yes
Calorie Mama API	Unknown	Yes

[a] API: application program interface.

Pretesting

Pretesting was done in order to select those platforms that could recognize 3 well-known food items considered easy to recognize (a whole banana, one slice of white bread, one round cookie); a clementine was also included in the test. These food items were photographed with a tablet by the same researcher. Foods were presented on a white round plate in a well-lit room on a table free of clutter. The whole plate was always in frame. Pictures were taken in 3264×2448 pixel size. The same picture of each food item was fed into all recognition platforms in random order on 1 day. Calorie Mama API required pictures to be resized to 544×544 pixels. Hence, for Calorie Mama API, the pictures were first resized using Pixlr Express web tool. Identifications of the image recognition process were gathered.

If a recognition platform succeeded in the correct top 1 identifications of at least 1 out of 3 food items combined with at least some form of recognition of the clementine as being a citrus, orange, or tangerine, it would be selected for more thorough testing.

Photographing Foods and Beverages

For further detailed testing of the image recognition platforms, pictures of a wide range of foods and beverages were taken in different settings. The ideal setting is in line with the guidelines defined by Rhyner et al [23]. As illumination and occlusion can play a role in the quality of the picture and consequently in the possibility of recognizing a food image, different conditions have been created. Both settings (eg, light, angle, height) and circumstances (eg, other objects on a table) can decrease the quality of the food image recognition. Using a stepwise approach, we have created different conditions. The settings used are defined in Table 2, and examples are shown in Figure 1. The use of different settings allowed us to compare the performance of the image recognition platforms in specific circumstances.

The original size of all pictures was 1920×1920 pixels, and all were resized to 544×544 pixels. Selection of food items, meals, and drinks was done to achieve a set of widely varying dishes in terms of components and difficulty while also keeping in mind selecting well-known, regularly consumed foods and beverages. The selection of the foods followed a stepwise approach. First, 5 simple, plain food items were selected. These were defined as single, unprocessed food items that can be eaten on their own. Next, we selected 5 plain, processed food items. Again, these are single food items that can be eaten on their own. Additionally, we selected 7 common hot and cold drinks. Last, 12 mixed dishes were selected with widely varying key components. The selected foods, beverages, and meals reflect the Belgian dietary pattern. All selected items are listed in Multimedia Appendix 1. Each food, beverage, or meal was photographed 6 times, each time in a different setting. All pictures were taken in the same room on the same solid colored table except for the real-life setting. All pictures were taken by the same researcher unless specified otherwise. All pictures are available in Multimedia Appendix 2.

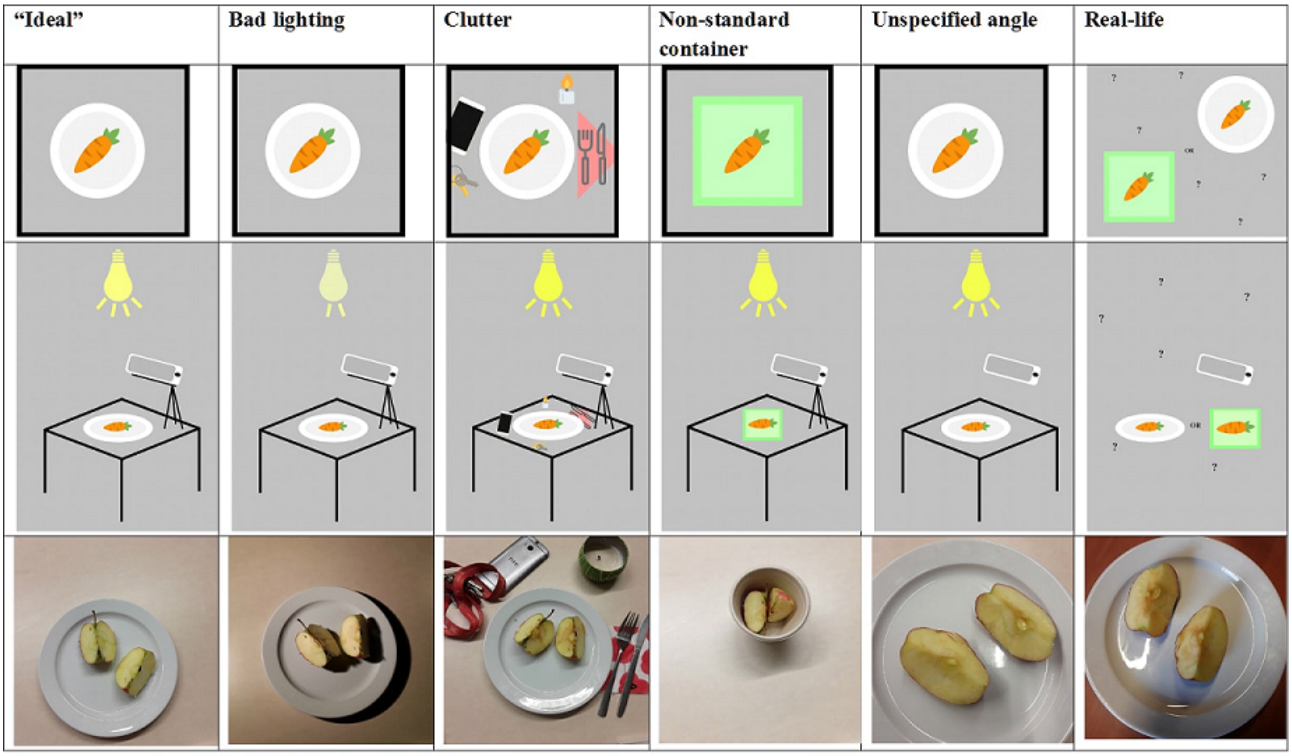
Table 2. Predefined settings and circumstances pictures were taken.

Setting	Circumstances			
	Lighting	Other objects on table	Container	Angle and height of tablet
Ideal	Well lit	None	Standard[a]	• Standardized height (whole plate needed to be in frame) • Standardized angle (15±3°) [23]
Bad lighting	Poorly lit	None	Standard	• Standardized height (whole plate needed to be in frame) • Standardized angle (15±3°) [23]
Clutter	Well lit	Fork, knife, napkin, keys, candle, and smartphone	Standard	• Standardized height (whole plate needed to be in frame) • Standardized angle (15±3°) [23]
Nonstandard container	Well lit	None	Nonstandard[b]	• Standardized height (whole plate needed to be in frame) • Standardized angle (15±3°) [23]
Unspecified angle	Well lit	None	Standard	Undefined: coworkers were asked to take the picture without guidance on how to hold the tablet; no specification about height and angle
Real-life	Undefined	Undefined	About one-half on standard container and one-half on nonstandard container	Undefined: coworkers were asked to take the picture without guidance on how to hold the tablet; no specification about height and angle

[a] Standard containers: white, round plate and translucent glass or white mug.

[a] Nonstandard containers: colored glasses, plates with designs and different shapes, colored mugs, bowls made out of different materials and colors.

Figure 1. Scheme and example pictures of the photographing settings used.



Automatic Image Recognition

All images were entered into the selected image recognition platforms (Table 1). This procedure was done in random order within 2 consecutive days during mid-December 2018 to avoid any automatic learning curve of any of the platforms. The top 1 and top 5 best identifications were saved.

Calculations

Reference Food Items for Identification

For each image, it was assessed whether the top 1 and top 5 identifications were correct; this was deemed correct when the key food item was identified correctly. More specifically, meats and poultry had to be identified to the level of species or specific dish (eg, stew, curry). Fruits and vegetables had to be identified at least to the level of a core term as defined in FoodEx2, the food classification system used by the European Food Safety Authority to standardize nutrient and exposure assessment. Carbohydrate sources had to be identified to the level of the key ingredient (eg, pasta, rice, or potatoes). Pancakes had to be identified as such. Cornflakes had to be identified as cereal, cornflakes, granola, or muesli. A Snickers bar could be identified as being Snickers, a candy bar, chocolate bar, or chocolate (fudge). Chocolate spread could be identified as such, as Nutella, or as chocolate. Milk had to be identified as such. Detail on fat content was not necessary or expected, but soy milk had to be identified specifically as soy milk to be deemed correct. Vanilla pudding had to be identified as custard, pudding, or a term describing a highly similar dish such as crème caramel or crème Catalan. Different sauces that were present in the picture needed to be identified at least in general as being a sauce. Water, coffee, tea, and beer needed to be identified as such, while for Coca Cola an identification of soda, soft drink, or Coke was deemed correct. For mixed dishes, an identification was deemed correct if at least one of the key components was identified correctly. Totality of the identification was also assessed for mixed dishes (see below). The top 1 and top 5 accuracy were calculated as described below.

Accuracy

Accuracy was calculated as the number of images correctly classified (top 1 and top 5) divided by the total number of images for each image recognition platform, each setting, and each of the predefined food categories (simple plain, simple processed, drinks, mixed dishes). To measure the accuracy of the different platforms, we used the food terms proposed by the platforms. These food terms are compared with the descriptions of the selected food items.

In addition, the image recognition tool should recognize all the individual components or use a term that describes these multiple components (eg, Quiche Lorraine) when provided with an image of a mixed dish. Hence, for mixed dishes, the proportion of components that were correctly classified within the top 5 identifications was also calculated.

Totality

Totality was calculated as the number of dish components correctly classified divided by the total number of dish components. A list of all mixed dishes along with their respective key components can be found in Multimedia Appendix 3. None of the mixed dishes had more than 5 key components.

Statistical Analysis

Accuracy was compared among the different image recognition platforms as well as among the different settings and dish categories using chi-square tests. All tests were considered significant when *P*<.05. Tests were performed using SPSS Statistics 25 (IBM Corporation).

Results

Pretesting

Based on the platforms' performances during pretesting, Bitesnap, Foodvisor, LogMeal, Clarifai, IBM Watson Recognition, Google Vision API, and Calorie Mama API were selected for further testing. Snappy Meal did not return any results during pretesting and was thus deemed to malfunction.

Accuracy

In total, 185 pictures were uploaded on the different platforms. This number consists of 30 images of simple, plain foods (5 foods × 6 settings), 36 images of simple, processed foods (6 foods × 6 settings), 47 images of beverages (8 drinks × 6 settings plus 1 for Coca Cola in a can, which cannot be photographed in a nonstandard container), and 72 images of mixed dishes (12 dishes × 6 settings). Recognition accuracy based on the top 1 and top 5 best identifications was compared among the different recognition platforms and varied widely. Both for top 1 and top 5 identifications, Calorie Mama API was most accurate with 62.9% (117/185) and 87.6% (163/185), respectively. Bitesnap also performed well with a top 1 and top 5 accuracy of 48.9% (91/185) and 71.0% (132/185), respectively. Foodvisor performed at a similar level with a top 1 and top 5 accuracy of 46.2% (86/185) and 71.5% (133/185), respectively. The results of all platforms among all settings are shown in Table 3.

Table 3. Accuracy of food recognition among different platforms and different settings (n=31/setting).

Platform	Ideal, n (%)	Bad lighting, n (%)	Nonstandard container, n (%)	Clutter, n (%)	Unspecified angle, n (%)	Real-life, n (%)	Total, n (%)
Top 1 accuracy							
LogMeal	8 (25.8)	6 (19.4)	5 (16.1)	6 (19.4)	12 (38.7)	8 (25.8)	45 (24.2)
Clarifai	12 (38.7)	9 (29.0)	12 (38.7)	8 (25.8)	17 (54.8)	13 (41.9)	71 (38.2)
Google Vision API[a]	2 (6.5)	3 (9.7)	1 (3.2)	0 (0)	7 (22.6)	4 (12.9)	17 (9.1)
IBM Watson Recognition	10 (32.3)	5 (16.1)	5 (16.1)	7 (226)	8 (25.8)	12 (38.7)	47 (25.3)
Calorie Mama API	20 (64.5)	16 (51.6)	18 (58.1)	19 (61.3)	24 (77.4)	20 (64.5)	117 (62.9)
Bitesnap	16 (51.6)	11 (35.5)	13 (41.9)	15 (45.2)	21 (67.7)	16 (51.6)	91 (48.9)
Foodvisor	10 (32.3)	13 (41.9)	15 (48.4)	16 (51.6)	19 (61.3)	13 (41.9)	86 (46.2)
Total	78 (35.9)	63 (29.0)	69 (32.9)	70 (32.3)	108 (49.8)	86 (39.6)	—[b]
Top 5 accuracy							
LogMeal	14 (45.2)	13 (41.9)	10 (32.3)	14 (45.2)	17 (54.8)	14 (45.2)	82 (44.1)
Clarifai	21 (67.7)	19 (61.3)	18 (58.1)	18 (58.1)	23 (74.2)	20 (64.5)	119 (64.0)
Google Vision API	7 (22.6)	7 (22.6)	2 (6.5)	3 (9.7)	16 (51.6)	10 (32.3)	45 (24.2)
IBM Watson Recognition	15 (48.4)	10 (32.3)	10 (32.3)	10 (32.3)	20 (64.5)	16 (51.6)	81 (43.5)
Calorie Mama API	27 (87.1)	28 (90.3)	25 (80.6)	26 (83.9)	29 (93.5)	28 (90.3)	163 (87.6)
Bitesnap	23 (74.2)	20 (64.5)	20 (64.5)	21 (67.7)	28 (90.3)	20 (64.5)	132 (71.0)
Foodvisor	19 (61.3)	19 (61.3)	24 (77.4)	21 (67.7)	26 (83.9)	24 (77.4)	133 (71.5)
Total	126 (58.1)	116 (53.5)	109 (51.9)	113 (52.1)	159 (73.3)	132 (60.8)	—

[a] API: application programming interface.

[b] Not applicable.

When comparing the different settings, the unspecified angle setting resulted in a better recognition accuracy compared with all other settings. A comparison of the recognition accuracy among the different settings is shown in Table 3.

The accuracy was also compared among different food and beverage categories (eg, simple plain foods, simple processed foods, beverages, and mixed dishes, Table 4). Simple, plain foods, simple, processed foods, and beverages were all recognized less accurately than mixed dish components at top 1 level (*P*<.001, *P*=.02, and *P*=.004, respectively). At the top 5 level, mixed dish components were recognized significantly more accurately than drinks (*P*=.01). Additionally, simple, plain foods were recognized significantly less accurately than simple, processed foods, and drinks (*P*=.02 and *P*=.04, respectively) at the top 1 level.

Table 4. Accuracy of image recognition among different food categories.

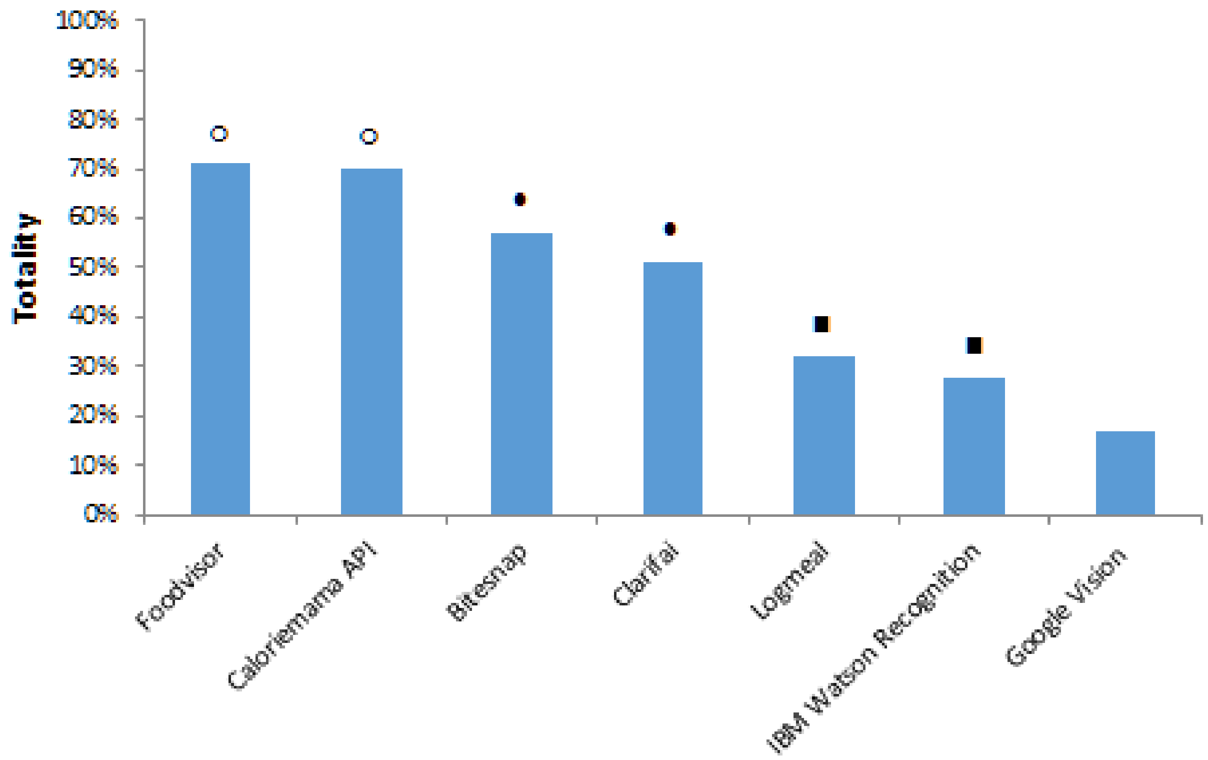
Platform	Food category n (%)							
	Simple plain		Simple processed		Beverages		Mixed dishes	
	Top 1	Top 5	Top 1	Top 5	Top 1	Top 5	Top 1	Top 5
Bitesnap	11 (36.6)	25 (83.3)	14 (46.6)	18 (60.0)	22 (53.6)	33 (80.4)	44 (61.1)	56 (77.7)
Calorie Mama API[a]	22 (73.3)	29 (96.6)	23 (76.6)	31 (100)	19 (46.3)	29 (70.7)	50 (69.4)	70 (97.2)
Clarifai	4 (13.3)	13 (43.3)	14 (46.6)	27 (90.0)	17 (41.4)	30 (73.1)	36 (50.0)	49 (68.1)
Foodvisor	3 (10.0)	15 (50.0)	16 (53.3)	22 (73.3)	23 (56.1)	32 (78.1)	44 (61.1)	64 (88.8)
Google Vision API	0 (0)	0 (0)	12 (40.0)	17 (56.6)	1 (2.4)	9 (21.9)	4 (5.5)	19 (26.4)
IBM Watson Recognition	5 (16.6)	16 (53.3)	9 (30.0)	16 (53.3)	4 (9.7)	12 (29.2)	29 (40.2)	37 (51.3)
LogMeal	9 (30.0)	23 (76.6)	3 (10.0)	13 (43.3)	13 (31.7)	15 (36.5)	20 (27.7)	31 (43.1)
All	54 (25.7)	121 (57.6)	91 (43.3)	148 (70.5)	99 (34.5)	160 (55.7)	227 (45.1)	326 (64.6)

[a] API: application programming interface.

Totality

Totality of identifications for each platform are shown in Figure 2. Foodvisor and Calorie Mama API succeeded in the estimation of most dish components in their top 5 identifications with a totality of 70.8% (119/168) and 69.6% (117/168), respectively. Bitesnap and Clarifai identified mixed dish components with a totality of 56.8% (96/168) and 50.6% (85/168), respectively.

Figure 2. Totality of estimates on mixed dishes among different recognition platforms. Non-significant differences are denoted with a symbol placed with both platforms. All other platforms estimated mixed dish content to a significantly different level of totality (all *P* values ≤.001).



Discussion

Principal Findings

A series of currently available image recognition platforms were tested for their performance on the recognition of foods, beverages, and mixed dishes conducted in a number of different standardized settings and in a real-life setting. Performance in terms of accuracy was found to vary widely between platforms ranging from 9% to 63% for top 1 accuracy. Top 5 accuracy ranged from 24% to 88% over all tested platforms. For mixed dishes, a high variance in the number of recognized dish components among the platforms was found. Foodvisor and Calorie Mama API succeeded in recognition of most of the components of the mixed dish (71% and 70%, respectively) while Google Vision API only succeeded in the recognition of 17% of the key components of the mixed dishes. None of the platforms was able to identify the respective portion size of the different foods and beverages.

Comparison of Platforms

The different tested platforms showed large differences in their performance. Calorie Mama API was the most accurate platform with a top 1 accuracy of 63% and a top 5 accuracy of 88%. Bitesnap and Foodvisor came in second place with a top 1 accuracy of 49% and 46%, respectively, and a top 5 accuracy of 71% and 72%, respectively. Third, Clarifai achieved a top 1 accuracy of 38% and a top 5 accuracy of 64%. Furthermore, IBM Watson Recognition and LogMeal were less accurate than all previously mentioned platforms. Google Vision API achieved the least accurate results, with a top 1 and top 5 accuracy of 9% and 24%, respectively. For Google Vision API, it seemed the recognition was influenced by how well the food or beverage filled the frame. This seems to be the reason Google Vision API appears to perform better under the unspecified angle setting (top 1 accuracy 23%; top 5 accuracy 52%). Under this setting, the majority of the pictures were taken closer to the food or beverage compared with the standard settings. When a food item was relatively small in volume, lying on a plate that was fully in frame, this would result in the food itself filling a small part of the full frame. The latter seemed to be a difficulty for Google Vision API. Among the tested platforms, Google Vision API is not specifically designed for food recognition, contrary to many other APIs and apps tested in this study. Google Vision API would often recognize a plate or dish in the picture but would give no indication as to what was on the plate. IBM Watson Recognition is also not specifically developed for the recognition of food, but they do provide a specific food module if the platform’s general model identifies the picture as a food. Overall, the tested platforms vary hugely in their performance. Due to the variation it remains difficult to recommend a specific platform in the context of dietary assessment.

Comparison of Settings

Based on the comparison of the different settings among all platforms, the setting influenced both the accuracy and totality. The accuracy achieved in the ideal setting was only slightly higher compared with the ones reached in the nonstandardized containers setting, the setting with clutter in frame, and the badly lit setting. Moreover, the unspecified angle setting resulted in a significantly higher recognition accuracy compared with all other settings. Although unexpected, this is actually a promising result since it means that recognition platforms perform better when supplied with pictures taken at angles and distances that feel natural to the consumer or participant compared with pictures that are taken at a standardized angle and height.

Furthermore, specific differences in recognition accuracy were noticed between platforms and settings. Google Vision API seemed unable to perform well when there was clutter in the frame. Additionally, the nonstandard container setting appeared to pose quite a problem for Google Vision API and LogMeal. On the other hand, Clarifai appeared to struggle most with bad lighting and having clutter in the frame. IBM Watson Recognition had the most difficulties when the lighting was bad or when the food or drink was in a nonstandard container. Bitesnap also appeared to have some problems with bad lighting. Foodvisor seemed to perform the worst in the ideal setting. It is clear that the performance of each platform is influenced by different factors in different ways; potentially the algorithms performing in the back end of these platforms play a key role.

Comparison of Food Categories

When comparing the recognition accuracy over the different food categories, mixed dish components were found to be recognized with the highest accuracy at the top 1 level. This implies that when multiple dish components are in the frame, often at least one of them can be identified correctly. However, it does not imply that the platforms succeeded in recognizing all dish components. When comparing how well the different platforms could recognize all different components of mixed dishes (totality), Foodvisor and Calorie Mama API were clearly best at recognizing a dish in its totality. Simple processed foods and beverages were recognized better than simple plain foods at the top 1 level but not at top 5 level. Simple processed foods seem to be less variable visually than simple plain foods and are consequently better recognized. Further, it appeared that some platforms simply could not properly recognize certain simple, plain foods, possibly because the food term was not in their system. Another explanation for this difference could be that pictures of simple, processed foods are just more common than pictures of simple, plain foods. However, this is just a hypothesis since it is unknown how each of the image recognition platforms were trained. For beverages, the accuracy remained low.

Limitations

The main limitation of this study is the use of demo versions of the recognition platforms. Demo versions do show the general capabilities of the platform but could potentially be improved to allow proper dietary assessment. Apart from the technological aspects of the different platforms, there are no specific guidelines for the creation of different settings and circumstances; however, benchmark datasets (eg, ETHZ-Food-101) are quite close to real-life images. As our results show, the setting determines the accuracy of the different platforms. Further, the use of large public available datasets (eg, FoodX-251 or Food-5K) would allow evaluation and testing of the external validity of our findings. Pictures from real-life settings, especially, would be of added value. Our real-life setting tried to assess the real-life recognition accuracy but since we used specifically selected foods and beverages, it can only serve as a surrogate for the real-life situation. However, we had to use this research approach in order to allow for a fair comparison between the different settings. Previously, He et al [17] tested their own image recognition system on 1453 real-life food images taken by 45 free-living or community-dwelling individuals and reached a recognition accuracy of 34%. When they also included contextual information such as food co-occurrence patterns and personalized learning models in the recognition process, the recognition accuracy increased to 44%.

We did not correct for multiple testing because this study was conducted as an exploratory study of the current recognition accuracy of commercial available image recognition platforms.

To our knowledge, this is the first study that compared performances of currently available image recognition platforms with a focus on consumer use. Results were highly variable but promising nonetheless. However, currently there is no freely available recognition system that is already useable for dietary assessment in the context of research. To obtain nutrient intake assessments, food images could be coupled to a food composition database, which translates the food weight into nutritional values. Previously, the lack of a comprehensive food density database was limiting, but this has now been improved [15,24]. Multiple groups are working toward the goal of successful estimation of food amounts [13,18,23]. For example, Rhyner et al [23] worked specifically on a system for automated estimation of carbohydrate content of a meal via image recognition. They tested whether their system could help type 1 diabetes patients in estimating the carbohydrate content of their meals. When patients were asked to calculate the carbohydrate content on their own as they normally would, the authors found a mean absolute error of 27.89 (SD 38.20) g. When the patients were asked to use their system for automated estimation of the carbohydrate content, the mean absolute error was reduced to 12.28 (SD 9.56) g [23]. Furthermore, Zhu et al [13] developed and tested an image recognition system. For weight estimation of the food, they reported mean percentage error rates of 3.4% to 56.4% when tested on pictures of garlic bread and yellow cake. Therefore, we expect that the accurate estimation of food amount is a hurdle that can and will be overcome in the future. Combining image recognition with other sources of information (eg, time of day, container in which the food is in, and location) could be used and gathered using a smartphone. We expect that certain foods or beverages will never be distinguished via image recognition. Hence, when the recognition system identifies multiple foods or beverages with a similar high likelihood (eg, soy milk versus cow’s milk), simply asking the participant which one it is could be the best solution, as suggested by Eldridge et al [25] in their stepwise approach for reporting dietary assessment using different technologies. When required, additional questions could also be asked of the participant (eg, asking the fat content when the system recognizes a drink as being milk). By using a blended assessment form like this, it is expected that the use of image recognition technology for dietary assessment could not only be possible but could solve some of the problems faced by traditional dietary assessment methods.

Conclusion

To our knowledge, this is the first comparative study investigating the recognition accuracy of currently available image recognition platforms on food and drinks. We found the recognition accuracy to vary widely between different platforms ranging from poor to excellent. The estimation of portion sizes of foods or beverages is lacking in the tested platforms. A blended form of assessment, automated image recognition, and asking questions using chatbot options could improve the overall dietary assessment.

Acknowledgments

The research was conducted as part of the SmartFoodLogging project (grant number SMARTFOODLOGGING [KAVA 18102-19091]) and has been funded by EIT Food [26], a body of the European Union. The European Institute of Innovation and Technology (EIT) receives support from the Horizon 2020 Research and Innovation Programme. The content of the paper reflects only the views of the authors. The European Commission and EIT Food are not liable for any use that may be made of the information contained in this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Food and drink items to be used.
[DOCX File , 14 KB - formative_v4i12e15602_app1.docx]

Multimedia Appendix 2
Food images in different settings.
[ZIP File (Zip Archive), 38179 KB - formative_v4i12e15602_app2.zip]

Multimedia Appendix 3
Mixed dishes composition.
[DOCX File , 13 KB - formative_v4i12e15602_app3.docx]

References

1. Hébert JR, Hurley TG, Steck SE, Miller DR, Tabung FK, Peterson KE, et al. Considering the value of dietary assessment data in informing nutrition-related health policy. Adv Nutr 2014 Jul;5(4):447-455. [doi: 10.3945/an.114.006189] [Medline: 25022993]
2. Wang X, Lin X, Ouyang YY, Liu J, Zhao G, Pan A, et al. Red and processed meat consumption and mortality: dose-response meta-analysis of prospective cohort studies. Public Health Nutr 2016 Apr;19(5):893-905. [doi: 10.1017/S1368980015002062] [Medline: 26143683]
3. Malik VS, Li Y, Tobias DK, Pan A, Hu FB. Dietary protein intake and risk of type 2 diabetes in US men and women. Am J Epidemiol 2016 Apr 15;183(8):715-728 [FREE Full text] [doi: 10.1093/aje/kwv268] [Medline: 27022032]
4. Boggs DA, Rosenberg L, Coogan PF, Makambi KH, Adams-Campbell LL, Palmer JR. Restaurant foods, sugar-sweetened soft drinks, and obesity risk among young African American women. Ethn Dis 2013;23(4):445-451 [FREE Full text] [Medline: 24392607]
5. Archer E, Hand GA, Blair SN. Validity of U.S. nutritional surveillance: National Health and Nutrition Examination Survey caloric energy intake data, 1971-2010. PLoS One 2013;8(10):e76632 [FREE Full text] [doi: 10.1371/journal.pone.0076632] [Medline: 24130784]
6. Ioannidis JPA. Implausible results in human nutrition research. BMJ 2013 Nov 14;347:f6698. [doi: 10.1136/bmj.f6698] [Medline: 24231028]
7. Mitka M. Do flawed data on caloric intake from NHANES present problems for researchers and policy makers? JAMA 2013 Nov 27;310(20):2137-2138. [doi: 10.1001/jama.2013.281865] [Medline: 24281443]
8. Probst Y, Nguyen DT, Tran MK, Li W. Dietary assessment on a mobile phone using image processing and pattern recognition techniques: algorithm design and system prototyping. Nutrients 2015 Jul 27;7(8):6128-6138 [FREE Full text] [doi: 10.3390/nu7085274] [Medline: 26225994]
9. Ishihara J. Challenges in dietary exposure assessment in epidemiology: research trends. J Nutr Sci Vitaminol (Tokyo) 2015;61 Suppl:S33-S35 [FREE Full text] [doi: 10.3177/jnsv.61.S33] [Medline: 26598878]
10. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing current criticism regarding the value of self-report dietary data. J Nutr 2015 Dec;145(12):2639-2645 [FREE Full text] [doi: 10.3945/jn.115.219634] [Medline: 26468491]
11. Gibney MJ. Introduction to Human Nutrition. 2nd Edition. West Sussex: Wiley-Blackwell; 2009.
12. Shim J, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. Epidemiol Health 2014;36:e2014009 [FREE Full text] [doi: 10.4178/epih/e2014009] [Medline: 25078382]
13. Zhu F, Bosch M, Woo I, Kim S, Boushey CJ, Ebert DS, et al. The use of mobile devices in aiding dietary assessment and evaluation. IEEE J Sel Top Signal Process 2010 Aug;4(4):756-766 [FREE Full text] [doi: 10.1109/JSTSP.2010.2051471] [Medline: 20862266]
14. Illner A, Freisling H, Boeing H, Huybrechts I, Crispim SP, Slimani N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol 2012 Aug;41(4):1187-1203. [doi: 10.1093/ije/dys105] [Medline: 22933652]
15. Adibi S. mHealth diet and nutrition guidance. In: Probst Y, Nguyen DT, Rollo M, Li W, editors. mHealth Multidisciplinary Verticals. Boca Raton: CRC Press; 2015.
16. Mezgec S, Koroušić Seljak B. NutriNet: a deep learning food and drink image recognition system for dietary assessment. Nutrients 2017 Jun 27;9(7) [FREE Full text] [doi: 10.3390/nu9070657] [Medline: 28653995]
17. Wang Y, He Y, Boushey CJ, Zhu F, Delp EJ. Context based image analysis with application in dietary assessment and evaluation. Multimed Tools Appl 2018 Aug;77(15):19769-19794 [FREE Full text] [doi: 10.1007/s11042-017-5346-x] [Medline: 30202237]
18. Zhang W, Yu Q, Siddiquie B, Divakaran A, Sawhney H. “Snap-n-Eat”: food recognition and nutrition estimation on a smartphone. J Diabetes Sci Technol 2015 May;9(3):525-533 [FREE Full text] [doi: 10.1177/1932296815582222] [Medline: 25901024]
19. Zhu F, Bosch M, Boushey CJ, Delp EJ. An image analysis system for dietary assessment and evaluation. Proc Int Conf Image Proc 2010:1853-1856 [FREE Full text] [doi: 10.1109/ICIP.2010.5650848] [Medline: 22025261]
20. Anthimopoulos MM, Gianola L, Scarnato L, Diem P, Mougiakakou SG. A food recognition system for diabetic patients based on an optimized bag-of-features model. IEEE J Biomed Health Inform 2014 Jul;18(4):1261-1271. [doi: 10.1109/JBHI.2014.2308928] [Medline: 25014934]
21. He H, Kong F, Tan J. DietCam: multiview food recognition using a multikernel SVM. IEEE J Biomed Health Inform 2016 May;20(3):848-855. [doi: 10.1109/JBHI.2015.2419251] [Medline: 25850095]
22. Knez S, Šajn L. Food object recognition using a mobile device: evaluation of currently implemented systems. Trends Food Sci Technol 2020 May;99:460-471. [doi: 10.1016/j.tifs.2020.03.017]
23. 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]
24. Stumbo PJ, Weiss R. Using database values to determine food density. J Food Composition Analysis 2011 Dec;24(8):1174-1176. [doi: 10.1016/j.jfca.2011.04.008]

25. Eldridge AL, Piernas C, Illner A, Gibney MJ, Gurinović MA, de Vries JHM, et al. Evaluation of new technology-based tools for dietary intake assessment: an ILSI Europe Dietary Intake and Exposure Task Force evaluation. Nutrients 2018 Dec 28;11(1) [FREE Full text] [doi: 10.3390/nu11010055] [Medline: 30597864]
26. EIT Food. URL: https://www.eitfood.eu/ [accessed 2020-10-09]

Abbreviations

API: application programming interface
EIT: European Institute of Innovation and Technology

Edited by G Eysenbach, R Leung; submitted 23.07.19; peer-reviewed by E Delp, F Zhu; comments to author 26.02.20; revised version received 02.09.20; accepted 02.10.20; published 07.12.20.

Please cite as:
Van Asbroeck S, Matthys C
Use of Different Food Image Recognition Platforms in Dietary Assessment: Comparison Study
JMIR Form Res 2020;4(12):e15602
URL: https://formative.jmir.org/2020/12/e15602
doi:10.2196/15602
PMID:33284118

Viewpoint

Psychological Screening and Tracking of Athletes and Digital Mental Health Solutions in a Hybrid Model of Care: Mini Review

Luke Balcombe[1,2], BA, MSc, PhD; Diego De Leo[2], MD, PhD, DSc, FRCPE, FRANZCP

[1]School of Health and Sport Science, University of the Sunshine Coast, Sunshine Coast, Australia

[2]Australian Institute for Suicide Research and Prevention, Griffith University, Brisbane, Australia

Corresponding Author:
Luke Balcombe, BA, MSc, PhD
School of Health and Sport Science
University of the Sunshine Coast
90 Sippy Downs Drive
Sunshine Coast
Australia
Phone: 61 0447505709
Email: lukebalcombe@gmail.com

Abstract

Background: There is a persistent need for mental ill-health prevention and intervention among *at-risk* and vulnerable subpopulations. Major disruptions to life, such as the COVID-19 pandemic, present an opportunity for a better understanding of the experience of stressors and vulnerability. Faster and better ways of psychological screening and tracking are more generally required in response to the increased demand upon mental health care services. The argument that mental and physical health should be considered together as part of a biopsychosocial approach is garnering acceptance in elite athlete literature. However, the sporting population are unique in that there is an existing stigma of mental health, an underrecognition of mental ill-health, and engagement difficulties that have hindered research, prevention, and intervention efforts.

Objective: The aims of this paper are to summarize and evaluate the literature on athletes' increased vulnerability to mental ill-health and digital mental health solutions as a complement to prevention and intervention, and to show relationships between athlete mental health problems and resilience as well as digital mental health screening and tracking, and faster and better treatment algorithms.

Methods: This mini review shapes literature in the fields of athlete mental health and digital mental health by summarizing and evaluating journal and review articles drawn from PubMed Central and the Directory of Open Access Journals.

Results: Consensus statements and systematic reviews indicated that elite athletes have comparable rates of mental ill-health prevalence to the general population. However, peculiar subgroups require disentangling. Innovative expansion of data collection and analytics is required to respond to engagement issues and advance research and treatment programs in the process. Digital platforms, machine learning, deep learning, and artificial intelligence are useful for mental health screening and tracking in various subpopulations. It is necessary to determine appropriate conditions for algorithms for use in recommendations. Partnered with real-time automation and machine learning models, valid and reliable behavior sensing, digital mental health screening, and tracking tools have the potential to drive a consolidated, measurable, and balanced risk assessment and management strategy for the prevention and intervention of the sequelae of mental ill-health.

Conclusions: Athletes are an *at-risk* subpopulation for mental health problems. However, a subgroup of high-level athletes displayed a resilience that helped them to positively adjust after a period of *overwhelming* stress. Further consideration of stress and adjustments in brief screening tools is recommended to validate this finding. There is an unrealized potential for broadening the scope of mental health, especially symptom and disorder interpretation. Digital platforms for psychological screening and tracking have been widely used among general populations, but there is yet to be an eminent athlete version. Sports in combination with mental health education should address the barriers to help-seeking by increasing awareness, from mental ill-health to positive functioning. A hybrid model of care is recommended, combining traditional face-to-face approaches along with innovative and evaluated digital technologies, that may be used in prevention and early intervention strategies.

(JMIR Form Res 2020;4(12):e22755) doi:10.2196/22755

KEYWORDS

athletes; screening; tracking; engagement; well-being; stress; adjustment; COVID-19; hybrid model of care; digital mental health; machine learning; artificial intelligence

Introduction

Elite Athlete Mental Health Phenomenon and Digital Mental Health Potential

Elite athlete mental health problems recently attracted great discussion as a worldwide relevant phenomenon [1], including reference to sporting factors, such as injury, overtraining, burnout, and career termination as well as nonsporting factors [2,3]. The array of athlete-specific mental health symptoms and disorders such as distress, depression and suicide, anxiety and stress, overtraining, sport-related concussion, substance misuse, sleeping and eating disorders, posttraumatic stress disorder (PTSD) and other trauma-related disorders, bipolar and psychotic disorders, attention-deficit/hyperactivity disorder, gambling disorder, and other behavioral addictions [4-6], are complicated by a range of social and psychological factors as well as attributes including personality, sexuality and gender issues, hazing, bullying, sexual misconduct, and transition from sport [6].

An increased interest in understanding the complexity of mental health among elite athletes has led to the empirical development [7] and validation [8] of athlete-specific questionnaires that are acceptable and appropriate for measuring mental health symptoms in the athlete environment. There remains an unrealized potential for universal electronic screening for mental health concerns, especially as a preparticipation examination when competitive athletes transition to a higher level (administered by sporting organizations) [9]. The early detection of mental health disorders has been proposed as part of an athlete-specific early intervention framework involving web-based consultations [10]. There is potential for this approach to be expanded at times of life changes, transition, and retirement. However, flexibility and pragmatism are required to apply digital mental health tools such as smartphone apps cognizant of the dichotomous directions and efforts that divide this space and limit its potential (randomized trials vs pragmatic studies, precision medicine vs population health, free market vs regulation, consumer vs clinical uses, big data vs privacy, and open vs proprietary software) [11].

Athlete Mental Health Challenges Amid COVID-19

Although preceded by many other epidemics caused by infectious disease, the COVID-19 pandemic is worthy of particular exploration because of a trend of increasing emergence of novel pathogens from nonhuman hosts. Major worldwide disruptions to habits and customs have impacted elite athletes' training and competition. Uncharted challenges have been posed for those engaged in sport, with preliminary issues associated with social isolation, career disruption, qualification process uncertainty, and unconventional and limited access to effective training environments and partners [12]. This has been compounded by uncertainty around sporting events, which led to a negative response over a period of weeks to months [13]. However, some positives were drawn by the adaptive and resourceful group of those with well-formed athletic identities: a chance to refresh, time for injury recovery, and time for honing practice to fill perceived gaps in development [12]. Examples of novel training methods were shared by some high-profile elite athletes via social media during the COVID-19 lockdown.

There are calls to better understand the disruption and psychosocial consequences arising from the COVID-19 pandemic and predicted mental health case surge across general populations [14,15]. The urgent mental health care response makes it worthwhile to consider opportunities to pivot from already ongoing studies [16]. Research that disentangles aspects of athletes' experience of stress and adjustment issues during changes and transitions can provide important contributions to early intervention frameworks or models of care as well as educational resources. A clearer picture of physical and mental activity behaviors, well-being, stressors, and adjustment is required from athlete-specific screening and monitoring to better understand the holistic health of high-level athletes [17,18].

Recent longitudinal research with high-level athletes (Simons et al, unpublished data, 2020) and college students [19] identified these subgroups as *at risk* of mental health problems amid COVID-19. These studies included online (smartphone compatible) assessments. The results suggested the influence of *overwhelming* stress and changes in mental health and behaviors related to the pandemic situation. In college students, this was negatively altered with a restrictive effect related to the pandemic proximity, media coverage, and policy. The mixed linear model of smartphone mobile sensing and self-reported mental health questions [19] was applied in this study. In synthesis, the main theoretical basis informing these studies is the hybrid model of care that identifies patients receiving both in-person and online interventions for diagnosis, therapy, and monitoring [10,20-22].

Digital Mental Health Opportunity Amid COVID-19

Digital mental health interventions have been proposed as a flexible and cost-effective way to reduce stigma and the treatment gap among university students [23]. However, the limitations are low adherence rates and questionable long-term efficacy in real-world settings. A collaborative effort to confront digital and system challenges within the general population will be required to "realize the vital goals of mental health prevention and support by providing a means to measure and track population mental health" [24]. A road map, a communication tool that articulates the strategic thinking of the goal and plan, is required to strengthen global mental health systems to tackle the impact of the COVID-19 pandemic [25]. Maulik et al [25] drew upon key sources and accumulated knowledge of mental health systems worldwide to provide a perspective on practical steps for implementation.

An integrated blueprint to deliver digital mental health screening and tracking highlighted the need to provide specifications for vulnerable subpopulations such as children, college and university students, domestic violence victims, frontline heath care workers, low socioeconomic groups, those with mental health disorders, and older adults [26]. Athletes were initially included with this group. However, it should be clarified that high-level athletes are a peculiar *at-risk* group because there is an increased vulnerability yet resilience that requires further investigation and articulation. There is an opportunity to enable digital tools and artificial intelligence in athlete mental health screening and tracking to collect real-time data in a hybrid model of care (combining online and offline data with in-person data) to enrich research data to support better models of care [26]. This leads to the research question: should digital screening and tracking tools be tailored toward adjustment and stress in *at-risk* subpopulations such as athletes to assist the management of their mental health?

Methods

This mini review shapes literature in the fields of athlete mental health and digital mental health by summarizing and evaluating journal and review articles drawn from PubMed Central and the Directory of Open Access Journals (generally from within the last 4 years). There are also internet sources (refereed electronic journals), a thesis, conference proceedings, books and websites to complement the primary research. A story has been developed to show relationships between athlete mental health problems and resilience as well as digital mental health innovation with recommendations presented for research prospects in the near future.

Results

Elite Athlete Mental Health Screening and Tracking

The use of technology in elite athlete mental health screening has been limited to internet-based surveys. A brief, fully automated internet-based pilot study of mental health help-seeking interventions [27] addressed mental health literacy and found that stigma can lead to underreporting and inhibit help-seeking behavior. This led to a model of care and an inference that young elite athletes are particularly unlikely to seek help compared to nonathletes. A self-report internet-based prevalence study [28] established that just under half of the surveyed Australian high-level athletes met the criteria for at least one mental health problem. However, there were engagement issues with insufficient power in the analyses to draw firm conclusions.

A qualitative study transformed from a psychological protective and risk factors screening (an online pilot study) in response to engagement difficulties [29]. An investigation of psychological distress in current and retired Australian elite athletes resulted in novel findings on how accumulative stress manifests into more intense and severe symptoms over time. The vortex of upward and downward spirals was described as the spiraling of states between positive and negative outcomes, a turbulent transitional period where retired elite athletes reported it took 1-3 years for the balance to be restored in their mental health. This highlights the importance of a tailored prevention and early intervention strategy that effectively engages athletes in the screening of psychological protective and risk factors, and tracking that seeks to pinpoint an understanding of how these factors are associated with psychological symptoms, disorders, and abnormal behavior.

Short-form mental health screening instruments were applied with substantial variability in American collegiate athletes, which led to a call for brief validated *electronic* survey instruments that can be interpreted by nonspecialists and have a utility of screening for symptoms or risk factors of specific mental health disorders [9]. An example of such is the Athlete Psychological Strain Questionnaire (APSQ), a brief self-report screening tool by Rice et al [30]. The APSQ was implemented with male elite professional athletes from three national Australian sporting codes (Australian football, cricket, and soccer; n=1007, 78.6% participation rate). Twelve items were developed in the three thematic areas of self-regulation, performance, and external coping to assess for psychological distress and well-being, difficulties with team-based interactions, impaired impulse control and frustration tolerance, worries related to athletic performance and training stress, and transition to life beyond professional athletic pursuits. However, the sensitivity and specificity of the APSQ was designed to identify the behaviors that suggested underlying distress. Modeled on the widely used and validated Kessler 10 and with validity and reliability established from psychometric evaluations [30], the APSQ includes a range of cut-off points reflecting moderate, high, and very high scores, and is the recommended triage screening tool within the International Olympic Committee's Sports Mental Health Assessment Tool [8].

The lack of robust empirical data on mental health symptoms in currently competing athletes [31] led to a cross-sectional, anonymous, online survey of prevalence and correlates of mental health symptoms in high-level athletes (n=810) at the Australian Institute of Sport (AIS). There was a good participation rate (51.7%) with a significantly higher report of *high to very high* psychological distress in high-level athletes compared to general community norms (17.7% vs 9.5%, respectively). These significantly high levels of distress (17.7%) correspond with findings of 16% of meta-analyses from 34 original studies, with 1 at low risk of bias [4]. The extent of other mental health symptoms and disorders ranged from 19% for alcohol misuse to 34% for anxiety and depression for current elite athletes and to 26% for anxiety and depression for former elite athletes. Valid comparisons with the general population were not established by these authors. However, it was surmised that these levels may be slightly higher than in the general population, leading to the call for further research that focuses on the development of reliable and valid screening instruments.

Knowledge of how the mental ill-health sequelae manifests in elite athletes is not yet established in the literature [30], which in turn limits knowledge of impeding factors. The impact of the COVID-19 lockdown upon the mental well-being of high-level athletes was captured in a longitudinal cohort study by Simons et al (unpublished data, 2020), with research support and a sport psychology network to assist athletes [32] from the Queensland Academy of Sport (Australia). A monitoring study (n=15) ensued after online screening of stressors with a smartphone-compatible questionnaire (face-to-face was nullified). Relocation and self-quarantine (coinciding with the implementation of the harshest restrictions of COVID-19) were related to a high risk of adjustment disorder. Mental well-being was negatively impacted during the lockdown period with reports of being *overwhelmed* beyond their sporting careers (into their holistic existence). However, the return to baseline well-being during this period indicated resilience and successful implementation of coping strategies and (self) interventions.

Elite Athlete Mental Health Literature Reviews and Consensus Statements

The generally shared viewpoint of meta-analytic reviews is that elite athletes experience a broadly comparable risk of high-prevalence mental disorders (eg, anxiety and depression) relative to the general population [2]. More specifically, there are comparable rates of depression [4,30,33] as well as anxiety, posttraumatic stress, and sleep disorder [4,30]. Athletes are unique in that there is an existing stigma of mental health and underrecognition of mental illness as well as the possibility of simultaneously having both positive mental health and experiences of mental illness [34]. The range of both athlete-specific and general risk factors associated with mental ill-health in elite athletes led to a comprehensive early intervention framework [10]. A hybrid model of care (face-to-face or by telephone or web-enabled consultations) was used to target the limiting factors of athletes' psychological processes and intervention via the mental health referral network at the AIS. The intent of the AIS early intervention model of care is for the referred practitioner to work one-on-one with the athlete to address their needs in a restoration approach to mental health and functioning [35].

Elite athlete mental health research has evolved as a subset of psychiatry and psychology. Significant recent contributions from sports psychiatry and sport psychology (including clinical psychology) resulted in consensus, expert, and position statements on the individual, cultural, and environmental factors that affect the well-being of the athlete and inform the prevention and treatment of mental health symptoms and disorders [1] (see [3,5,36-40]). A key recommendation is the need for action toward greater consistency in the development, evaluation, and reporting of mental health awareness programs [36]. It was noted that there is "no evidence or consensus-based guidelines for diagnosis and management of mental health symptoms and disorders in elite athletes" [5]. A previous systematic review [41] concluded that a cautious approach is required in determining effective evidence and theory-based intervention programs with unbiased, high-quality methods. Breslin et al [41] outlined blind randomized longitudinal studies as being needed with larger sample sizes and outcomes measured with validated measurement tools that account for behavior change related to mental health literacy. There has been a lack of evidence and discussion of how technology may be applied in the development and evaluation of research or mental health awareness, prevention, and intervention programs.

There are calls for a better understanding of sport as a subculture within society [5,37,42] and for future studies to highlight positive mental health outcomes within athlete subgroups. Investigations into the athletic identity, individual and subgroup character, and cultural traits led to further issues being identified with regard to help-seeking barriers [42]. In particular, Castaldelli-Maia et al [42] found that elite sport culture includes some features that increase the likelihood of athletes with mental health symptoms and disorders. A summary of barriers to athletes seeking mental health treatment was presented including stigma, low mental health literacy, negative past experiences with mental health treatment seeking, busy schedules, and hypermasculinity. Male high-performance athletes were previously reported as 52% less likely than their female counterparts to report mild or more severe depressive symptoms [34]. The proposed link between disclosure of mental health symptoms and disorders, and the stigma of being perceived as *weak* [42] is central to the suggestion that brief antistigma interventions offer promise as an initial step to overcome help-seeking barriers. Further research is required for more focused and tailored interventions for subgroups. For example, it was advised to expand upon gender differences (larger, representative female athlete samples are required) [8].

Acknowledgement of the global challenge of increasing mental health literacy of elite athletes and reducing help-seeking barriers prompted the call for management strategies to address all contributors to mental health symptoms and consider biopsychosocial factors relevant to athletes [5]. Rice et al [30] noted the existence of athlete-centric models of care [3,43]. However, an early intervention framework or model of care is required to support and respond to the mental health needs of this group [10]. The design, implementation, and evaluation of intervention programs may be derived from appropriate theories and models [36]. This implies the merits of reverse engineering. It is not yet established how technology may be applied by functional specialists for the delivery of specific mental health prevention and care outcomes. This gap in the literature highlights the importance of developing and using professionals who can bridge the gap between data science, technology, and organizational strategy.

A key finding for engagement between sporting organizations, elite athletes, and mental health care services highlighted the need to "screen for psychological distress in athletes who may otherwise appear to be well-functioning, to provide timely, optimal treatment" [31]. This is an important consideration as national funding agencies, Olympic committees, federal governments, and sport organizations roll out strategies to combat mental ill-health during COVID-19 [12]. Schinke et al [12] suggested to learn from lessons gained through "autonomy, ingenuity, resilience, life balance, mindfulness" among other skill sets. There is a suggestion to consider cases of *false negatives* and *false positives*, and to capture such distinctions in the screening of the range of psychological function, from positive attributes to severe distress. However, elite athlete mental health prevention and early intervention frameworks have yet to expand beyond a web-based consultative approach to incorporate the potential of digital mental health, especially with regard to the use of technology and human-computer interactions (HCIs).

A theoretical framework needs to recognize the impact of general and athlete-specific risk factors, engage key individuals that may identify and promote athlete mental health, and be adaptable and responsive to varied career stages and mental health states [10]. There is a wealth of theoretical knowledge and insight in the recent literature, yet there is a need for summary, evaluation, and recommendations on ways to effectively engage athletes via technology and make well-designed use of models of care, and adapt or extend them to increase their explanatory power. Furthermore, there is a welcome emphasis on commitment to holistic health and a prevention and early intervention framework. These developments help to redress the balance in favor of mental and physical health being considered together with support for biopsychosocial studies of mental health. It offers a balance with regard to the elite sporting culture, and there is clearly global interest in the results and implications.

Discussion

COVID-19 Pandemic Impact on Mental Health a Catalyst for Change in Screening

There is a need to transform key sources and worldwide accumulated knowledge of mental health systems into insightful findings that can be effectively grasped. The limitations of organizing data scientists and difficulties in upskilling mental health researchers in data science (database systems, programming and data analytics, data mining and management, machine learning, and visual analytics) has hindered the practical steps needed to strengthen mental health systems. The quality of collaboration in elite athlete mental health consensus statements, systematic reviews, and models of care suggests that this field is suitable for the testing of digital mental health. Inroads made thus far, with some sporting organizations having opened up their doors to research via a biopsychosocial approach, lead to the proposal that persistence, data science, and digital mental health methods will result in articulation of how the mental ill-health sequelae manifests in athletes.

A lack of availability and accessibility of acceptable resources and use of counselling during the COVID-19 pandemic [44] points toward evidence of the void that a hybrid model of care (telehealth, digital mental health, and face-to-face) has a great opportunity to contribute to. There is a high potential for data-driven innovation strategies to expand knowledge of how the mental ill-health sequelae manifests in athletes. Machine learning in particular may lead to change more generally if more precise ethics are implemented, resulting in better and faster screening, assessment, and management.

The COVID-19 pandemic has had a profound psychological and social effect [45] with severe psychiatric impacts on the community [46] and increased suicide rates expected [47], especially linked to higher unemployment [48]. Research studies are needed on how mental health consequences can be mitigated during and after the COVID-19 pandemic with a focus on decreasing stress, anxiety, fears, and loneliness in the general population [45]. Sher [45] pressed for urgency but predicted suicidal behavior is likely to be present for a long time and peak later than the actual pandemic. As a result of this catalyst for change, it was suggested to focus on vulnerable subpopulations. Patients at risk for suicide require psychiatric services and monitoring to ensure safety [49]. Schrieber and Culpepper [49] explained that a holistic understanding of the clinical features and course of mental ill-health in subpopulations are prerequisites for the screening, assessment, and management (including prevention and intervention) of people and patients who are vulnerable to develop a clearer picture of the “underlying factors of psychiatric disorders, precipitating events, and ongoing life circumstances” with recommended remedy from “medications, counselling, and involvement of friends, family, and religious/community groups as appropriate.”

There is preliminary evidence among general populations that the COVID-19 pandemic may be associated with psychiatric symptoms that do not necessarily rise to the level of a psychiatric disorder as well as full-blown anxiety disorders, depressive disorders, insomnia disorder, and PTSD [50]. Stein [50] found the pathogenesis of psychiatric symptoms and disorders may include biologic and psychosocial factors, which were noted as possibly increasing the risk of suicidal ideation and behavior because of the array of hardships imposed by the pandemic. This review was especially concerned for patients with COVID-19 and clinicians who treat patients as well as individuals in quarantine. It expresses the overwhelming need to simultaneously screen, track, and treat the “new onset or exacerbation of subsyndromal psychiatric symptoms as well as full-blown psychiatric disorders, including anxiety disorders, depressive disorders, PTSD, or substance use disorders.”

Answering the call for a generalized prevention and intervention tool [50], a first-line screening tool called the Mental Health Quotient (MHQ) was developed to provide for the individual perspective (a profile of mental health challenges and positive well-being) [24]. It screens for concerns and abilities across a results profile while developing a complete picture and unbiased insights of reported symptoms and functions. It is an efficient and user-friendly online assessment of population mental health and well-being. It is suitable for use in primary care and psychiatric clinics to identify *at-risk* individuals and subgroups. It provides information for diagnosis across 10 disorders after a comprehensive review of symptoms across 126 commonly used psychiatric assessment tools. A distinguishing feature is that it is not constrained by the clinical classification systems, the Diagnostic and Statistical Manual of Mental Disorders (DSM) [51] or the International Classification of Diseases (ICD) [52]. The MHQ is useful in identifying people who do not qualify for a particular disorder yet need help because of a large number of severe clinical symptoms (from an anonymous screening that provides a score and full individual report that encourages honest self-reporting). Validity was suggested from demonstration of a close alignment between MHQ scores and the degree to which people meet DSM-5 diagnostic criteria as well as alignment with known epidemiological estimates along with various dimensions including annual prevalence rates of mental health disorders and age and gender difference findings [24]. However, the MHQ is not known to have been tested by others for its validity.

The MHQ is generally useful to target subsyndromal symptoms before it becomes distress or impairment [24]. There is expected to be a greater demand for evidence-based mental health interventions as cultural change prompts more positive perceptions of help-seeking [35]. The athletic culture and dichotomous mental health states and associated biopsychosocial investigations provide a context to compare screening methods and results. The evaluation of the APSQ screening tool [30] was a timely development especially with regard to its validity and reliability [8]. It is a triage screening tool that is relevant worldwide, feeds into an early intervention framework [10], and is connected with the established national Mental Health Referral Network, which presents opportunities for expansion and data capture, informing future education-based initiatives and aims at positively affecting the broader culture around mental health and help-seeking behaviors [35].

A synthesis of a select range of psychological issues and mental health disorders in competitive athletes was intended to assist team physicians and other members of the athletic care network with detection, treatment, and prevention [6]. Screening with Qualtrics (brief psychological survey accessible on a digital device from anywhere) clinically assessed the coping and psychological functioning of nearly 6000 student athletes during the COVID-19 pandemic [44]. Petrie [44] noted small but sizable levels of severe impairment in terms of depressive symptoms, psychological distress, and dissatisfaction with life but a higher incidence of moderate (or subclinical) levels on these measures. Findings between elite [42] and student athletes [44] indicate that these subgroups of athletes face significant challenges with ongoing mental health stigma. Cross-sectional results of high-level (Simons et al, unpublished data, 2020) and student athlete [44] studies suggest a prevalence of depression and distress during the COVID-19 pandemic. However, the study with high-level athletes (Simons et al, unpublished data, 2020) demonstrated a return to baseline levels during the lockdown period, suggesting a possible differentiation for the positive (limited by a small number of participants in the monitoring study).

Broadening the Scope of Elite Athlete Mental Health Symptom and Disorder Interpretation

There have been concerns raised for more than a decade about misdiagnosis arising from current symptom-based DSM and ICD diagnostic criteria for mental disorders, which “are prone to yielding false positives because they ignore the context of symptoms” [53]. However, *false positives* have been widely argued to not be a benign flaw because support and treatment for “disorders” are beneficial to those experiencing emotional problems. Harmful dysfunction analysis has been described for “disordered” classification if human beings fail to satisfy the new social demands [53].

There are *false negative* cases that may arise in elite athletes from dishonest screening responses out of fear of selection or losing a position and, thus, presenting difficulties for detection by functional specialists. There are also different interpretations of the diagnostic criteria for mental disorders in elite athletes. In consideration of the biopsychosocial context in the etiology, diagnosis, treatment, and prevention of disorders, the model of maladjustment [54] reconceptualized the overtrained state of athletes as an adjustment disorder and, thus, presented a rationale that it should be treated from the perspective of this diagnosis. In contrast, longitudinal cohort research with high-level athletes found a subgroup that were *at risk* (subsyndromal) for adjustment disorder because of a range of stressors (injury, relocation, being on tour and long periods away from home, and the COVID-19 lockdown impact on training and competition; Simons et al, unpublished data, 2020). Therefore, the scope for targeting subsyndromal conditions and psychological symptoms and disorders in athletes can be altered by various measures, diagnostic criteria, and interpretations.

Psychological Risk, Protective Factors, and Outcome Measures for Tracking

As an example of formative research for tracking, a qualitative study of current and retired elite athletes’ mental health [29] held semistructured in-depth interviews with discussion around their perceptions and expectations of their self, with a focus on achievements and challenges, stressors inside and outside of sporting careers, injuries, transition to life after sport, psychological distress, suicidal ideation, perfectionism, life satisfaction, support, coping strategies, and substance use. Outcome measures for quantitative studies with athletes include knowledge of mental health (ie, disorder and symptom recognition) or behavior regarding mental health (intended or actual help-seeking), mental health competencies (eg, mindfulness and coping) or specific mental health (eg, anxiety, depressive symptoms, and positive affect), and well-being (eg, subjective and psychological well-being domains, and life satisfaction) [41].

Future research is required to conceptualize and gain a better understanding of the combination of multiple stressors and coping resources in athletes. It is also important to verify which factors aid and impede positive adjustment to stressors from a holistic perspective. However, it should not be assumed that there is a need for increased monitoring, surveillance, and medication to enact control (at a cost to recovery). There is a need to promote a balance with a risk assessment and management framework so that the subjective experiences of people with mental health problems are at the forefront of prevention and intervention [55].

Sporting Culture and Athletic Identity

Culture provides socially constructed myths about natural phenomena, resulting in reshaped systems of belief, which are digested to become part of a worldview and thus influences interpretation of natural phenomena [56]. It is necessary to observe subcultures in athletes to effectively address help-seeking barriers and stigma [42]. For example, athletes *play for each other* in team sports and senior athletes take a leadership role, in addition to coaches and staff, to materialize successful performance. The interconnectedness of sickness and health, and the global effects of COVID-19 means that athletes need each other to adhere to protocols to effectively participate in sports. They also require support and a platform to build a winning culture. However, the “sporting bubble” may distort athletes’ perception of the world (including that mental ill-health is a sign of weakness), which affects their ability to live in it. Therefore, it is important for researchers to gradually peel back the layers acknowledging this group’s attributes such as mental toughness and resilience. Elite athletes may not admit up front to their vulnerabilities such as the effect of the accumulation of pressure and psychological, emotional, professional, and physical stress. Partnered with injuries and recovery, life changes, selection worries, or retirement, their depth of physical and mental strength may be tested such that a breaking point is reached.

Athletes who possess a strong and exclusive level of athletic identity take longer to adapt to their postsport life and are more prone to experiencing psychosocial and career identity difficulties later in life [57]. Athletes who face an involuntary retirement continue to be at the greatest risk of experiencing adjustment issues [58]. The transition out of sport can be a challenging time for athletes [59,60], so preparations for such time should be made well in advance in conjunction with an *exit health examination* [4]. The instability may be inhibited further if they are not in touch with their own feelings and thoughts or these are rendered unacceptable by parents, coaches and staff, peers, or former athletes. The digital age and generational differences also call for innovative ways of connecting athletes with their networks and traditions. Although, this may not always be positive, as high achievers tend to overlay their predisposition to succeed in other areas of their life such that, when the “sporting bubble” has burst, negative coping strategies may be applied with similar intensity (eg, substance misuse).

Digital Mental Health Solutions

Although not a panacea to mental ill-health issues in athletes or general populations, digital technology has garnered increasing interest and evidence for mental health improvement, especially with younger people [61]. Bevan Jones et al [61] cited a lack of literature to inform the co-design of digital mental health technologies. However, it was established that the involvement of stakeholders is important throughout the life and research cycle of the program. The biggest challenges are the changing face of technology, methods of engaging with diversity in the user group, and evaluation of the co-design process and its impact on the technology.

Risk assessment should consider the dichotomy of digital technology use. It can aid with information and support for those with mental health problems, but it can also be of detriment because of internet-related risk behavior (digital risk) such as bullying or prosuicide websites [62]. Therefore, Aref-Adib et al [62] suggested training and standardized risk pro forma to promote best practice digital risk awareness in mental health care general risk assessments. Digital mental health projects may benefit from a risk management approach that focuses on collaboration, ideation, and innovation, and deals with uncertainty with augmented responses in obtaining objectives, monitoring along the way, communicating, and better integrating and reporting.

Ecological perspectives are important to gather insight for educational strategies to increase mental health literacy [1]. The ecological design perspective provides a conceptual framework within which to investigate more complex interactions between persons and environments [63]. As new ways of being connected to each other and artificial intelligence are becoming more possible and needed, it is important to consider the central setting that digital technology already has in the broader institutional and community contexts. The digital infrastructure is expanding along with its own influence.

The benefits of digital mental health include broad accessibility and choices with free or low-cost availability. It may also be an introduction to psychological intervention that helps participants to overcome stigma or embarrassment. Before technology and changing demand enabled digital mental health to grow, the health industry was dealing with a range of other issues such as inequitable outcomes; unaffordable growth in expenditure; and varied outcomes in wellness, prevention, and intervention. Although there are other pertinent issues, it is important to future proof with data science (programming languages, visualization tools, relevant data platforms). There is also a requirement for practical skills and the knowledge and ability to create complex ecosystems of smart devices as well as the use of the data collected by these devices. However, the *design to innovate* agenda needs to draw from lessons learned and themes such as *the struggle to use digital tools* and *how to overcome difficulties in connecting with them*. A different approach is required to better turn data into information, knowledge, and insight. The design ecosystem needs to give attention to the layout of key aspects of an architecturally designed digital landscape including problem definition, strategy, process and solution, and improvements. Advisory knowledge hubs and dashboards should provide solutions and opportunities to upskill and build capability.

Up-to-date, peer-reviewed online algorithms are useful, for example, in differentiating intentional self-injurious thoughts and behavior into suicidal or nonsuicidal categories [49]. A next frontier in technology is designing artificial intelligence to deliver a range of mental health care services according to algorithms. Artificial intelligence–powered solutions currently include predictive analytics, machine learning, natural language processing or generation, voice recognition and response, virtual personal assistants and chatbots, and diagnosis and recommendation engines. Large-scale and rigorous studies with artificial intelligence–delivered mental health care are still in a preliminary stage [64]. However, Fiske et al [64] noted more evidence is required with regard to patient acceptance and contingent treatment outcomes of embodied artificial intelligence applications in mental health. The social and ethical implications of artificial intelligence remain pertinent issues. Yet, the demand and toll of COVID-19 upon health care workers [50] and the inevitability of further epidemics or disruptions to life calls for virtual mental health care in a practical complementary approach.

Virtual reality (VR) was found to outperform Skype as a therapeutic tool in remote therapy in terms of perceived realism of the session and the degree of presence (face-to-face but in a virtual environment) [65]. The VR concept will be adapted to pioneer digital mental health solutions, including artificial intelligence and deep learning algorithms [66-69], offering promise as a supplementary approach to boosting mental skills and improving the diagnosis and treatment of patients with mental health conditions. A pilot study found VR may benefit the therapeutic alliance [70], especially for people with a psychotic disorder with regard to emotion perception [71]. It was noted that there are limited studies in this field thus far.

There are various mental health apps with links to digital interventions (eg, cognitive behavior therapy and mindfulness exercises). The potential for smartphones to become a real-time diagnostic tool is not yet realized, although the emerging 5G mobile network technology will enable better connection of devices, which will aid this. In the future, 6G may enable the creation of virtual figures or replicas from application of immersive extended reality (XR), high-fidelity mobile holograms, and digital replicas. These HCIs may be programmed for comprehension of algorithms for mental health issues including the intricacies of psychology and psychiatry, somatization, cognition, mind-body connection, behavior, attributes, drive, motivation, attitude, mood, socialization, and personality. However, this is not an exhaustive list. It will be necessary to define and describe HCI and how it impacts user experience design for XR.

Digital solutions with machine learning and artificial intelligence at their core have been developed as flexible tools for identifying patterns in data (via deep learning algorithms) with powerful diagnostic and treatment capabilities to serve large populations at reasonable costs and without human prejudice [72]. The how and why of artificial intelligence delivering virtual mental health care requires an understanding of the Internet of Things (IoT)—the connection of machines for information exchange without human intervention.

The self-improvement capabilities of machine learning and artificial intelligence as new data becomes available is an advantage that assists in early case detection and tracking. However, there is potential for harm and threats arising from the increasing adoption and power [73]. It was recommended by these authors for information literacy and community building in deep learning as tools for getting ahead of deep fakes. To give contexts to the risks, Caldwell et al [73] referred to useful systematic reviews [74], overview of strategic policy in the next 5 years [75], and the cybersecurity risks associated with the interconnectedness of everyday devices (IoT) [76]. A user-friendly description of the technological background of artificial intelligence was presented to instill an understanding of the associated criminal potential and with the intention of reducing potential harm from it [73]. To understand the issue of deep fakes, these authors defined artificial intelligence, machine learning, supervised and unsupervised learning, training data, deep learning, reinforcement learning, active learning, natural language processing, adversarial perturbation, and criminal potential. Patient privacy and the potential for machine learning discrimination requires effective measures to be in place. An emerging example is federated learning whereby an algorithm is trained to function across multiple decentralized devices or servers holding local data samples without exchanging them.

There is an opportunity to innovate developing technologies with the delivery of consistent and cybersecure screening, tracking, and well-being programs in an engaging format that helps users to overcome stigmatization as well as increase mental health literacy, screening honesty, and help-seeking behaviors. In effect, the process of measuring and tracking population mental health is working toward relieving the burden of mental health care "by facilitating the development of relevant and effective interventions and policies before symptoms escalate to clinical levels" as well as "manage and improve the lives and well-being of all people, and not just those with a clinical disorder" [24]. As researchers reimagine scientific or health processes, digital enablement offers promise as a collaborative process that can quickly adapt from client and consumer touchpoint reviews and lessons learned to realizing the total value of its ownership. To grow and evolve a digital mental health ecosystem, it is important to document, store, and publish successful innovation activities; seek ways to increase the rate of adoption; and help shape further opportunities and improvements.

Recommendations

The following novel recommendations for the design and building of a digital mental health platform and associated screening and tracking technologies and services are intended as a guide for consultation and a consideration for construction. A standard of care and fitness of purpose should be considered with expert and legal advice to ensure a prototype is factual, experienced, and applicable to industry standards before entering a design or construction contract.

Athlete-Specific Mental Health Screening Questionnaire

A worldwide-relevant and smartphone compatible athlete-specific mental health screening questionnaire based on the APSQ [30] is recommended to be deployed (weekly), partnered with real-time automation and machine learning models. It should be presented as part of a screening process associated with a well-being framework with consideration of stress and adjustments. A complementary follow-up involving a tracking phase should be implemented weekly with the identified *at-risk* and vulnerable subgroups. It should be designed to improve athlete buy-in with honest feedback and self-evaluation. Referrals to licensed professionals are recommended where clinical levels of psychological distress are identified. The athlete is then to be treated as a patient depending on the symptoms and appropriate treatment recommended by algorithms, such as cognitive behavior therapy or assistance with coping and resilience skills.

The APSQ should be applied as an education concept for the development and validation of subpopulation-specific screening tools. In seeking a consistent and consolidated knowledge management strategy, a hybrid model of care is recommended for the prevention and intervention of stress and adjustment issues as well as psychological disorders. The promotion of mental health improvement goals should be recognized as an investment for athlete well-being programs in addressing issues around engagement, identity, privacy, and confidentiality.

Championing of Mental Health Awareness and Antistigma Interventions

Concise guidelines, consensus statements, and brief antistigma interventions are recommended for consolidation and consistency in mental health awareness. It is recommended to highlight the range of mental health states with elite athlete and mental health champions to raise mental health awareness in society. It is recommended to specifically target peculiar subgroups of athletes with more focused prevention and tailored interventions that recognize varied career stages and mental health states.

Pivot From Existing Subpopulation Research to Quickly Respond to Disruptive Global Events

Further research is recommended to expand the finding that high-level athletes are at increased vulnerability to mental ill-health, yet there are positive adjustments and resilience in response to *overwhelming stress*, resulting in a subsyndromal condition (Simons et al, unpublished data, 2020). It is recommended to pivot from existing studies such as with college student, high-level, and elite athletes to investigate and articulate "how behaviors and mental health change and interact in the face of monumental adverse global events" [19]. Verification of behavior research is required such as a study with the Circadian Rhythm for Mood app, which found that a wearable device is effective in the prevention, intervention, and improvement of mood disorders [77].

In targeting *at-risk* subpopulations such as athletes for early intervention, it is recommended to implement brief, valid, and reliable smartphone-compatible online surveys (with consideration of the offline options such as paper and text messaging alternatives for the underserved). It should be tailored to the relevant psychological risk and protective factors. Tracking (monitoring of symptoms, in particular where there is an indication of high stress and adjustment issues but not psychological distress or disorder) is recommended to be run by unlicensed professionals such as academic researchers or from other areas of psychology, psychiatry, or mental health that do not require direct clinical contact with patients. A rationale for tracking is that descriptive and experimental research is required to increase positive engagement with *at-risk* and vulnerable subpopulations, support mental health care services, and counter the psychological impact of disruptive global events such as the COVID-19 pandemic.

Integration of Subpopulation and General Population Research With Algorithms

Valid and reliable subpopulation research should be compared with findings from the general population. Blind randomized longitudinal studies with large-scale data collection via MHQ screening in general populations may expand its validation. Evidence- or consensus-based guidelines for diagnosis and management of mental health symptoms and disorders are recommended to evolve into standardized, up-to-date online algorithms. Real-time automation with machine learning models is recommended to apply the MHQ to test its effectiveness in tracking compared with results from clinical psychologists and psychiatrists as well as unlicensed professionals, who may contribute more of a supporting role. The intent is to inform effective and efficient pathways of mental health care with valid, reliable, and replicable identification of at-risk individuals and subgroups, and assess the impact of any support and wellness programs.

The MHQ provides for the geographical, cultural, and experiential factors of the general population, and covers the complete range of clinical mental health symptoms as well as positive mental assets [24]. It is recommended for general population self-screening because it encompasses a fluid state of how these symptoms and assets interact based on an extensive review of the way mental health is assessed in clinical and research fields [78]. The MHQ is an epidemiological mental health assessment tool that provides personalized insight into an individual's mental health partnered with feedback that is generated based on the individual scoring profile. The delivery of solutions for influencing positive mental health is recommended to consider the mixed linear model of smartphone mobile sensing and self-reported mental health questions [19], and a hybrid model of care [10,20-22] as part of a general guide.

Development and Evaluation of Digital Mental Health Platforms

A road map [25] and integrated blueprint [26] for platform-based solutions are urgently needed to strengthen mental health systems with consolidation, consensus, and concise guidelines. There are opportunities for growth in digital mental health risk surveillance (screening and tracking) extending from an honest and transparent platform, which promotes trust and confidentiality. It is recommended to progress opportunities in digital mental health (especially with platforms) while acknowledging that there are constraints and rigorous evaluation is required [26]. It is recognized that these emerging developments are "not explicitly covered by the existing guidelines" [21]. However, it is recommended that a certification framework, applicable industry standards, and expert advice (eg, a road map, integrated blueprint, and innovation ecosystem for digital mental health services) guide entrepreneurial activities. In the absence of a government-endorsed certification framework, it is recommended to consult with the terms and conditions of a reliable example that has been evaluated to an international standard. It is also important to ensure there is a synergy of activities for mental health care with efforts in outreach and dissemination of peer-reviewed literature as well as development of a skilled, digitally capable workforce.

A new type of patient-consumer is fueling remote screening and tracking's rapid growth. Although there is a wide range of choices available for digital mental health, there is an absence of a useful and reliable athlete-specific digital mental health platform. If a prototype is developed, it should serve as an example to other subpopulations for tailored prevention and intervention. Mental health resources and services should be sourced from trusted organizations, including links to evidence-based apps and programs, forums, phone chats, emails, and websites.

Digital mental health requires consultative leadership in advising, designing, and delivering the platforms and infrastructure for a transformation in client and consumer delivery. An organization's track record is important for building or adopting a brand for an individual. It is recommended to not be reliant on a small, distant vendor for a core platform, which could be easily disrupted and cause a negative flow-on effect. Therefore, a well-established company may help build trust in the adoption of digital mental health technologies. A future-ready technology and health company is recommended for an end-to-end service, offering and drawing on depth and capability to deliver on promises.

Web-Based and Smartphone Mental Health App Considerations

Web-based and smartphone mental health apps are recommended to be presented in a visually engaging way with an appropriate balance of human characteristics in the graphics (representative of diversity and ethnicity). Brief scalable survey questions, which can be skipped or with options to go to the next or previous one, may be supported with an open-ended answer box that prompts the user to explain their thoughts and feelings after a score is provided. An information package should include the conditions and terms of engagement, and links to the app, app help, translated information, resources, what it is for, how it works, frequently asked questions (FAQs), deleting the app, and privacy. There should also be information provided about data collection, processing, and storage, with details on the purpose, data source and access, storage period and location, contact information, and complaints.

Human-Computer Interactions in a Hybrid Model of Care

The digital therapeutic alliance (DTA) as a concept [79-83] suggests that there is a future for HCI, but there is a need for testing of emerging interactive techniques. More research is required to support acceptability, feasibility, and safety around digital interventions [84]. The efficacy of a hybrid model of care requires evaluation of its impact on the therapeutic alliance, clinical and social outcomes, cost-effectiveness, and engagement [23]. A therapeutic alliance may be refined in digital interventions with its own unique features (although further investigation is required to affirm this) [69].

Some mental health functional specialists may be required to lead, upskill, and strategize in data science and artificial intelligence. It is important to understand the potential and limitations of smart devices and apply a design-thinking approach as a way to come up with creative solutions. The design, implementation, and analysis of smart systems that gather data can inform and drive more effective and insightful mental health policies and strategies, for example, Lederman et al’s [85] suggestions to aim for adaptability of multifaceted tools through *concordance in HCI*. This concept takes account of the self-determination theory and is useful for design improvements from peer and moderator support as well as automated feedback [85]. Future developments with digital interventions are recommended to measure HCI for engagement and adherence effectiveness as well as individual client and consumer satisfaction with the DTA experience, including feedback on its functionality, trustworthiness, and expertise as well as confidence in its recommendations.

It is recommended for longitudinal research of screening and tracking with real-time automation and machine learning to determine whether there is improvement in engagement levels, ongoing participation rates, mental health literacy, and help-seeking behaviors. This strategy is recommended to be tailored toward a hybrid model of care including self-help, face-to-face, and digital interventions with *at-risk* and vulnerable subpopulations.

Artificial Intelligence, Deep Learning Algorithms, VR, and Digital Interventions

Artificial intelligence, object-oriented, and information systems as well as deep learning algorithms are increasingly of interest for therapeutic outcomes from digital interventions. It is recommended for an experimental rationale to better understand the *bond* within the human-smartphone connection. The IoT structure is recommended for database, statistical, and communication systems, including edge devices, data mining, and visualization as well as security and cloud computing.

The ecological perspective is suitable to be adopted with the co-design of digital mental health technologies to describe how the system or product works within its external environment. For example, VR modules are used for the treatment of multiple conditions from addiction to eating disorders, depression and anxiety, stress, and trauma as well as for immersion and repetition of scenes that may be useful for sport performance and mindfulness. It is recommended to expand research with VR and emerging technologies to produce specialized immersive interactive experiences for *at-risk* and vulnerable subpopulations, and to quantify the effect of creative problem solving and critical reflection upon stress and adjustment to see if there is a relationship with subsyndromal conditions or symptoms and disorders.

In determining a DTA for virtual clinics, it is recommended to consult up-to-date algorithms especially for evidence-based psychological protective and risk factors. A subset of HCI, affective computing, is recommended to be included in the co-design for user state–tailored response and an appealing balance of artificial intelligence and human characteristics such as traits, emotions, and intentions.

Conclusions

The novel finding that the COVID-19 pandemic had a dichotomous impact on high-level athletes’ mental well-being, with a period of *overwhelming stress* followed by a return to baseline well-being during lockdown (Simons et al, unpublished data, 2020), suggests a peculiarity in an *at-risk* but resilient subgroup of a subpopulation who are resourceful in achieving positive adjustments. A next step is validation with elite athletes. Brief psychological strain screening tools such as the APSQ should extend with stress and adjustment-focused methods [17,18]. The unrealized potential of this early intervention strategy should also be more generally explored (eg, MHQ) for how it may assist the management of mental health (from mental ill-health to positive functioning).

Future research should draw comparisons between athletes and the general population with regard to broadening the scope of mental health, especially symptom and disorder interpretation. There is a high potential for reliable and evaluated digital mental health platforms to inform associated prevention and early intervention strategies. Collaboration between humans and machines will be critical to the innovation of mental health care in the future. Abilities or willingness to embrace and adopt machine learning (with real-time automation) and artificial intelligence in a hybrid model of care should be guided by the right relevance in a road map and integrated blueprint to drive valid and reliable studies and address concerns over the potential for threats or harm. Studies adopting inclusive design or human-centric design principles should demonstrate how embracing digital mental health (especially deriving from digital platforms and algorithms from machine learning and artificial intelligence) is a critical success factor in a post–COVID-19 world.

Sports have a great potential to become more than a fertile platform for awareness of mental health issues and education for seeking help. Athletes are a suitable subpopulation for which to implement digital mental health solutions for stress and adjustment as a complement to prevention and intervention. Validation of the dichotomy of athlete mental health problems and resilience (in or outside of a pandemic-affected situation) offers a dual opportunity to advance digital mental health screening and tracking, and enact faster and better treatment algorithms for *at-risk* and vulnerable subpopulations.

Acknowledgments

The authors acknowledge Simon Rice, who provided suggestions on this article, and the Queensland Academy of Sport, who supported associated research at the University of the Sunshine Coast [17,18, Simons et al, unpublished data, 2020] under Grant 99.

Conflicts of Interest

None declared.

References

1. Gorczynski P, Currie A, Gibson K, Gouttebarge V, Hainline B, Castaldelli-Maia JM, et al. Developing mental health literacy and cultural competence in elite sport. J Appl Sport Psychol 2020 Feb 06:1-15. [doi: 10.1080/10413200.2020.1720045]
2. Rice S, Purcell R, De Silva S, Mawren D, McGorry PD, Parker AG. The mental health of elite athletes: a narrative systematic review. Sports Med 2016 Sep;46(9):1333-1353 [FREE Full text] [doi: 10.1007/s40279-016-0492-2] [Medline: 26896951]
3. Moesch K, Kenttä G, Kleinert J, Quignon-Fleuret C, Cecil S, Bertollo M. FEPSAC position statement: mental health disorders in elite athletes and models of service provision. Psychol Sport Exerc 2018 Sep;38:61-71. [doi: 10.1016/j.psychsport.2018.05.013]
4. Gouttebarge V, Castaldelli-Maia JM, Gorczynski P, Hainline B, Hitchcock ME, Kerkhoffs GM, et al. Occurrence of mental health symptoms and disorders in current and former elite athletes: a systematic review and meta-analysis. Br J Sports Med 2019 Jun;53(11):700-706 [FREE Full text] [doi: 10.1136/bjsports-2019-100671] [Medline: 31097451]
5. Reardon C, Hainline B, Aron CM, Baron D, Baum AL, Bindra A, et al. Mental health in elite athletes: International Olympic Committee consensus statement (2019). Br J Sports Med 2019 Jun;53(11):667-699. [doi: 10.1136/bjsports-2019-100715] [Medline: 31097450]
6. Chang C, Putukian M, Aerni G, Diamond A, Hong G, Ingram Y, et al. Mental health issues and psychological factors in athletes: detection, management, effect on performance and prevention: American Medical Society for Sports Medicine Position Statement-Executive Summary. Br J Sports Med 2020 Feb;54(4):216-220. [doi: 10.1136/bjsports-2019-101583] [Medline: 31810972]
7. Donohue B, Galante M, Maietta J, Lee B, Paul N, Perry JE, et al. Empirical development of a screening method to assist mental health referrals in collegiate athletes. J Clin Sport Psychol 2019;13(4):561-579. [doi: 10.1123/jcsp.2018-0070]
8. Rice S, Olive L, Gouttebarge V, Parker AG, Clifton P, Harcourt P, et al. Mental health screening: severity and cut-off point sensitivity of the Athlete Psychological Strain Questionnaire in male and female elite athletes. BMJ Open Sport Exerc Med 2020;6(1):e000712. [doi: 10.1136/bmjsem-2019-000712] [Medline: 32231792]
9. Kroshus E. Variability in institutional screening practices related to collegiate student-athlete mental health. J Athl Train 2016 May;51(5):389-397 [FREE Full text] [doi: 10.4085/1062-6050-51.5.07] [Medline: 27111587]
10. Purcell R, Gwyther K, Rice SM. Mental health in elite athletes: increased awareness requires an early intervention framework to respond to athlete needs. Sports Med Open 2019 Nov 28;5(1):46 [FREE Full text] [doi: 10.1186/s40798-019-0220-1] [Medline: 31781988]
11. Torous J, Haim A. Dichotomies in the development and implementation of digital mental health tools. Psychiatr Serv 2018 Dec 01;69(12):1204-1206 [FREE Full text] [doi: 10.1176/appi.ps.201800193] [Medline: 30256182]
12. Schinke R, Papaioannou A, Henriksen K, Si G, Zhang L, Haberl P. Sport psychology services to high performance athletes during COVID-19. Int J Sport Exerc Psychol 2020 Apr 22;18(3):269-272. [doi: 10.1080/1612197x.2020.1754616]
13. Mehrsafar A, Gazerani P, Moghadam Zadeh A, Jaenes Sánchez JC. Addressing potential impact of COVID-19 pandemic on physical and mental health of elite athletes. Brain Behav Immun 2020 Jul;87:147-148 [FREE Full text] [doi: 10.1016/j.bbi.2020.05.011] [Medline: 32387513]
14. Kola L. Global mental health and COVID-19. Lancet Psychiatry 2020 Aug;7(8):655-657 [FREE Full text] [doi: 10.1016/s2215-0366(20)30235-2]

15. Torjesen I. Covid-19: mental health services must be boosted to deal with "tsunami" of cases after lockdown. BMJ 2020 May 15;369:m1994. [doi: 10.1136/bmj.m1994] [Medline: 32417756]
16. Keshavan M. Pandemics and psychiatry: repositioning research in context of COVID-19. Asian J Psychiatr 2020 Jun;51:102159 [FREE Full text] [doi: 10.1016/j.ajp.2020.102159] [Medline: 32442894]
17. Simons C, Martin L, Balcombe L, Dunn PK. Amendment of a screening tool for adjustment disorder in the athlete population. 2019 Presented at: 15th European Congress of Sport and Exercise Psychology; July 2019; Münster, Germany.
18. Simons C, Martin L, Balcombe L, Dunn PK. Stressor incidence and adjustment disorder prevalence among elite athletes: a pilot study. 2019 Presented at: 15th European Congress of Sport and Exercise Psychology; July 2019; Münster, Germany.
19. Huckins J, daSilva AW, Wang W, Hedlund E, Rogers C, Nepal SK, et al. Mental health and behavior of college students during the early phases of the COVID-19 pandemic: longitudinal smartphone and ecological momentary assessment study. J Med Internet Res 2020 Jun 17;22(6):e20185 [FREE Full text] [doi: 10.2196/20185] [Medline: 32519963]
20. Chan S, 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]
21. Smith K, Ostinelli E, Macdonald O, Cipriani A. COVID-19 and telepsychiatry: development of evidence-based guidance for clinicians. JMIR Ment Health 2020 Aug 28;7(8):e21108 [FREE Full text] [doi: 10.2196/21108] [Medline: 32658857]
22. Valentine L, McEnery C, Bell I, O'Sullivan S, Pryor I, Gleeson J, et al. Blended digital and face-to-face care for first-episode psychosis treatment in young people: qualitative study. JMIR Ment Health 2020 Jul 28;7(7):e18990 [FREE Full text] [doi: 10.2196/18990] [Medline: 32720904]
23. Becker T, Torous JB. Recent developments in digital mental health interventions for college and university students. Curr Treatment Options Psychiatry 2019 Jun 14;6(3):210-220. [doi: 10.1007/s40501-019-00178-8]
24. Newson JJ, Thiagarajan TC. Assessment of population well-being with the Mental Health Quotient (MHQ): development and usability study. JMIR Ment Health 2020 Jul 20;7(7):e17935 [FREE Full text] [doi: 10.2196/17935] [Medline: 32706730]
25. Maulik P, Thornicroft G, Saxena S. Roadmap to strengthen global mental health systems to tackle the impact of the COVID-19 pandemic. Int J Ment Health Syst 2020;14:57 [FREE Full text] [doi: 10.1186/s13033-020-00393-4] [Medline: 32742305]
26. Balcombe L, De Leo D. An integrated blueprint for digital mental health services amidst COVID-19. JMIR Ment Health 2020 Jul 22;7(7):e21718 [FREE Full text] [doi: 10.2196/21718] [Medline: 32668402]
27. Gulliver A, Griffiths KM, Christensen H. Barriers and facilitators to mental health help-seeking for young elite athletes: a qualitative study. BMC Psychiatry 2012 Sep 26;12:157 [FREE Full text] [doi: 10.1186/1471-244X-12-157] [Medline: 23009161]
28. Gulliver A, Griffiths KM, Mackinnon A, Batterham PJ, Stanimirovic R. The mental health of Australian elite athletes. J Sci Med Sport 2015 May;18(3):255-261. [doi: 10.1016/j.jsams.2014.04.006] [Medline: 24882147]
29. Balcombe L. Responses of elite athletes to the negative consequences of turning points during and after their sport careers. Griffith University. 2015. URL: https://research-repository.griffith.edu.au/bitstream/handle/10072/366240/Balcombe_2015_02Thesis.pdf [accessed 2020-07-13]
30. Rice SM, Parker AG, Mawren D, Clifton P, Harcourt P, Lloyd M, et al. Preliminary psychometric validation of a brief screening tool for athlete mental health among male elite athletes: the Athlete Psychological Strain Questionnaire. Int J Sport Exerc Psychol 2019 May 03;18(6):850-865. [doi: 10.1080/1612197x.2019.1611900]
31. Purcell R, Rice S, Butterworth M, Clements M. Rates and correlates of mental health symptoms in currently competing elite athletes from the Australian National High-Performance Sports System. Sports Med 2020 Sep;50(9):1683-1694. [doi: 10.1007/s40279-020-01266-z] [Medline: 32026315]
32. Performance health and wellbeing at the Queensland Academy of Sport. Queensland Government. 2020. URL: https://www.qld.gov.au/recreation/sports/academy/services/performance-health
33. Gorczynski P, Coyle M, Gibson K. Depressive symptoms in high-performance athletes and non-athletes: a comparative meta-analysis. Br J Sports Med 2017 Sep;51(18):1348-1354. [doi: 10.1136/bjsports-2016-096455] [Medline: 28254747]
34. Uphill M, Sly D, Swain J. From mental health to mental wealth in athletes: looking back and moving forward. Front Psychol 2016;7:935. [doi: 10.3389/fpsyg.2016.00935] [Medline: 27445903]
35. Rice S, Butterworth M, Clements M, Josifovski D, Arnold S, Schwab C, et al. Development and implementation of the National Mental Health Referral Network for elite athletes: a case study of the Australian Institute of Sport. Case Stud Sport Exerc Psychol 2020;4(S1):S1-27-S1-35. [doi: 10.1123/cssep.2019-0016]
36. Breslin G, Smith A, Donohue B, Donnelly P, Shannon S, Haughey TJ, et al. International consensus statement on the psychosocial and policy-related approaches to mental health awareness programmes in sport. BMJ Open Sport Exerc Med 2019;5(1):e000585. [doi: 10.1136/bmjsem-2019-000585] [Medline: 31673406]
37. Gorczynski P, Gibson K, Thelwell R, Papathomas A, Harwood C, Kinnafick F. The BASES expert statement on mental health literacy in elite sport. The British Association of Sport and Exercise Sciences. 2019. URL: https://www.bases.org.uk/imgs/7879_bas_expert_statement__pages_735.pdf [accessed 2020-07-07]

38. Henriksen K, Schinke R, Moesch K, McCann S, Parham W, Larsen CH, et al. Consensus statement on improving the mental health of high performance athletes. Int J Sport Exerc Psychol 2019 Jan 31;18(5):553-560. [doi: 10.1080/1612197x.2019.1570473]
39. Schinke R, Stambulova NB, Si G, Moore Z. International society of sport psychology position stand: athletes' mental health, performance, and development. Int J Sport Exerc Psychol 2017 Apr 11;16(6):622-639. [doi: 10.1080/1612197x.2017.1295557]
40. Van Slingerland KJ, Durand-Bush N, Bradley L, Goldfield G, Archambault R, Smith D, et al. Canadian Centre for Mental Health and Sport (CCMHS) position statement: principles of mental health in competitive and high-performance sport. Clin J Sport Med 2019 May;29(3):173-180. [doi: 10.1097/JSM.0000000000000665] [Medline: 31033609]
41. Breslin G, Shannon S, Haughey T, Donnelly P, Leavey G. A systematic review of interventions to increase awareness of mental health and well-being in athletes, coaches and officials. Syst Rev 2017 Aug 31;6(1):177 [FREE Full text] [doi: 10.1186/s13643-017-0568-6] [Medline: 28859666]
42. Castaldelli-Maia J, Gallinaro J, Falcão RS, Gouttebarge V, Hitchcock ME, Hainline B, et al. Mental health symptoms and disorders in elite athletes: a systematic review on cultural influencers and barriers to athletes seeking treatment. Br J Sports Med 2019 Jun;53(11):707-721. [doi: 10.1136/bjsports-2019-100710] [Medline: 31092400]
43. Reardon C, Factor RM. Sport psychiatry: a systematic review of diagnosis and medical treatment of mental illness in athletes. Sports Med 2010 Nov 01;40(11):961-980. [doi: 10.2165/11536580-000000000-00000] [Medline: 20942511]
44. Petrie T. Impact of COVID-19 on college student athletes' health, performance, and psychological well-being: an executive summary on baseline data. University of North Texas Center for Sport Psychology. 2020. URL: https://sportpsych.unt.edu/sites/default/files/covid_executive_summary.final2.july8.2020.pdf [accessed 2020-07-13]
45. Sher L. The impact of the COVID-19 pandemic on suicide rates. QJM 2020 Oct 01;113(10):707-712. [doi: 10.1093/qjmed/hcaa202] [Medline: 32539153]
46. Gautam M, Thakrar A, Akinyemi E, Mahr G. Current and future challenges in the delivery of mental healthcare during COVID-19. SN Compr Clin Med 2020 Jun 11:1-6. [doi: 10.1007/s42399-020-00348-3] [Medline: 32838140]
47. Devitt P. Can we expect an increased suicide rate due to Covid-19? Ir J Psychol Med 2020 May 21:1-5. [doi: 10.1017/ipm.2020.46] [Medline: 32434598]
48. McIntyre R, Lee Y. Projected increases in suicide in Canada as a consequence of COVID-19. Psychiatry Res 2020 Aug;290:113104. [doi: 10.1016/j.psychres.2020.113104] [Medline: 32460184]
49. Schrieber J, Culpepper L. Suicidal ideation and behavior in adults. UpToDate. 2020. URL: https://www.uptodate.com/contents/suicidal-ideation-and-behavior-in-adults [accessed 2020-07-18]
50. Stein MB. Coronavirus disease 2019 (COVID-19): psychiatric illness. UpToDate. 2020. URL: https://www.uptodate.com/contents/coronavirus-disease-2019-covid-19-psychiatric-illness [accessed 2020-07-18]
51. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Association; 2013.
52. International Statistical Classification of Diseases and Related Health Problems (ICD). World Health Organization. 2018. URL: http://www.who.int/classifications/icd/en/ [accessed 2020-07-21]
53. Wakefield JC. False positives in psychiatric diagnosis: implications for human freedom. Theor Med Bioeth 2010 Feb;31(1):5-17. [doi: 10.1007/s11017-010-9132-2] [Medline: 20232254]
54. Jones C, Tenenbaum G. Adjustment disorder: a new way of conceptualizing the overtraining syndrome. Int Rev Sport Exerc Psychol 2009 Sep;2(2):181-197. [doi: 10.1080/17509840903110962]
55. Felton A, Repper J, Avis M. The construction of people with mental health problems as risk objects: findings of a case study inquiry. J Psychiatr Ment Health Nurs 2018 Nov;25(9-10):558-568. [doi: 10.1111/jpm.12501] [Medline: 30372573]
56. Dake K. Myths of nature: culture and the social construction of risk. J Soc Issues 1992:48-37. [doi: 10.1111/j.1540-4560.1992.tb01943.x]
57. Fraser L, Fogarty G, Albion M. A longitudinal analysis of the transition to retirement from elite sport: athletic identity, life satisfaction and career indecision. J Sci Med Sport 2010 Jan;12:e174. [doi: 10.1016/j.jsams.2009.10.366]
58. Martin L, Fogarty GJ, Albion MJ. Changes in athletic identity and life satisfaction of elite athletes as a function of retirement status. J Appl Sport Psychol 2013 Nov 18;26(1):96-110. [doi: 10.1080/10413200.2013.798371]
59. Wylleman P, Reints A. A lifespan perspective on the career of talented and elite athletes: perspectives on high-intensity sports. Scand J Med Sci Sports 2010 Oct;20 Suppl 2:88-94. [doi: 10.1111/j.1600-0838.2010.01194.x] [Medline: 20840566]
60. Chambers T, Harangozo G, Mallett CJ. Supporting elite athletes in a new age: experiences of personal excellence advisers within Australia's high-performance sporting environment. Qualitative Res Sport Exerc Health 2019 May 05;11(5):650-670. [doi: 10.1080/2159676x.2019.1605404]
61. Bevan Jones R, Stallard P, Agha SS, Rice S, Werner-Seidler A, Stasiak K, et al. Practitioner review: co-design of digital mental health technologies with children and young people. J Child Psychol Psychiatry 2020 Aug;61(8):928-940. [doi: 10.1111/jcpp.13258] [Medline: 32572961]
62. Aref-Adib G, Landy G, Eskinazi M, Sommerlad A, Morant N, Johnson S, et al. Assessing digital risk in psychiatric patients: mixed methods study of psychiatry trainees' experiences, views, and understanding. JMIR Ment Health 2020 Jul 29;7(7):e19008 [FREE Full text] [doi: 10.2196/19008] [Medline: 32726288]

63. Steinberg L. Contextual studies: methodology. In: Wright JD, editor. International Encyclopedia of the Social & Behavioral Sciences (Second Edition). Amsterdam: Elsevier Ltd; 2015:786-789.
64. Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. J Med Internet Res 2019 May 09;21(5):e13216 [FREE Full text] [doi: 10.2196/13216] [Medline: 31094356]
65. Pedram S, Palmisano S, Perez P, Mursic R, Farrelly M. Examining the potential of virtual reality to deliver remote rehabilitation. Comput Hum Behav 2020 Apr;105:106223. [doi: 10.1016/j.chb.2019.106223]
66. Bini SA. Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care? J Arthroplasty 2018 Aug;33(8):2358-2361. [doi: 10.1016/j.arth.2018.02.067] [Medline: 29656964]
67. Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim HC, et al. Artificial intelligence for mental health and mental illnesses: an overview. Curr Psychiatry Rep 2019 Nov 07;21(11):116 [FREE Full text] [doi: 10.1007/s11920-019-1094-0] [Medline: 31701320]
68. Victor E, Aghajan ZM, Sewart AR, Christian R. Detecting depression using a framework combining deep multimodal neural networks with a purpose-built automated evaluation. Psychol Assess 2019 Aug;31(8):1019-1027. [doi: 10.1037/pas0000724] [Medline: 31045384]
69. Su C, Xu Z, Pathak J, Wang F. Deep learning in mental health outcome research: a scoping review. Transl Psychiatry 2020 Apr 22;10(1):116. [doi: 10.1038/s41398-020-0780-3] [Medline: 32532967]
70. 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]
71. Nijman SA, Veling W, Greaves-Lord K, Vos M, Zandee CER, Aan Het Rot M, et al. Dynamic interactive social cognition training in virtual reality (DiSCoVR) for people with a psychotic disorder: single-group feasibility and acceptability study. JMIR Ment Health 2020 Aug 07;7(8):e17808 [FREE Full text] [doi: 10.2196/17808] [Medline: 32763880]
72. Bansal A, Padappayil RP, Garg C, Singal A, Gupta M, Klein A. Utility of artificial intelligence amidst the COVID 19 pandemic: a review. J Med Syst 2020 Aug 01;44(9):156. [doi: 10.1007/s10916-020-01617-3] [Medline: 32740678]
73. Caldwell M, Andrews JTA, Tanay T, Griffin LD. AI-enabled future crime. Crime Sci 2020 Aug 5;9(1). [doi: 10.1186/s40163-020-00123-8]
74. King TC, Aggarwal N, Taddeo M, Floridi L. Artificial intelligence crime: an interdisciplinary analysis of foreseeable threats and solutions. Sci Eng Ethics 2020 Feb;26(1):89-120. [doi: 10.1007/s11948-018-00081-0] [Medline: 30767109]
75. Brundage M, Avin S, Clark J, Toner H, Eckersley P, Garfinkel B, et al. The malicious use of artificial intelligence: forecasting, prevention, and mitigation. arXiv. 2018. URL: https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf [accessed 2020-07-22]
76. Wilner A. Cybersecurity and its discontents: artificial intelligence, the Internet of Things, and digital misinformation. Int J 2018 Jul 26;73(2):308-316. [doi: 10.1177/0020702018782496]
77. Cho C, Lee T, Lee J, Seo JY, Jee H, Son S, et al. Effectiveness of a smartphone app with a wearable activity tracker in preventing the recurrence of mood disorders: prospective case-control study. JMIR Ment Health 2020 Aug 05;7(8):e21283 [FREE Full text] [doi: 10.2196/21283] [Medline: 32755884]
78. Newson J, Hunter D, Thiagarajan TC. The heterogeneity of mental health assessment. Front Psychiatry 2020;11:76. [doi: 10.3389/fpsyt.2020.00076] [Medline: 32174852]
79. Brailas A. Psychotherapy in the era of artificial intelligence: Therapist Panoptes. Homo Virtualis 2019 Mar 27;2(1):68. [doi: 10.12681/homvir.20197]
80. Berry K, Salter A, Morris R, James S, Bucci S. Assessing therapeutic alliance in the context of mHealth interventions for mental health problems: development of the Mobile Agnew Relationship Measure (mARM) questionnaire. J Med Internet Res 2018 Apr 19;20(4):e90 [FREE Full text] [doi: 10.2196/jmir.8252] [Medline: 29674307]
81. D'Alfonso S. AI in mental health. Curr Opin Psychol 2020 Dec;36:112-117. [doi: 10.1016/j.copsyc.2020.04.005] [Medline: 32604065]
82. Henson P, Wisniewski H, Hollis C, Keshavan M, Torous J. Digital mental health apps and the therapeutic alliance: initial review. BJPsych Open 2019 Jan;5(1):e15 [FREE Full text] [doi: 10.1192/bjo.2018.86] [Medline: 30762511]
83. Torous J, Hsin H. Empowering the digital therapeutic relationship: virtual clinics for digital health interventions. NPJ Digit Med 2018;1:16. [doi: 10.1038/s41746-018-0028-2] [Medline: 31304301]
84. Rice S, O'Bree B, Wilson M, McEnery C, Lim MH, Hamilton M, et al. Leveraging the social network for treatment of social anxiety: pilot study of a youth-specific digital intervention with a focus on engagement of young men. Internet Interv 2020 Apr;20:100323 [FREE Full text] [doi: 10.1016/j.invent.2020.100323] [Medline: 32435600]
85. Lederman R, Gleeson J, Wadley G, D’Alfonso S, Rice S, Santesteban-Echarri O, et al. Support for carers of young people with mental illness. ACM Trans Comput-Hum Interaction 2019 Feb 23;26(1):1-33. [doi: 10.1145/3301421]

Abbreviations

AIS: Australian Institute of Sport

APSQ:	Athlete Psychological Strain Questionnaire
DSM:	Diagnostic and Statistical Manual of Mental Disorders
DTA:	digital therapeutic alliance
FAQ:	frequently asked question
HCI:	human-computer interaction
ICD:	International Classification of Diseases
IoT:	Internet of Things
MHQ:	Mental Health Quotient
PTSD:	posttraumatic stress disorder
VR:	virtual reality
XR:	extended reality

Edited by J Torous; submitted 22.07.20; peer-reviewed by K Ohara, L Guo; comments to author 09.09.20; revised version received 16.10.20; accepted 17.10.20; published 14.12.20.

Please cite as:
Balcombe L, De Leo D
Psychological Screening and Tracking of Athletes and Digital Mental Health Solutions in a Hybrid Model of Care: Mini Review
JMIR Form Res 2020;4(12):e22755
URL: http://formative.jmir.org/2020/12/e22755/
doi:10.2196/22755
PMID:33271497

Original Paper

Mentoring Young African American Men and Transgender Women Who Have Sex With Men on Sexual Health: Formative Research for an HIV Mobile Health Intervention for Mentors

Michelle R Kaufman[1], MA, PhD; Albert Casella[1], MDP; John Mark Wiginton[1], MPH, MSW; Wenjian Xu[2], PhD; David L DuBois[3], PhD; Renata Arrington-Sanders[4], MSc, MPH, MD; Jeannette Simon[5], BSc; Deb Levine[6], MA

[1]Department of Health, Behavior & Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[2]Department of Sociology & Psychology, School of Public Administration, Sichuan University, Chengdu, China

[3]Division of Community Health Sciences and Institute for Health Research and Policy, University of Illinois Chicago, Chicago, IL, United States

[4]School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[5]Ananizach, LLC, Baltimore, MD, United States

[6]DKF Consulting, Oakland, CA, United States

Corresponding Author:
Michelle R Kaufman, MA, PhD
Department of Health, Behavior & Society
Johns Hopkins Bloomberg School of Public Health
624 N. Broadway, Room 257
Baltimore, MD, 21205
United States
Phone: 1 4109552313
Email: michellekaufman@jhu.edu

Abstract

Background: African American men who have sex with men (MSM) and transgender women bear a disproportionate burden of HIV. Young MSM account for 75% of this burden for youth. When youths lack socially protective resources such as strong networks of adults, including parents, teachers, or community members, mentors may play a critical role in promoting health behaviors. This is especially true for youth at risk for HIV, such as African American youth with sexual and gender minority (SGM) identities. In the past decade, natural mentoring and mentoring programs have proliferated as a key prevention and intervention strategy to improve outcomes for young people at risk for poor academic, social, and health issues. Mentors appear to be able to facilitate health promotion among young SGM by modeling healthy behaviors; however, mentors' knowledge and resource needs regarding sexual health topics including HIV are understudied, as is the potential role of mobile technology in enhancing mentoring relationships and the ability of mentors to learn about sensitive issues faced by youth.

Objective: The aim of this study is to explore how mentoring plays a role in the sexual health of African American SGM youth and understand how mentoring relationships can be strengthened through mobile technology to promote youth HIV prevention behaviors.

Methods: In-depth interviews were conducted with African American SGM youth mentees (n=17) and mentors (n=20) to such youths in 3 Mid-Atlantic cities. Mentee interviews focused on discussions regarding sexual health and HIV and how a mentor could broach such topics. Mentor interviews explored whether sexual health and HIV are currently mentoring topics, mentors' knowledge and confidence in mentoring on these issues, and barriers to discussions. All participants were asked if a mobile app could help facilitate mentoring on sensitive health issues, particularly HIV and sexual health. Data were transcribed, coded, and analyzed for relevant themes.

Results: Sexual health was a common topic in mentoring relationships, occurring more in natural mentorships than in mentoring program pairs. Mentors and mentees felt positive about such discussions. Mentors expressed having limited knowledge beyond condom use and HIV testing, and expressed a need for more complete resources. Both mentors and mentees had mixed comfort levels when discussing sexual health. Sufficient trust and shared lived experiences made discussions easier. Mentees have multifaceted needs; however, mentors stated that an app resource that provided self-training, resources, support from other mentors, and tips for better mentoring could prove beneficial.

Conclusions: For the African American SGM community, access to natural mentors is crucial for young people to learn healthy behaviors. A mobile resource to assist mentors in confidently having discussions with mentees may be a promising way to promote healthy practices.

(JMIR Form Res 2020;4(12):e17317) doi:10.2196/17317

KEYWORDS

mentoring; HIV; mobile app; mHealth; men who have sex with men; transgender; African Americans

Introduction

Background

Young African American gay, bisexual, and other men who have sex with men (MSM) continue to bear a disproportionate burden of HIV [1-5]. In 2018, 26% of all incident HIV infections were among African American MSM, accounting for 37% of incident HIV infections among all MSM. In addition, 75% of African American MSM diagnosed as having HIV were younger than 35 years [5]. Between 2000 and 2010, there was a 133% increase in HIV infections among young African American men aged between 13 to 24 years. This number has since seen a downward trend (11%); however, these young men continue to bear a significant burden of infection in the United States [6,7]. Although less is known about HIV incidence among African American transgender women, a recent meta-analysis found a 44% HIV prevalence in this population, compared with 7% among White transgender women [8]. Moreover, research has highlighted increased vulnerabilities for HIV acquisition among young transgender women, including those reporting African American race or ethnicity [9].

These sexual health disparities are contextualized by oppressive sociostructural factors with which young, urban African American MSM and transgender women must contend, including racism, homophobia, transphobia, and classism [10,11]. These factors underpin experiences of stigma, bullying, rejection, abuse, and home eviction [12-14]; a lack of appropriately tailored sexuality education and accessible, culturally competent health care [13-15]; and fractured support networks, socioeconomic challenges, and limited community resources [16,17]. These conditions set the stage for poor sexual and other health outcomes for young African American MSM and transgender women, including a greater likelihood of engaging in sex without a condom, exchanging sex for money and basic needs, or an increased number of risky encounters with multiple sex partners, all of which contribute to disparities in HIV infection [12,18,19].

Mentoring could be leveraged to mitigate the effects of such conditions and facilitate the sexual health of these populations and other sexual and gender minority (SGM) youth. In the youth development context, a *mentor* is commonly defined as a nonparental adult who builds a relationship with a young person, plays a supportive role in their life, and provides guidance and encouragement to cultivate positive and healthy development [20]. In the past decade, mentoring in formal programs and naturally occurring relationships between youth and trusted adults has proliferated as a key prevention and intervention strategy for young people at risk of poor academic, social, and health outcomes [20]. A meta-analysis by DuBois et al [20] synthesized evaluation findings of one-to-one and group-based youth mentoring programs and found positive, though modest, effects on emotional, psychological, behavioral, social, academic, and career outcomes across diverse mentoring program types. Mentors of SGM youth could likewise provide social and emotional support, sex and sexuality-related guidance and education, assistance with accessing health care and other resources, and social capital through connection to a broader social network. In fact, SGM young adults with a mentor report achieving higher levels of education than those without mentors, with the strongest difference observed among young MSM compared with their heterosexual peers [21]. Mentorship has also been associated with positive sexual identity development and improved quality of life among MSM [22].

Few mentoring programs tailored to the needs of SGM youth exist today, despite the calls for such programs over the past two decades [23-25]. Furthermore, extant non–SGM-specific formal mentoring programs may be perceived as unsafe by SGM youth, particularly those who are also racial or ethnic minorities [26]. This lack of mentoring resources may explain why available data suggest that 60% to 80% of SGM youth tend to receive mentorship from individuals with whom they have naturally occurring relationships, such as nonparental relatives, teachers, and individuals in the lesbian, gay, bisexual, transgender, and queer (LGBTQ) community [21,27,28].

Young African American MSM have explicitly expressed the need for mentors because of family rejection and consequent lack of guidance [25], and they have articulated a need for support and direction regarding sex, sexuality, and HIV prevention [29]. As in other SGM youth populations, some young African American MSM have found mentors naturally [30]. For others, mentorship has come through membership in constructed families (families formed with nonbiologically related individuals) or participation in the house and ball community (an LGBTQ underground subculture) [31,32]. Individuals receiving mentorship in these latter contexts have indicated that such support has helped them build healthy relationships, engage in HIV-preventive behaviors, and has conferred social, emotional, and instrumental support [32-34]. This limited body of research leaves much unknown about mentoring relationships among young African American MSM, and there is a general lack of research on mentoring relationships among young African American transgender women.

Previous research has shown the perceived benefits of mobile phone–based health interventions, and such interventions have great potential for improving young people’s sexual health outcomes [35-38]. Although the incorporation of technology into youth mentoring relationships is new [39], everyday life is now intimately intertwined with mobile technology. The intersection between youth mentoring, mobile technology, and improvements in sexual health deserves further exploration.

Study Objectives

This research sought to explore the role of mentoring among African American MSM and transgender female populations. In addition, this research aimed to learn how mentoring on sexual health issues among this group takes place, if at all, and how mobile technology might enhance mentoring on such sensitive issues. The objectives of this study are to (1) explore the extent to which existing mentoring relationships play a role in the sexual health of African American SGM youth who have sex with men and (2) understand ways in which these mentoring relationships can be strengthened to promote HIV prevention behaviors among the youth, particularly through the use of mobile technology resources. Findings from this formative research informed the development of a mobile app designed to facilitate sexual health conversations between African American SGM and their mentors.

Methods

Participant Description

Participants were mentors and mentees living in Baltimore, Philadelphia, and Washington, DC. Mentee participants were included if they identified as African American; identified as cisgender male, transgender female, or nonbinary assigned male at birth; were aged between 15-24 years; expressed a sexual interest in men; and reported currently having a nonfamilial adult mentor whom they had known for at least three months. Mentor participants were included if they were aged 18 years or older, serving as a nonfamilial mentor to a youth meeting the mentee inclusion criteria, and had at least three months of mentoring experience with such a mentee at the time of participation. The sexual and gender identities of mentors were not considered for eligibility. Study participation was not contingent on one’s mentor or mentee also participating.

Recruitment

Participants were recruited through organizations focused on mentoring and/or LGBTQ services in each city and through passive recruitment methods (eg, posted flyers, flyer distribution at LGBTQ events, advertisements on streaming radio, and social media). Participants in natural mentoring relationships were screened to determine whether the relationship met the definition of mentoring (ie, a relationship between the youth and a nonfamilial adult that focuses on the positive development and well-being of the youth). Formal mentoring program participants were not screened for mentoring relationships because of programs completing this process as part of onboarding procedures.

Data Collection Procedures

Participants completed an oral consent process, followed by an individual, semistructured qualitative interview lasting 45-60 min. The interviews began with a small number of structured questions to assess the characteristics of the mentors and mentees and their relationships. Interviewers were an African American gay male (for the mentee interviews) and an African American cisgender female with more than 25 years of mentoring program experience (for the mentor interviews). Mentee interviews focused on discussions with their mentors regarding sexual health and HIV risk and/or treatment (or why such conversations were absent) and how a mentor could broach such conversations in a youth-friendly manner. Mentor interviews explored whether sexual health and HIV risk and/or treatment are mentoring topics, how much mentors know about these topics, their confidence in mentoring on these issues, and current and/or potential barriers to such discussions. Consistent with the intervention development aim of the research, all participants were also asked whether and how a mobile app might facilitate mentoring on sexual health topics. An abridged interview guide is available in Multimedia Appendix 1.

Data Analysis Procedures

All interviews were audio recorded, transcribed verbatim, and imported into Atlas.ti (version 8.2). Data analysis included inductive and deductive coding techniques, whereby team members independently read transcripts and reached a consensus on the code list. Codebook development was based on the aims of this study and previous literature. Two research team members (both males, 1 gay and 1 heterosexual) independently coded the transcripts. To ensure reliability, 25% of the transcripts were randomly selected for double coding. Coding was considered reliable once the intercoder agreement reached ≥85%. Research team members reviewed code outputs and discussed their relevance to the objectives of this study. Additional data-driven codes were added throughout the analysis, provided there was agreement on their inclusion (in total, 88 codes were established). Data were analyzed thematically (each theme described under a subheading below), a method shown to allow for various epistemological and ontological stances [40].

Results

Participants

Mentee participants’ (n=17) mean age was 20.5 years (SD 1.85; range 18-24), and the mean age for mentors (n=20) was 30.1 years (SD 6.16; range 22-45). Recruiters did not identify any prospective participants in the age range of 15 to 17 years.

Participants were categorized into 4 types: formal mentor, naturally occurring mentor, formal mentee, and naturally occurring mentee. A description of the mentee and mentor demographics by city is available in Table 1. Only 1 mentee reported having a formal mentor, whereas 25% (5/20) of mentors reported being formal mentors. The remaining participants described their mentoring relationship as originating in a natural setting rather than through a formal mentoring program. Examples of these settings include the house-ball community, high school environments where the mentor was a teacher or coach, LGBTQ pride events, and the mentee’s social network. Nearly all participants reported being in their mentoring relationship for at least 1 year (15/17, 88% of mentees and 18/20, 90% of mentors), with an average duration of 2.5 years (mentee mean 2.34 years, SD 1.78; mentor mean 2.7 years, SD 2.08).

Table 1. Demographic and mentoring characteristics (N=37).

Characteristics	Mentee participants, n (%)	Mentor participants, n (%)
City		
Baltimore	5 (29)	2 (10)
Philadelphia	4 (24)	9 (45)
Washington, DC	8 (47)	9 (45)
Gender identity		
Cisgender male	8 (47)	14 (70)
Transgender female	6 (35)	5 (25)
Nonbinary	3 (18)	1 (5)
Age (years)		
15-19	6 (35)	0 (0)
20-29	11 (65)[a]	12 (60)
30-39	N/A[b]	6 (30)
40-49	N/A	2 (10)
Educational attainment		
Less than high school	2 (12)	0 (0)
High school graduate	14 (82)	8 (40)
Some college	1 (6)	5 (25)
Bachelor’s degree	0 (0)	5 (25)
Master’s degree	0 (0)	2 (10)
Mentorship type		
Formal	1 (6)	5 (25)
Naturally occurring	16 (94)	15 (75)
Length of relationship		
Less than 1 year	2 (12)	2 (10)
1-2 years	10 (59)	10 (50)
3-5 years	3 (18)	4 (20)
5 or more years	2 (12)	4 (20)

[a] Mentee eligibility included a maximum age of 24 years.

[b] N/A: not applicable.

The reported communication frequency between mentors and mentees ranged from once per week to multiple times per day. Reported communication tended to be more frequent and used multiple modes (in person, text, and video chat) in naturally occurring relationships as compared with formal mentoring relationships.

We did not conduct a comparison of findings between the transgender and MSM participants given the limited sample (n=5) of transgender women. However, it should be noted that conversations with transgender mentees and their mentors tended to reveal a heightened focus on immediate basic needs.

Sexual Health as a Common Topic of Conversation

Sexual health was often reported to be a prominent conversation topic between mentors and mentees. Almost all participants reported having conversations focused on sexual health within their mentoring relationship. These conversations seemed to be more common in longer (≥1 year), naturally occurring relationships, with a vast majority of longer mentorships reporting this discussion topic. All mentors who were part of a formal program cited the program’s structure (eg, curriculum) as an impediment to such conversations occurring, as the programs were focused on addressing other youth development issues (eg, academic performance). In line with this trend, only 1 participant from a formal mentoring relationship reported having such conversations. Although conversations on sexual health were reported by most participants, mentors revealed reluctance to broach these topics unsolicited:

> *I don’t want her to be like, ‘Why are you asking me questions [about HIV]?’...I’m comfortable with [having a conversation about sexual health], but I*

> *don't want her to look like 'why are you asking me that?'* [Mentor, age 29 years, Washington, DC]

Despite this reluctance, discussions of sexual health were reported to be initiated by mentors in the majority of cases. Mentees described these conversations as occurring in instances where the mentor and the mentee were talking about a mentee's romantic relationship or recent sexual encounter. Mentors reported that they often initiated these conversations by embedding sexual health topics into a broader discussion of healthy relationships, such as the importance of open communication with partners and practicing safe sex.

There was little difference in conversation occurrence as reported by mentors and mentees. Mentors tended to describe these discussions as responses to observing risk factors in their mentee's behavior, such as engaging with multiple partners. Mentees tended to describe discussions of sexual health as embedded in broader topics focused on dating, relationships, or current events.

Positive Attitudes but Limited Mentor Knowledge Regarding Sexual Health

Attitudes toward conversations about sexual health topics between mentors and mentees, once broached, were overwhelmingly positive. Mentees cited confidence in the accuracy of information provided by their mentor because their mentor is a trusted adult with a shared life experience and has, in many cases, *been there*. Mentor participants typically expressed that they had a general knowledge of sexual health; however, a subset described themselves as less knowledgeable about sexual health care or standard prevention practice information. Almost all mentors mentioned that these conversations did not go into great detail or lead to service referral (eg, guiding their mentee to an HIV testing site).

Observing risk behavior in a mentee was a motivating factor for many mentors to introduce a conversation about sexual health. Key elements mentors used when planning to discuss sexual health were ensuring requisite trust was built in the relationship and a concern for the appearing objective:

> *You know, what's really challenging is when you can see them [the mentee engaging in sexual risk behavior] without the other person knowing you can see, but you still have to remain neutral. I feel like [Mentee] and I have a trusting relationship, so I don't think I would be afraid to bring it up. If we didn't have that trust, it would be a different story.* [Mentor, age 45 years, Washington, DC]
>
> *I've talked to him multiple, multiple times about PrEP. And did he take the advice? I don't know....But since I'm aware of [mentee's behavior]...I was like, I'm gonna at least give you [the information]—you can prevent yourself from getting HIV if you take your PrEP every day.* [Mentor, age 26 years, Baltimore]

When approached by a mentor regarding sexual health in general, mentees reported being generally receptive. For example, if a mentor wanted to casually discuss a mentee's romantic relationships or healthy relationships, mentees were attentive. However, when these discussions were focused on mentee risk behavior, such as engaging with multiple partners, mentors talked about being less comfortable and sometimes felt embarrassed to bring it up. Mentees in longer mentoring relationships (>1 year) reported that they trusted the mentor to have their best interests in mind when introducing conversations on sexual health, even if it did lead to some embarrassment.

Mixed Levels of Comfort Mentoring on Sexual Health Topics

When there was sufficient trust in the relationship or observation of potential risk behavior, mentors reported generally feeling comfortable discussing some sexual health topics such as knowing when one is ready to start having sex and using condoms. However, mentors' stated that confidence in mentoring on other sexual health issues, particularly HIV, varied. Mentors who were living with HIV or who reported being in serodiscordant partnerships communicated higher confidence than those without direct experience:

> *Besides that, we talked about, you know, [HIV] status...I had wound up finding out not too long before that I was positive, and when I was young, I didn't have the luxury of having that much insight into the do's and don'ts of how [HIV] is transmitted. Because of my status, I did my homework and so I know what to say when I talk to [Mentee] about it.* [Mentor, age 28 years, Philadelphia]

Mentors who reported feeling less comfortable with HIV-related conversations also reported feeling less confident in their knowledge of HIV prevention practices beyond condom use. Conversely, mentors who reported feeling very comfortable and knowledgeable about HIV also reported detailed conversations with their mentee regarding risk reduction strategies through safer sex practices. Mentors who reported identifying as an SGM individual also reported a need to stay updated on information regarding pre-exposure prophylaxis (PrEP); however, in most cases, mentors did not report having detailed conversations about PrEP.

Likewise, mentees did not recall having in-depth conversations with mentors about sexual health and/or HIV unless they had a long-standing, trusting relationship with the adult. For those who did report such discussions, mentees' reported comfort level appeared to be dependent on multiple factors, including the level of trust, the types of sexual experiences they believed their mentors may have encountered, and whether their mentor has ever implicitly mentioned HIV in previous conversations with the mentee (such as in the context of condom use). When mentees reported being highly trusting of their mentors, they also reported higher comfort levels discussing sensitive health topics with them. In addition, mentees who reported a shared sexual or gender identity with their mentor reported a greater willingness to engage in discussions about sexual health as opposed to a potential mentor without a shared experience:

> *In the lifestyle I live, people ain't really there for you. So sometimes you having somebody in your community to mentor you, that can mean a lot because they understand the issues I'm going through. Instead of going to someone that really don't understand you,*

> *I'd rather go to somebody that understands me and accepts me for me.* [Mentee, transgender female, age 22 years, Washington, DC]

Mentees also reported feeling more comfortable and confident in their mentors' abilities to provide accurate information about sexual health when mentors had direct experience with the SGM community or identified as SGM individuals.

Mentees Have Multifaceted Needs and Mobile Health Tools Could Help

This formative research was conducted in large part to learn whether mentors and their mentees would benefit from an app focused on improving mentorship around sexual health issues. Our original conception was that the app would be primarily used by the mentors, with content to be shared with mentees to spark conversations. When asked about their thoughts regarding such a tool, mentors and mentees reported that an app may be useful in facilitating conversations around sexual health topics, particularly as it relates to sharing knowledge, local resources, and communicating about sexual health in a relationship. We did not provide much detail on our conceptualization of such an app because we wanted to learn what would be most useful for the users themselves:

> *The app sounds like an important thing. Like, someone to remind you that, hey, you have to go to your doctor's appointment today. Or there's these new condoms...Did you get these? Have you seen these? There's a new transgender group going on, I think you should attend, maybe you wanna check it out. So, those things [on the app] would probably make it easy [to use].* [Mentor, age 36 years, Washington, DC]

We also asked if such an app is something that is needed by the SGM mentoring community, and participants overwhelmingly agreed that it would be highly beneficial to both mentors and mentees.

Most participants felt HIV was an important issue in their community and needs to be addressed, particularly with African American MSM and transgender women. Mentors cited the need for an app to feature a space (such as a forum) where mentors could talk with each other about common HIV-related issues and share information regarding other health conditions and life challenges that the SGM youth face. Mentors pointed out the multifaceted needs their mentees may face, ranging from day-to-day survival assistance to longer term health conditions, and that these factors should be addressed in the app as well. Notably, mental health was highlighted as a critical issue by most mentors:

> *If you get nothing else from me in this interview, mental health is number one. Mental health, self-esteem, self-worth, depression, anxiety - those are the things that are getting in the way. Those are some of the factors that are leading to this risky behavior.* [Mentor, age 38 years, Philadelphia]

A potential app's data security and a visually discreet interface were identified as important components to consider in its development, particularly if the app is focused on HIV and/or SGM youth. Mentors also identified a need for self-training and suggested that the app include skills building resources tailored to the mentors of SGM youth. Responses from mentees regarding what should be incorporated into the app included a way to be linked with resources and information in a discreet way:

> *I feel like a lot of people wanna do stuff like [test for HIV, learn about PrEP] secretly. So it will be good if we could like, you know, learn about it without everyone knowing.* [Mentee, cisgender male, age 21 years, Baltimore]

Mentors in each city felt an app focused on improving their knowledge and self-efficacy to mentor on health topics relevant to SGM youth, combined with a place to view successful mentoring on the relevant issues, may be useful in their mentoring practice. Most participants stated that the prospect of an app tailored to the needs of mentors of high-risk youth would be well received in the community.

Discussion

Key Findings and Comparison With Previous Research

Participants in our sample tended to form relationships with mentors and mentees in natural ways rather than through formal youth mentoring programs. This is consistent with studies of LGBTQ youths who have mentors, which have shown that such youths are more likely to report having informal mentors than programmatic mentors [27,28]. Natural mentoring relationships in this study were reported to be of longer duration than formal mentoring relationships, which may be a significant advantage for the mentees, as longer lasting relationships have been linked to more positive youth outcomes [41,42]. In addition to LGBTQ youth in general receiving mentorship through informal channels as opposed to formal programs, research has shown that African American youth tend to use natural mentors on a regular basis as well [43], and those with natural mentorships are more likely to experience positive youth outcomes than those who do not [44].

Naturally occurring relationships, as opposed to those tied to a mentoring program, may be especially important to consider when designing programs that use mentors as the intervention vehicle for African American SGM youth, given that such youths are more likely to be reached through informal networks. One study of mature (older) Black MSM reported that most participants agreed that mature Black MSM should mentor younger Black MSM to help them learn from their mistakes and teach them about safer sex [45]. This was seen as a top priority of the mature Black MSM in the study, especially when considering HIV prevention for younger generations. If older generations of African American SGM are eager to impart their knowledge and provide guidance, this may be a key intervention vehicle for both mentoring and health practitioners alike to promote the health and well-being of younger SGM generations. Although evidence regarding the impact of mentoring on health and other outcomes for SGM youth is just starting to emerge [46], this study begins to unpack the mentoring relationship dynamics and characteristics that would make crucial conversations around sexual health and other sensitive issues possible.

Mentors in this study believed that sexual health topics are important in a mentoring relationship, especially when it concerns African American young SGM, and they expressed a desire to improve their capacity to have such conversations. However, mentors generally did not feel prepared to have in-depth discussions beyond sexual debut and condom use, unless they themselves had direct HIV experience. Furthermore, the few mentors from formal programs interviewed did not feel comfortable discussing sexual health topics because such discussions are not in line with the program's goals. Few mentors felt confident being a strong guide in all things related to sexual health; however, when given the right tools and resources, particularly in the form of a mobile app, the mentors were confident they could grow into such a role.

Participants in this study also expressed a need to focus on the mental health of the mentees. This is consistent with previous research showing that the emotion work involved in navigating a queer identity and maintaining a sense of belonging and security among youth has a stark impact on youth mental health and well-being [47]. Although mental health concerns were not the original focus of the app and this study, this was subsequently added as a module to HIV-focused content based on current findings. In addition, this finding has prompted the research team to make mental health a priority area for future app versions targeting LGBTQ youth populations.

Although research on the mentoring of young African American SGM is still in a nascent stage, the results of this study are consistent with previous findings highlighting the importance of mentorship for this group of youths, particularly to promote health and well-being. A study examining the natural mentoring relationships of young Black MSM showed that participants preferred mentors to be people with whom they shared key attributes and whom they considered successful and resilient in their own lives [48]. The findings of this study echo these previous results, in that mentee participants felt most comfortable with a mentor who shared a lived experience, and mentors who had such a shared experience felt comfortable imparting their knowledge to the young people they mentor.

Previous research on HIV-related apps also highlights areas where mentors can fill crucial gaps for SGM youth at risk. Although there is a proliferation of HIV-related apps, and participants in this study were open to an app focused on mentoring around sexual health issues, the mentors were reluctant to use existing tools because they were not relevant to mentoring relationships. Previous HIV apps have tended to focus on supporting individuals to test or facilitate adherence for those living with HIV [49-51]. Most of these apps also operate with a connection to clinical settings. Although this is important, adolescent and young adult populations are among the least likely demographic groups to seek health care in a clinical setting, considering routine barriers to access, availability, and affordability that they face [52,53]. In addition, few studies have used adult-youth social support networks such as mentoring in developing a mobile health HIV prevention intervention, with available interventions largely focused directly on the individual [54]. No mentors in this sample mentioned using such resources to help guide their mentoring, sometimes even citing stigma around accessing an HIV-related app. Perhaps an app providing sexual health materials for mentors in the context of their mentoring relationship (rather than separately as HIV prevention education) could add a critical element of discretion, potentially benefiting both mentors and mentees in increasing sexual health and reducing risk in this population (SGM of color).

The results of this study also help to throw light on the utility of mentors in addressing sensitive health issues affecting youth, particularly for populations at increased risk. Literature on health outcomes as a result of mentoring is scant and is largely focused on helping youth adapt to chronic illness or disability or looks at substance use prevention (MR Kaufman, unpublished data, 2021, [55,56]). There has not been a comparison of natural mentoring with formal mentoring programs and their respective impact on health outcomes. This would be an important area for future research, especially for marginalized youth at risk of negative health outcomes.

Limitations

Participants in this study were largely recruited by word of mouth and advertising through mentoring and LGBTQ organizations, establishments, and events, which may have produced a biased sample. We were able to recruit only 1 mentee from a formal program and 6 mentors associated with programs, despite our best efforts to reach out to all known mentoring programs (particularly those serving African American youth) in the 3 cities. In our informal conversations with community members, it became clear that this population (African American SGM) simply does not trust formal mentoring programs. Such youths appear to trust mentors that they come to know through their own communities and social networks. Regardless, the small representation of mentors and mentees in formal mentorships likely had an impact on the results, as the dynamics of such relationships and associated mentoring needs may be different from naturally formed relationships. In addition, given the stigmas regarding sexual and gender minorities and HIV in the African American community [39], participants may have been hesitant to discuss the issues. We attempted to minimize this by employing interviewers whom we hoped would be relatable to the participants.

Eligible mentee participants in this study were older (aged 15-24 years) than the age group for which most youth mentoring literature exists (younger than 18 years). We chose the older age group because our focus was on conversations regarding sexual behaviors. Focusing on younger mentees may not have revealed the presence of sexual health conversations because they were not yet relevant for younger youth.

Although we included transgender female participants, we did not have a large enough subgroup sample size to make meaningful comparisons with the cisgender male participants. Furthermore, all of the transgender participants were referred to the study by a social support organization that focuses on addressing the basic needs of youth. Interviews with transgender mentees and their mentors tended to focus on mentoring regarding immediate basic needs such as food security, employment, and housing, as these are the services provided by that organization. Future research should further investigate the broader needs of transgender youths who have mentors using a larger and more representative sample of these young persons.

Conclusions

For the African American SGM community, access to natural mentors may be crucial for young people at risk of HIV to learn to develop healthy behaviors and relationships. Such youth also require mentorship on other health areas (particularly mental health) and life skills. This study highlights the role of mentors for such youth and what is needed for mentors to confidently and effectively address sexual health and other issues with these young people. A resource for mentors to confidently have discussions on critical but sensitive issues with mentees appears promising as a way to promote healthy practices in this vulnerable adolescent population. A mobile app that is discrete and up to date and provides support and resources for the mentors holds promise for meeting some of these needs. Future research on the mentorship of SGM youth of color may also benefit from focusing on naturally occurring mentorship in addition to formal mentoring programs. A mobile app resource for mentors must take into account the fact that many African American SGM seek mentors through informal networks. Through a better understanding of the needs of this particular youth group and their areas of comfort and discomfort when it comes to being mentored by trusted adults, mentoring programs and health practitioners will be better able to address their needs and promote youth health and well-being.

Acknowledgments

The authors would like to thank the following organizations for their assistance on this project: Joy Baltimore, Sascha Bruce Youthwork, The Attic Youth Center, COLOURS, Big Brothers Big Sisters Independence Region, MENTOR Independence Region, Maryland MENTOR, The Children's Hospital of Philadelphia, and Casa Ruby. In addition, they are immensely grateful to James Conley, III, for his assistance with this study and to the Johns Hopkins Adolescent and Young Adult Working Group for their support. The authors would also like to thank Carl Latkin and Nick Ialongo for their guidance and support on this line of research. This research was funded by an Administrative Supplement to the Johns Hopkins University Center for AIDS Research, a National Institutes of Health (NIH)–funded program (P30AI094189), which is supported by the following NIH Co-Funding and Participating Institutes and Centers: National Institute of Allergy and Infectious Diseases, National Cancer Institute, National Institute of Child Health and Human Development, National Heart, Lung, and Blood Institute, National Institute on Drug Abuse, National Institute on Minority Health and Health Disparities, National Institute on Aging, National Institute of General Medical Sciences, and 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 NIH.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mentor interview guide summary.
[DOCX File , 16 KB - formative_v4i12e17317_app1.docx]

References

1. Monitoring Selected National HIV Prevention and Care Objectives By Using HIV Surveillance Data - United States and 6 Dependent Areas - 2013. Centers for Disease Control and Prevention. 2015. URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-supplemental-report-vol-24-3.pdf [accessed 2020-11-17]
2. Dasgupta S, Oster AM, Li J, Hall HI. Disparities in consistent retention in HIV care--11 states and the district of Columbia, 2011-2013. MMWR Morb Mortal Wkly Rep 2016 Feb 5;65(4):77-82 [FREE Full text] [doi: 10.15585/mmwr.mm6504a2] [Medline: 26844978]
3. Maulsby C, Millett G, Lindsey K, Kelley R, Johnson K, Montoya D, et al. HIV among black men who have sex with men (MSM) in the United States: a review of the literature. AIDS Behav 2014 Jan;18(1):10-25. [doi: 10.1007/s10461-013-0476-2] [Medline: 23620241]
4. Mimiaga MJ, Reisner SL, Bland S, Skeer M, Cranston K, Isenberg D, et al. Health system and personal barriers resulting in decreased utilization of HIV and STD testing services among at-risk black men who have sex with men in Massachusetts. AIDS Patient Care STDS 2009 Oct;23(10):825-835 [FREE Full text] [doi: 10.1089/apc.2009.0086] [Medline: 19803696]
5. HIV Surveillance Report: Diagnoses of HIV Infection in the United States and Dependent Areas, 2018 (Preliminary). Centers for Disease Control and Prevention. 2019. URL: https://www.cdc.gov/hiv/library/reports/hiv-surveillance.html [accessed 2020-11-17]
6. NCHHSTP Atlasplus. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchhstp/atlas/index.htm [accessed 2019-05-20]

7. Johnson AS, Hall HI, Hu X, Lansky A, Holtgrave DR, Mermin J. Trends in diagnoses of HIV infection in the United States, 2002-2011. J Am Med Assoc 2014;312(4):432-434 [FREE Full text] [doi: 10.1001/jama.2014.8534] [Medline: 25038362]
8. Becasen JS, Denard CL, Mullins MM, Higa DH, Sipe TA. Estimating the prevalence of HIV and sexual behaviors among the US transgender population: a systematic review and meta-analysis, 2006–2017. Am J Public Health 2019 Jan;109(1):e1-e8. [doi: 10.2105/ajph.2018.304727]
9. Garofalo R, Kuhns LM, Reisner SL, Mimiaga MJ. Behavioral interventions to prevent hiv transmission and acquisition for transgender women. J Acquired Immun Def Synd 2016;72:S220-S225. [doi: 10.1097/qai.0000000000001084]
10. Fields E, Morgan A, Sanders RA. The intersection of sociocultural factors and health-related behavior in lesbian, gay, bisexual, and transgender youth: experiences among young black gay males as an example. Pediatr Clin North Am 2016 Dec;63(6):1091-1106 [FREE Full text] [doi: 10.1016/j.pcl.2016.07.009] [Medline: 27865335]
11. Lacombe-Duncan A. An intersectional perspective on access to HIV-related healthcare for transgender women. Transgend Health 2016;1(1):137-141 [FREE Full text] [doi: 10.1089/trgh.2016.0018] [Medline: 29159304]
12. Hatzenbuehler ML, Pachankis JE. Stigma and minority stress as social determinants of health among lesbian, gay, bisexual, and transgender youth: research evidence and clinical implications. Pediatr Clin North Am 2016 Dec;63(6):985-997. [doi: 10.1016/j.pcl.2016.07.003] [Medline: 27865340]
13. Sentse M, Lindenberg S, Omvlee A, Ormel J, Veenstra R. Rejection and acceptance across contexts: parents and peers as risks and buffers for early adolescent psychopathology. the TRAILS study. J Abnorm Child Psychol 2010 Jan;38(1):119-130 [FREE Full text] [doi: 10.1007/s10802-009-9351-z] [Medline: 19707865]
14. Bird JD, LaSala MC, Hidalgo MA, Kuhns LM, Garofalo R. 'I had to go to the streets to get love': pathways from parental rejection to HIV risk among young gay and bisexual men. J Homosex 2017;64(3):321-342. [doi: 10.1080/00918369.2016.1179039] [Medline: 27092441]
15. Kubicek K, Beyer WJ, Weiss G, Iverson E, Kipke MD. In the dark: young men's stories of sexual initiation in the absence of relevant sexual health information. Health Educ Behav 2010 Apr;37(2):243-263 [FREE Full text] [doi: 10.1177/1090198109339993] [Medline: 19574587]
16. Wilson W. The Ghetto Underclass: Social Science Perspectives. London, UK: Sage Publications; 1993.
17. Wilson W. The Truly Disadvantaged. Chicago, IL: University of Chicago Press; 1987.
18. Hafeez HT, Zeshan M, Tahir MA, Jahan N, Naveed S. Health care disparities among lesbian, gay, bisexual, and transgender youth: a literature review. Cureus 2017 Apr 20;9(4):e1184 [FREE Full text] [doi: 10.7759/cureus.1184] [Medline: 28638747]
19. Institute of Medicine. The Health of Lesbian, Gay, Bisexual,Transgender People: Building A Foundation for A Better Understanding. Washington, DC: National Academies of Sciences, Engineering, & Medicine; 2011.
20. DuBois DL, Portillo N, Rhodes JE, Silverthorn N, Valentine JC. How effective are mentoring programs for youth? A systematic assessment of the evidence. Psychol Sci Public Interest 2011 Aug;12(2):57-91. [doi: 10.1177/1529100611414806] [Medline: 26167708]
21. Johnson D, Gastic B. Natural mentoring in the lives of sexual minority youth. J Community Psychol 2015 Apr 13;43(4):395-407. [doi: 10.1002/jcop.21692]
22. Ross F. Achieving Cultural Competence: the Role of Mentoring in Sexual Minority Identity Development. Indiana: Indiana University; 2005.
23. Alexander CJ. Mentoring for Gay and Lesbian Youth. J Gay Lesbian Soc Serv 2008 Oct 22;10(2):89-92. [doi: 10.1300/j041v10n02_05]
24. Davis TS, Saltzburg S, Locke CR. Supporting the emotional and psychological well being of sexual minority youth: youth ideas for action. Child Youth Serv Rev 2009 Sep;31(9):1030-1041. [doi: 10.1016/j.childyouth.2009.05.003]
25. Lyons T, Johnson AK, Garofalo R. 'What could have been different': a qualitative study of syndemic theory and HIV prevention among young men who have sex with men. J HIV AIDS Soc Serv 2013;12(3-4):- [FREE Full text] [doi: 10.1080/15381501.2013.816211] [Medline: 24244112]
26. Weiston-Serdan T. Critical Mentoring: A Practical Guide. Sterling, VA: Stylus Publishing, LLC; 2017.
27. Gastic B, Johnson D. Teacher-mentors and the educational resilience of sexual minority youth. J Gay Lesbian Soc Serv 2009 Apr 23;21(2-3):219-231. [doi: 10.1080/10538720902772139]
28. Bird JD, Kuhns L, Garofalo R. The impact of role models on health outcomes for lesbian, gay, bisexual, and transgender youth. J Adolesc Health 2012 Apr;50(4):353-357 [FREE Full text] [doi: 10.1016/j.jadohealth.2011.08.006] [Medline: 22443838]
29. Seal DW, Kelly JA, Bloom FR, Stevenson LY, Coley BI, Broyles LA. HIV prevention with young men who have sex with men: what young men themselves say is needed. Medical College of Wisconsin CITY Project Research Team. AIDS Care 2000 Feb;12(1):5-26. [doi: 10.1080/09540120047431] [Medline: 10716014]
30. Torres RS, Harper GW, Sánchez B, Fernández MI, The Adolescent Medicine Trials Network for HIV/AIDS Interventions (co-author). Examining natural mentoring relationships (NMRs) among self-identified gay, bisexual, and questioning (GBQ) male youth. Child Youth Serv Rev 2012 Jan 1;34(1):8-14 [FREE Full text] [doi: 10.1016/j.childyouth.2010.12.018] [Medline: 23408225]

31. Zarwell MC, Robinson WT. The influence of constructed family membership on HIV risk behaviors among gay, bisexual, and other men who have sex with men in New Orleans. J Urban Health 2018 Apr;95(2):179-187 [FREE Full text] [doi: 10.1007/s11524-017-0203-9] [Medline: 29047021]
32. Dickson-Gomez J, Owczarzak J, St Lawrence J, Sitzler C, Quinn K, Pearson B, et al. Beyond the ball: implications for HIV risk and prevention among the constructed families of African American men who have sex with men. AIDS Behav 2014 Nov;18(11):2156-2168 [FREE Full text] [doi: 10.1007/s10461-014-0836-6] [Medline: 24980248]
33. Phillips G, Peterson J, Binson D, Hidalgo J, Magnus M, YMSM of color SPNS Initiative Study Group. House/ball culture and adolescent African-American transgender persons and men who have sex with men: a synthesis of the literature. AIDS Care 2011 Apr;23(4):515-520. [doi: 10.1080/09540121.2010.516334] [Medline: 21271403]
34. Arnold EA, Bailey MM. Constructing home and family: how the ballroom community supports African American GLBTQ youth in the face of HIV/AIDS. J Gay Lesbian Soc Serv 2009 Jan 1;21(2-3):171-188 [FREE Full text] [doi: 10.1080/10538720902772006] [Medline: 23136464]
35. Holloway IW, Rice E, Gibbs J, Winetrobe H, Dunlap S, Rhoades H. Acceptability of smartphone application-based HIV prevention among young men who have sex with men. AIDS Behav 2014 Feb;18(2):285-296 [FREE Full text] [doi: 10.1007/s10461-013-0671-1] [Medline: 24292281]
36. Feroz A, Abrejo F, Ali SA, Nuruddin R, Saleem S. Using mobile phones to improve young people's sexual and reproductive health in low- and middle-income countries: a systematic review protocol to identify barriers, facilitators and reported interventions. Syst Rev 2019 May 18;8(1):117 [FREE Full text] [doi: 10.1186/s13643-019-1033-5] [Medline: 31103044]
37. Levine D. Sex Education in the Digital Age. In: Nixon PG, Dusterhoft IK, editors. Sex in the Digital Age. New York, USA: Routledge; 2017.
38. 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]
39. Kaufman M. E-Mentoring. National Mentoring Resource Center. 2017. URL: http://nationalmentoringresourcecenter.org/images/PDF/E-Mentoring_Model_Review.pdf [accessed 2020-11-17]
40. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
41. Grossman JB, Rhodes JE. The test of time: predictors and effects of duration in youth mentoring relationships. Am J Community Psychol 2002;30(2):199-219. [doi: 10.1023/a:1014680827552]
42. Zimmerman MB, Behrendt DE. Handbook of Youth Mentoring. In: DuBois DL, Karcher MJ, editors. Natural Mentoring Relationships. Thousand Oaks, CA: Sage Publishers; 2005:143-157.
43. Culyba AJ, Ginsburg KR, Fein JA, Branas CC, Richmond TS, Wiebe DJ. Protective effects of adolescent-adult connection on male youth in urban environments. J Adolesc Health 2016 Feb;58(2):237-240 [FREE Full text] [doi: 10.1016/j.jadohealth.2015.10.247] [Medline: 26802994]
44. DuBois DL, Silverthorn N. Characteristics of natural mentoring relationships and adolescent adjustment: evidence from a national study. J Prim Prev 2005 Mar;26(2):69-92. [doi: 10.1007/s10935-005-1832-4] [Medline: 15977043]
45. Tobin KE, Winiker AK, Smith C. Understanding the needs of older (mature) black men who have sex with men: results of a community-based survey. J Health Care Poor Underserved 2018;29(4):1558-1569. [doi: 10.1353/hpu.2018.0112] [Medline: 30449763]
46. LGBTQ Supplement to the Elements of Effective Practice for Mentoring. MENTOR: The National Mentoring Partnership. 2019. URL: https://www.mentoring.org/wp-content/uploads/2019/11/MENTOR_LGBTQ-Supplement-to-EEP-for-Mentoring.pdf [accessed 2020-11-17]
47. McDermott E, Gabb J, Eastham R, Hanbury A. Family trouble: heteronormativity, emotion work and queer youth mental health. Health (London) 2019 Jul 24:1363459319860572 [FREE Full text] [doi: 10.1177/1363459319860572] [Medline: 31339365]
48. Reed SJ, Strzyzykowski T, Chiaramonte D, Miller RL. 'If it weren’t for them, I’d probably be lost': the diversity and function of natural mentors among young black men who have sex with men. Youth Soc 2018 May 3;51(7):961-980. [doi: 10.1177/0044118x18771246]
49. Dworkin MS, Lee S, Chakraborty A, Monahan C, Hightow-Weidman L, Garofalo R, et al. Acceptability, feasibility, and preliminary efficacy of a theory-based relational embodied conversational agent mobile phone intervention to promote HIV medication adherence in young HIV-positive African American MSM. AIDS Educ Prev 2019 Feb;31(1):17-37. [doi: 10.1521/aeap.2019.31.1.17] [Medline: 30742481]
50. Escobar-Viera C, Zhou Z, Morano JP, Lucero R, Lieb S, McIntosh S, et al. The Florida mobile health adherence project for people living with HIV (FL-MAPP): longitudinal assessment of feasibility, acceptability, and clinical outcomes. JMIR Mhealth Uhealth 2020 Jan 8;8(1):e14557 [FREE Full text] [doi: 10.2196/14557] [Medline: 31913127]
51. Fan X, She R, Liu C, Zhong H, Lau JT, Hao C, et al. Evaluation of smartphone APP-based case-management services among antiretroviral treatment-naïve HIV-positive men who have sex with men: a randomized controlled trial protocol. BMC Public Health 2020 Jan 20;20(1):85 [FREE Full text] [doi: 10.1186/s12889-020-8171-5] [Medline: 31959139]

52. Rider GN, McMorris BJ, Gower AL, Coleman E, Eisenberg ME. Health and care utilization of transgender and gender nonconforming youth: a population-based study. Pediatrics 2018 Mar;141(3):- [FREE Full text] [doi: 10.1542/peds.2017-1683] [Medline: 29437861]
53. Macapagal K, Bhatia R, Greene GJ. Differences in healthcare access, use, and experiences within a community sample of racially diverse lesbian, gay, bisexual, transgender, and questioning emerging adultsDifferences in healthcare access, use, and experiences within a community sample of racially diverse lesbian, gay, bisexual, transgender, and questioning emerging adults. LGBT Health 2016 Dec;3(6):434-442. [doi: 10.1089/Lgbt.2015.0124]
54. Schnall R, Travers J, Rojas M, Carballo-Diéguez A. eHealth interventions for HIV prevention in high-risk men who have sex with men: a systematic review. J Med Internet Res 2014 May 26;16(5):e134 [FREE Full text] [doi: 10.2196/jmir.3393] [Medline: 24862459]
55. Dowd AJ, Chen MY, Jung ME, Beauchamp MR. 'Go girls!': psychological and behavioral outcomes associated with a group-based healthy lifestyle program for adolescent girls. Transl Behav Med 2015 Mar;5(1):77-86 [FREE Full text] [doi: 10.1007/s13142-014-0285-9] [Medline: 25729456]
56. Helfrich CM, McWey LM. Substance use and delinquency. J Fam Issues 2013 Mar 5;35(10):1322-1338. [doi: 10.1177/0192513x13478917]

Abbreviations

LGBTQ: lesbian, gay, bisexual, transgender, and queer
MSM: men who have sex with men
NIH: National Institutes of Health
PrEP: pre-exposure prophylaxis
SGM: sexual and gender minority

Edited by G Eysenbach; submitted 06.12.19; peer-reviewed by B Kerr, T Kuhn, H Cho; comments to author 16.03.20; revised version received 30.06.20; accepted 02.10.20; published 17.12.20.

Please cite as:
Kaufman MR, Casella A, Wiginton JM, Xu W, DuBois DL, Arrington-Sanders R, Simon J, Levine D
Mentoring Young African American Men and Transgender Women Who Have Sex With Men on Sexual Health: Formative Research for an HIV Mobile Health Intervention for Mentors
JMIR Form Res 2020;4(12):e17317
URL: https://formative.jmir.org/2020/12/e17317
doi:10.2196/17317
PMID:33331822

Original Paper

A Smartphone App for Patients With Acute Coronary Syndrome (MoTER-ACS): User-Centered Design Approach

Nazli Bashi[1,2], RN, MAppSc; Marlien Varnfield[1], PhD; Mohanraj Karunanithi[1], PhD

[1]Australian eHealth Research Centre, Commonwealth Scientific and Industrial Research Organisation, Brisbane, Australia
[2]School of Medicine, The University of Queensland, Brisbane, Australia

Corresponding Author:
Nazli Bashi, RN, MAppSc
Australian eHealth Research Centre
Commonwealth Scientific and Industrial Research Organisation
Level 5 - UQ Health Sciences
Building 901/16, Royal Brisbane and Women's Hospital
Brisbane, 4029
Australia
Phone: 61 732533611
Email: ghafouryan@hotmail.com

Abstract

Background: Postdischarge interventions are limited for patients with acute coronary syndrome (ACS) due to few scheduled visits to outpatient clinics and the need to travel from remote areas. Smartphones have become viable lifestyle technology to deliver home-based educational and health interventions.

Objective: The aim of this study was to develop a smartphone-based intervention for providing postdischarge support to patients with ACS.

Methods: The content of Mobile Technology–Enabled Rehabilitation for Patients with ACS (MoTER-ACS) was derived from a series of small studies, termed prestudy surveys, conducted in 2017. The prestudy surveys were conducted in Prince Charles Hospital, Queensland, Australia, and consisted of questionnaires among a convenience sample of patients with ACS (n=30), a focus group discussion with health care professionals (n=10), and an online survey among cardiologists (n=15). Responses from the patient survey identified educational topics of MoTER-ACS. The focus group with health care professionals assisted with identifying educational materials, health monitoring, and self-management interventions. Based on the results of the cardiologists' survey, monitoring of symptoms related to heart failure exacerbation was considered as a weekly diary.

Results: The MoTER-ACS app covers multimedia educational materials to adopt a healthy lifestyle and includes user-friendly tools to monitor physiological and health parameters such as blood pressure, weight, and pain, assisting patients in self-managing their condition. A web portal that is linked to the data from the smartphone app is available to clinicians to regularly access patients' data and provide support.

Conclusions: The MoTER-ACS platform extends the capabilities of previous mobile health platforms by providing a home-based educational and self-management intervention for patients with ACS following discharge from the hospital. The MoTER-ACS intervention narrows the gap between existing hospital-based programs and home-based interventions by complementing the postdischarge program for patients with ACS.

(JMIR Form Res 2020;4(12):e17542) doi:10.2196/17542

KEYWORDS

mobile health; mHealth; mobile health apps; smartphone; mobile phone; self-management; patient education; cardiovascular disease; acute coronary syndrome

Introduction

Acute coronary syndrome (ACS) includes a broad spectrum of clinical presentations of ST-segment elevation and non–ST-segment elevation myocardial infarction (MI) and different types of angina. Coronary heart disease affects 7.7% of Australians, and it is the leading cause of total burden of disease across life stages and consequently one of the most common causes of medical admissions [1]. In 2015, 61,600 people had an acute coronary event in the form of a heart attack or unstable angina, which equates to around 170 events every day. Furthermore, coronary heart disease contributed to 12% of all deaths in Australia in 2016 [2]. This high number of hospital admissions could be reduced by providing education, health monitoring, and support. However, due to providers' workloads, the time-intensive nature of health services that provide face-to-face appointments, and the high number of remote and outer metropolitan patients, many Australian health services can only provide follow-up by phone for the majority of patients with ACS [3].

Currently, the ACS postdischarge model of care places limitations on conducting a full assessment and efficient follow-up of a patient's condition and progress. Traditional postdischarge models include fixed appointments for clinical review regardless of patients' requirements [4,5]. These episodic reviews do not provide clinicians with sufficient and updated information about a patient's condition, as they may not happen at the right time to address the patient's requirements. Furthermore, traditional hospital-based health care programs require significant resources, with patients experiencing long waiting times; these resources could be more appropriately used if prioritized based on patients' needs [6]. A new ambulatory care model is required to enable discharged patients to receive either traditional face-to-face follow-up, especially if they have high risks and needs, or a home-based approach [5].

Advances in mobile communication technologies and medical devices enable mobile health (mHealth) interventions to combine the accessibility of home-based rehabilitation with the clinical expertise, supervision, and coaching that has traditionally been limited to hospital-based practices [7]. Previous research demonstrated a positive impact of remote patient monitoring, including mHealth, on patients with cardiovascular disease, mainly heart failure [8-10]. However, little is known about the feasibility and effects of smart phone–based remote monitoring on the outcomes of patients with ACS. If the mHealth intervention encourages homebound patients to actively engage in their disease management by self-monitoring physiological parameters under the supervision of health care professionals and adhering to medications, then the mHealth intervention has potential to improve patient health outcomes and quality of life [11,12]. The development of such interventions requires an iterative process of obtaining information and guidance from all stakeholders, including patients, software engineers, and health care providers [13]. The aim of this study was to develop a smartphone-based intervention to provide postdischarge support to patients with ACS. In this study, we involved stakeholders in prestudy surveys to develop a theory-based intervention that provides health monitoring, education, and support. This paper first defines the design objectives and theoretical framework of an mHealth postdischarge intervention, then describes the steps we performed to extend a mobile technology–enabled rehabilitation platform and its key components to postdischarge support of patients with ACS.

Methods

Design Objectives

Mobile Technology–Enabled Rehabilitation (MoTER) is a platform designed for home-based monitoring in cardiac rehabilitation using a smartphone app and the web [14]. The platform consists of a smartphone app (Android or iOS) with educational videos, health measures, exercise review, goal setting, motivational text messages, and a web portal. We extended the platform to create MoTER for patients with ACS (MoTER-ACS), which integrates a smartphone app (Android or iOS), Bluetooth-enabled devices (blood pressure [BP] cuff and body weight [BW] scale), and a web portal to provide health monitoring, evidenced-based education, and health care provider support for patients with ACS following discharge from the hospital. Furthermore, we aimed to optimize the app's usability and to simplify navigation by including different levels of menu choices, creating a highly engaging appearance using quality images and adequate text, and reducing app errors through frequent testing. To achieve these aims, we studied mHealth apps by conducting literature reviews, including a systematic review on mHealth strategies and structures [15].

Core Components of Secondary Prevention

Based on the Australian guidelines [16], interventions to reduce modifiable risk factors for cardiovascular disease (CVD) include advice and treatment on (1) smoking cessation, (2) maintenance of normal BP, (3) lowering of the serum cholesterol concentration and maintenance of the lipid profile within guideline levels, (4) lowering of serum glucose to within guideline levels, and (5) information on lifestyle risk factors, such as physical activity, diet, obesity, and alcohol consumption. The core components of secondary prevention include health evaluation, patient education, exercise training, lifestyle risk factor management, medication, and psychosocial management [16]. In this study, we aimed to address these components in the development of an mHealth intervention for postdischarge management of patients with ACS.

Self-Efficacy Theory

Theory-based interventions have been demonstrated to be more effective than those without theoretical underpinnings. Applying behavior change theories to mHealth intervention designs significantly increases the likelihood of success [9]. The use of a theoretical framework enables researchers to address the complexity of mHealth interventions when involving a diverse set of stakeholders and their perspectives in order to create change [13].

The development of the MoTER-ACS intervention was informed by the social cognitive theory and used components of self-efficacy informational sources. The self-efficacy theory was adopted to reinforce behavior change, including problem solving, goal setting, action planning, e-diaries, self-monitoring, educational instructions, role modeling, and health care providers' persuasion.

Self-efficacy stands at the core of the social cognitive theory and consists of all the thoughts that affect human functioning [17]. Self-efficacy is defined as an individual's belief in his or her capacity to achieve behaviors necessary to produce specific performance skills [17]. Self-efficacy reflects confidence in the ability to control one's own motivation, behavior, and social environment [17]. MoTER-ACS pedagogical principles were based on self-efficacy elements, including mastery experience, role modeling, and verbal persuasion. Previous research has recommended using a combination of 3 to 4 sources of the theory to promote a stronger sense of self-efficacy and a greater willingness to undergo behavioral change and thus produce optimal results [18]. Mastery experience was implemented by asking participants to self-monitor and record their health measures (for example, BP and BW) and lifestyle parameters, such as smoking, drinking alcohol, and consuming fruits and vegetables.

Role modeling through observing others' actions is considered effective in increasing self-efficacy [19]. Peer role modeling was reinforced through educational materials linked to the Australian Heart Foundation website and the stories of people who had similar conditions. This aimed to resemble role models and encourage patients' engagement with the intervention. A further source of self-efficacy, verbal persuasion, was applied to the intervention through motivational SMS text messages and feedback from health care professionals.

Platform Development and Customization

The purpose of the MoTER-ACS was to overcome accessibility barriers and enable patients to either receive traditional face-to-face outpatient follow-up or a smartphone-based approach. The major platform design objectives were to provide patient-centered educational materials and symptom monitoring, support, and feedback by health care professionals based on the theory-driven strategies. To achieve these objectives and extend the platform, we conducted prestudy surveys that consisted of questionnaires among a group of patients (n=30), a focus group discussion with health care professionals (n=10), and an online survey among cardiologists (n=15). The study procedure was approved by the relevant human research ethics committee.

To address the research methodology as an iterative process and to modify the MoTER-ACS intervention, the survey among patients and the focus group with health care professionals will be repeated following feasibility testing. Repeating the prestudy surveys aims to determine any modifications required.

Prestudy Survey

Patient Survey

We conducted a systematic review to investigate smartphone-based educational interventions for patient self-management. The review also explored the mHealth structures and strategies (including format, interactivity, use of theory, duration of education, and health care professionals' follow-up) of the educational interventions, along with any documented theory or framework that informed the design of such interventions. The results of the systematic review were published separately [15].

After conducting the systematic review, we explored the perceived learning and educational needs of patients with ACS (n=30) after an episode of MI or angina, and we assessed their health-related literacy. We used validated questionnaires and recruited a convenience sample of patients who attended an outpatient clinic at a metropolitan hospital located in Brisbane, Australia. The participant characteristics are provided in Multimedia Appendix 1. Since we could only involve patients during outpatient clinic visits, it was not possible to invite a group of patients to attend a focus group discussion at a specific day and time. In addition, many of these patients travelled about an hour to come to the clinic, so we decided to conduct a survey with each patient individually. After signing the consent form, participants completed a demographic information questionnaire, the Cardiac Patients Learning Needs Inventory (CPLNI) [20,21], and the Australian version of the Short Test of Functional Health Literacy in Adults (S-TOFHLA) [22].

Statistical analysis was performed using IBM SPSS Statistics version 21 for Windows (IBM Corp). Descriptive statistics (mean, standard deviation, frequencies, and percentage) were used to describe the study sample and patients' perceptions of learning needs and health literacy.

The CPLNI is a Likert scale consisting of 38 items. Participants were asked to score each question from 1 to 5 according to the level of importance (1=not important, 2=somewhat important, 3=moderately important, 4=important, and 5=very important). According to the CPLNI scoring method, we assessed patients' learning needs based on the mean score of each domain, which ranged from 1 to 5. Higher scores reflected greater learning needs. To identify the educational topics most important for patients with ACS, the mean of each of the 8 educational topics was calculated and the topics were ranked from highest to lowest. Reasons for MI and signs and symptoms were ranked as the most important topics for learning, followed by medications management, lifestyle factors, diet, psychological factors, and lastly, physical activity (Multimedia Appendix 2). The identified topics were considered as part of the intervention's educational information. Therefore, we sought authorization from the Australian Heart Foundation to embed links from their website in the MoTER-ACS intervention. Furthermore, we included related video clips based on the educational topics identified by patients.

Monitoring health literacy assists health care providers in identifying patients who have difficulties with the educational instructions of cardiac rehabilitation and require further education to obtain adequate disease-related knowledge [22]. In this study, we observed adequate health literacy among the study participants. The S-TOFHLA results were scored as inadequate (0-16 correct answers), marginal (17-22 correct answers), and adequate (23-36 correct answers) health literacy. Of 30 participants, 1 (3%) was identified as having inadequate health literacy and 28 (94%) were identified as having adequate health literacy. One participant found the survey frustrating and did not answer it. No patient was identified as having marginal health literacy.

Focus Group

The purpose of the focus group was to identify components of MoTER-ACS for the postdischarge management of patients. In May 2017, we conducted a focus group with health care professionals (n=10) from a metropolitan hospital located in

Brisbane, Australia. Participants from a multidisciplinary team including cardiologists, nurse practitioners, clinical nurses, and a physiotherapist contributed to a 1-hour discussion by responding to 8 questions on the applicability of smartphone-based educational and health interventions (Multimedia Appendix 3). Transcriptions of the audiotaped session were generated and then imported into NVivo 11.0 (QSR International) for thematic analysis. The NVivo software was used for managing and organizing data, facilitating the process of analysis, identifying themes, collecting insight, and drawing conclusions [23]. Based on the focus group questions, relevant codes were assigned to the text fragments, reflecting the words spoken by the participants in a more abstract way. The coding process assisted with structuring and revealing themes within the text. Four major themes and their subthemes emerged from the qualitative analysis. The complete results of the focus group were published in 2018 [24].

Health care providers indicated that comprehensive education on diet, particularly by providing a daily meal plan, is essential for patients with ACS. For ACS symptoms, clinicians recommended mainly focusing on educating patients instead of monitoring chest pain and shortness of breath daily, as these are subjective and may not sufficiently inform clinicians. Results of the focus group also suggested that monitoring health measures such as BP and BW may result in increased awareness of patient physical health, yet may not be sufficient to support patients with ACS via the smartphone-based intervention [24]. Therefore, monitoring pain, emotional status, and other health measures was recommended. Real-time support via FaceTime or video conferencing was indicated to be motivational and supportive for patient engagement with the intervention. Higher age, low educational level, and lack of computer skills were identified as potential barriers for patients with ACS to engage with the smartphone-based intervention [24]. The items related to educational topics, meal plans, and measuring health parameters identified by health care professionals shaped different parts of the MoTER-ACS intervention.

Cardiologist Survey

We used an online survey to investigate specialists' postdischarge practices for patients with ACS. Investigating cardiologists' perspectives aimed to provide insight on methods of patient treatment and follow-up. This assisted us in translating specialists' practices into knowledge for developing a smartphone-based postdischarge intervention.

We developed the online survey based on the key objectives for developing a model of care introduced by the US Department of Health and Human Services' Center for Disease Control and Prevention [25]. The survey consisted of 12 multiple choice questions that covered topics such as systematic management to meet the needs of patients with ACS, multidisciplinary care coordination and communication, partnership between patients and cardiologists, and patient risk assessment. Through SurveyMonkey, the survey was sent to the email addresses of cardiologists from the hospital located in Brisbane, Australia. Of 45 cardiologists, 15 responded to the survey. Descriptive statistics were used to analyze the responses. The frequency of responses to each question were calculated and responses with higher frequency were considered within the intervention design if applicable.

The specialists considered regular medical assessment and follow-up at an outpatient clinic an effective method for the prevention of ACS readmission in the first 12 months. Based on the results of the online survey, the majority of cardiologists provided their services based on the patient's risk level. Consultation services were provided to patients with high risk every 3 or 6 months. Patients with low risk were followed up every 12 months. The results also showed that follow-up appointments for patients with ACS were mainly to assess response to treatment, prevent readmission, provide support, and adjust medication (Multimedia Appendix 4).

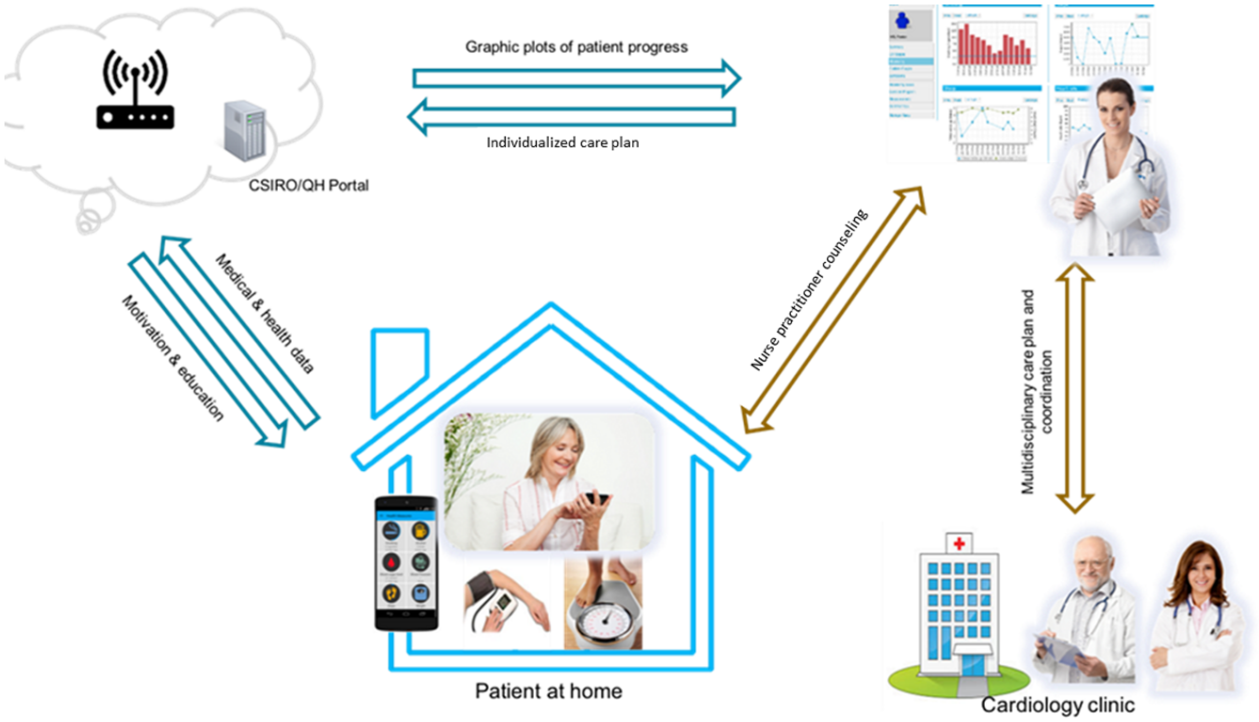
Onset of heart failure (HF) was identified as the main reason patients with ACS presented to the hospital. Other reasons, including chest pain, shortness of breath, unstable angina, and ST-elevation myocardial infarction or non–ST-elevation myocardial infarction, were also reported by cardiologists. Based on the results of the online survey with cardiologists, monitoring symptoms related to heart failure exacerbation was considered as a tool (weekly diary) in the MoTER-ACS intervention.

Results

Platform Components

We extended the MoTER platform (Figure 1) based on the results from the prestudy. The MoTER-ACS platform consists of smartphone apps (Android and iOS), Bluetooth-enabled medical devices, and a web portal. Apps are used for providing education and personalized feedback, collecting physiological data, recording patients' self-observations of their health-related behaviors, and facilitating health care providers' consultations via audio or video. All the data entered into the apps by patients are synchronized daily to a web portal on a secure server with a user-friendly front-end web portal, where clinicians can regularly (every 2 to 3 days) monitor physiological data.

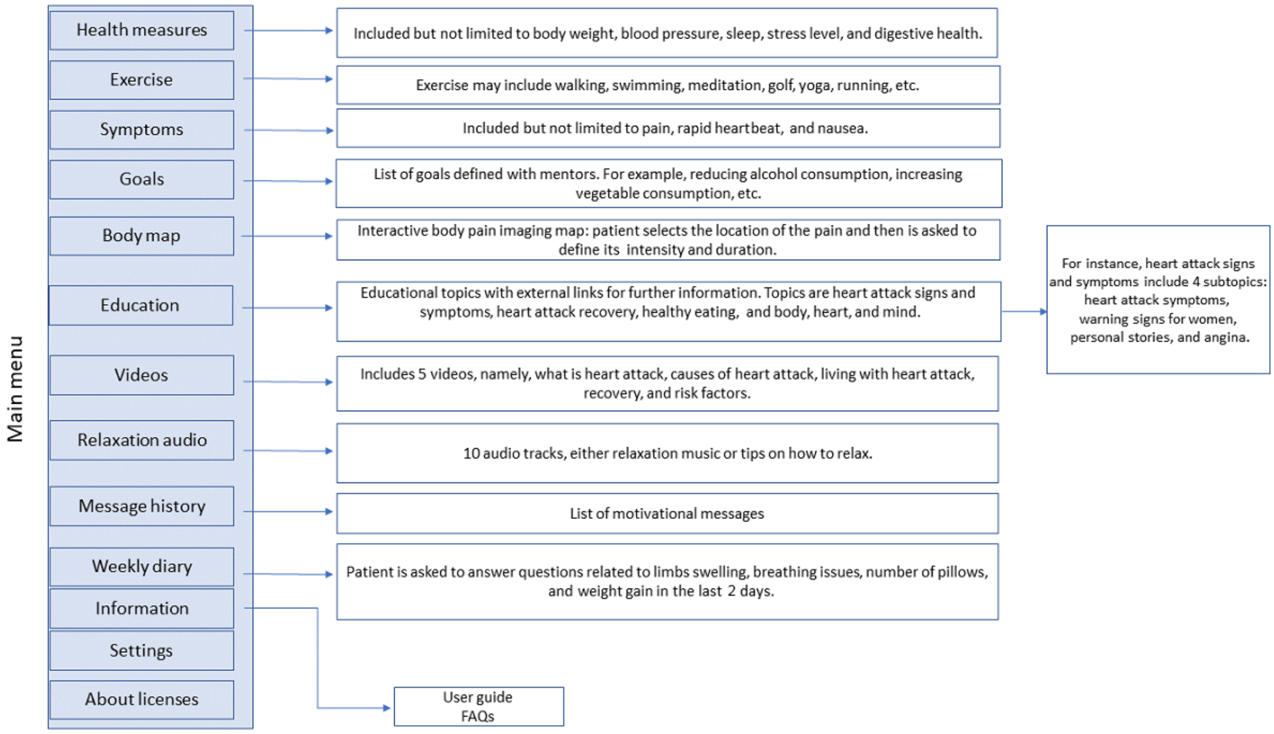
Figure 1. MoTER-ACS platform. CSIRO: Commonwealth Scientific and Industrial Research Organisation; MoTER-ACS: Mobile Technology–Enabled Rehabilitation for Patients With Acute Coronary Syndrome; QH: Queensland Health.



MoTER-ACS App

The MoTER-ACS apps consist of a number of software components, including health measures, multimedia educational materials, a body pain map, a weekly diary, relaxation audio, and motivational messages. Figure 2 outlines the navigation of the MoTER-ACS app, with examples from each category.

Figure 2. Navigation of the MoTER-ACS app. FAQ: frequently asked question; MoTER-ACS: Mobile Technology–Enabled Rehabilitation for Patients With Acute Coronary Syndrome.



Included but not limited to body weight, blood pressure, sleep, stress level, and digestive health.

Exercise may include walking, swimming, meditation, golf, yoga, running, etc.

Included but not limited to pain, rapid heartbeat, and nausea.

List of goals defined with mentors. For example, reducing alcohol consumption, increasing vegetable consumption, etc.

Interactive body pain imaging map: patient selects the location of the pain and then is asked to define its intensity and duration.

Educational topics with external links for further information. Topics are heart attack signs and symptoms, heart attack recovery, healthy eating, and body, heart, and mind.

For instance, heart attack signs and symptoms include 4 subtopics: heart attack symptoms, warning signs for women, personal stories, and angina.

Includes 5 videos, namely, what is heart attack, causes of heart attack, living with heart attack, recovery, and risk factors.

10 audio tracks, either relaxation music or tips on how to relax.

List of motivational messages

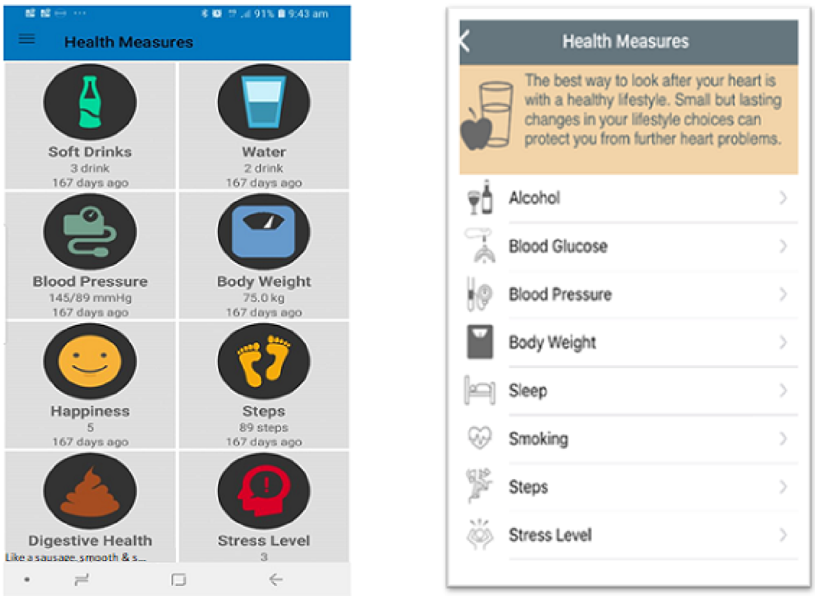
Patient is asked to answer questions related to limbs swelling, breathing issues, number of pillows, and weight gain in the last 2 days.

User guide
FAQs

Health Measures

Following enrollment, each patient sets goals (smoking, alcohol intake, etc) with the help of their nurse, or mentor, to start the program. During the intervention, the mentor discusses the patient's progress in comparison with the set goals and assists in setting new goals. Therefore, patients report a number of health and lifestyle measures, such as BP, BW, alcohol consumption, and servings of fruit consumed (Figure 3). Patient medication adherence is measured through agreement between the patient and mentor and as part of the health measures, where patients can enter their daily medication intake. To establish agreement between the patient and mentor, the mentor asks the patient if they agree to provide information about their medications, then considers this as the patient inputs information into the MoTER-ACS app.

Figure 3. Health measures screen for iPhone.

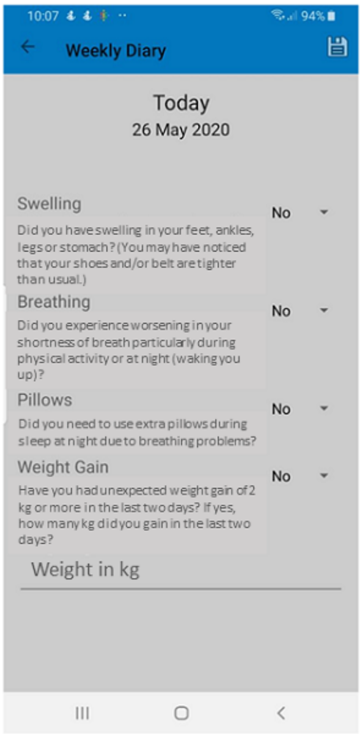


The best way to look after your heart is with a healthy lifestyle. Small but lasting changes in your lifestyle choices can protect you from further heart problems.

Weekly Diary

We developed a weekly diary to monitor HF symptoms. In addition to the daily monitoring of different health measures, patients are asked to answer 4 questions related to signs and symptoms of HF exacerbation, including swelling of the limbs, breathing patterns, number of pillows (when sleeping), and weight gain (more than 2 kg in the past 2 days). The symptoms are recorded weekly and the results are reported to specialists (Figure 4).

Figure 4. Weekly diary.



Today
26 May 2020

Swelling: No
Did you have swelling in your feet, ankles, legs or stomach? (You may have noticed that your shoes and/or belt are tighter than usual.)

Breathing: No
Did you experience worsening in your shortness of breath particularly during physical activity or at night (waking you up)?

Pillows: No
Did you need to use extra pillows during sleep at night due to breathing problems?

Weight Gain: No
Have you had unexpected weight gain of 2 kg or more in the last two days? If yes, how many kg did you gain in the last two days?

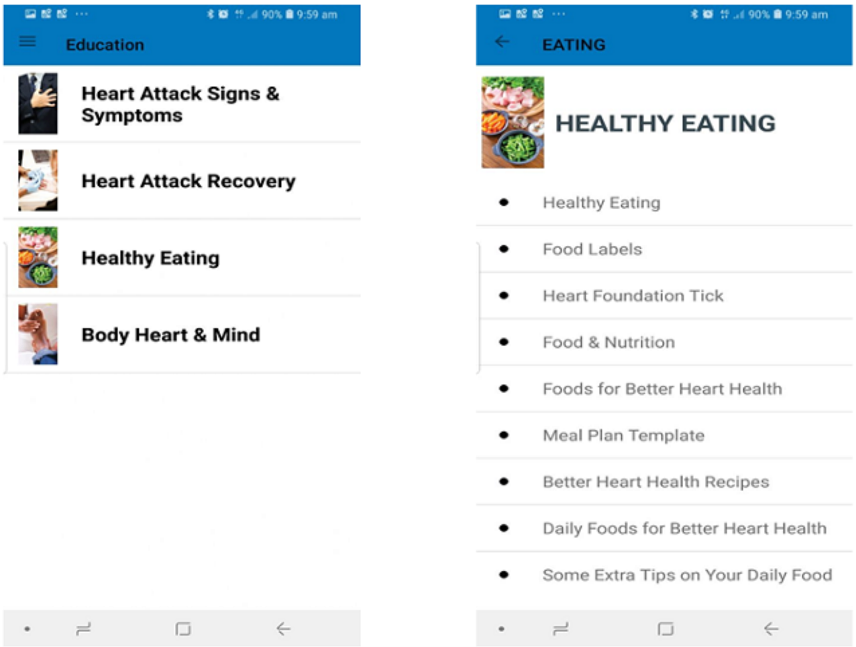
Weight in kg

Educational Materials

In addition to 5 video clips that provide information on CVD conditions, patients have access to a wide range of information through links to the Australian Heart Foundation website, including MI signs and symptoms, MI recovery, healthy eating, and body, heart, and mind (Figure 5).

Nutritional education has a beneficial impact on the dietary habits and nutritional knowledge of patients with CVD [26]. Hence, to encourage healthy eating, we developed diet-related instructions that cover topics such as foods that are better for heart health, daily meal plans, and tips to prepare daily foods and recipes.

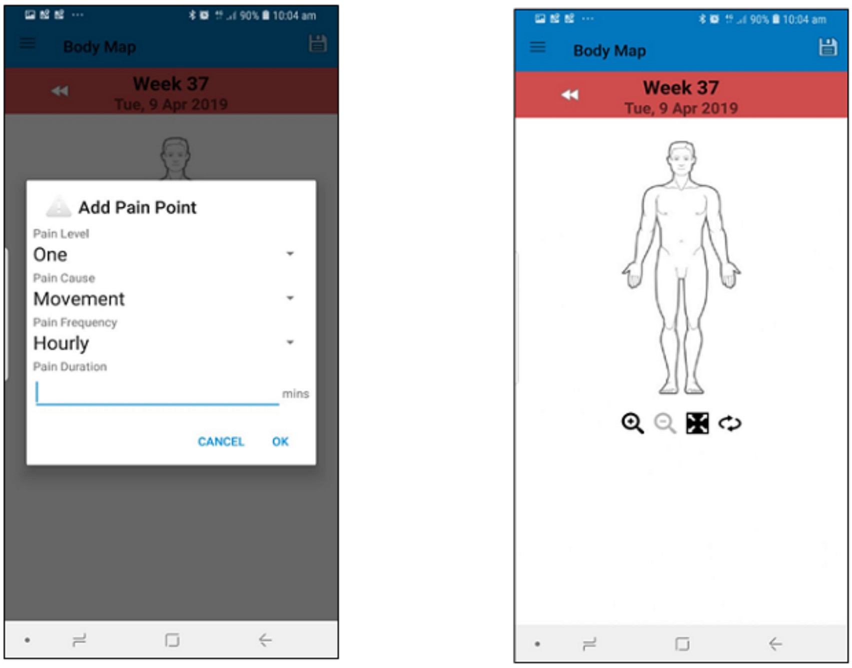
Figure 5. Educational materials.



Body Pain Map

Pain interferes with many daily activities, and the goal of pain management is to reduce the effect of pain on patient function and quality of life, including the ability to resume daily activity, maintain a positive mood, and get adequate sleep. The body map is a tool designed for patients to identify the location of their pain and score their pain level from 0 to 10. We developed the pain assessment tool to identify pain characteristics by asking questions about pain cause, intensity, aggravation factors, frequency, and duration (Figure 6).

Figure 6. Body pain map screen on Android phone.



Pain Level: One
Pain Cause: Movement
Pain Frequency: Hourly
Pain Duration: mins

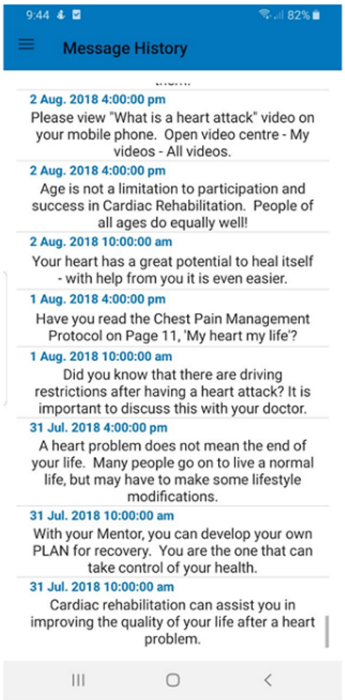
Relaxation Audio

To decrease patients’ emotional stress, the MoTER-ACS app contains relaxation audio. The audio was developed by the Australian Cancer Council, and authorization to use the audio in the MoTER-ACS intervention was obtained. The audio consists of 10 tracks, including exercises for relaxation, practical tips, and types of relaxation.

Motivational Messages

Previous research has shown that mHealth studies that reported low patient adherence used more basic and repetitive content, while successful studies used several educational and motivational strategies to engage users (ie, tailored or personalized messages) [27]. Based on health care professionals’ input, a collection of motivational messages was developed to deliver through SMS text messages throughout the program. Message content was adapted from a previous study that demonstrated positive effects on lifestyle behaviors [14]. Daily SMS text messaging aims to enhance patients’ self-management, competence, and relatedness. However, in this study, the messages were general and not tailored or personalized (Figure 7).

Figure 7. Motivational messages.



2 Aug. 2018 4:00:00 pm
Please view "What is a heart attack" video on your mobile phone. Open video centre - My videos - All videos.

2 Aug. 2018 4:00:00 pm
Age is not a limitation to participation and success in Cardiac Rehabilitation. People of all ages do equally well!

2 Aug. 2018 10:00:00 am
Your heart has a great potential to heal itself - with help from you it is even easier.

1 Aug. 2018 4:00:00 pm
Have you read the Chest Pain Management Protocol on Page 11, 'My heart my life'?

1 Aug. 2018 10:00:00 am
Did you know that there are driving restrictions after having a heart attack? It is important to discuss this with your doctor.

31 Jul. 2018 4:00:00 pm
A heart problem does not mean the end of your life. Many people go on to live a normal life, but may have to make some lifestyle modifications.

31 Jul. 2018 10:00:00 am
With your Mentor, you can develop your own PLAN for recovery. You are the one that can take control of your health.

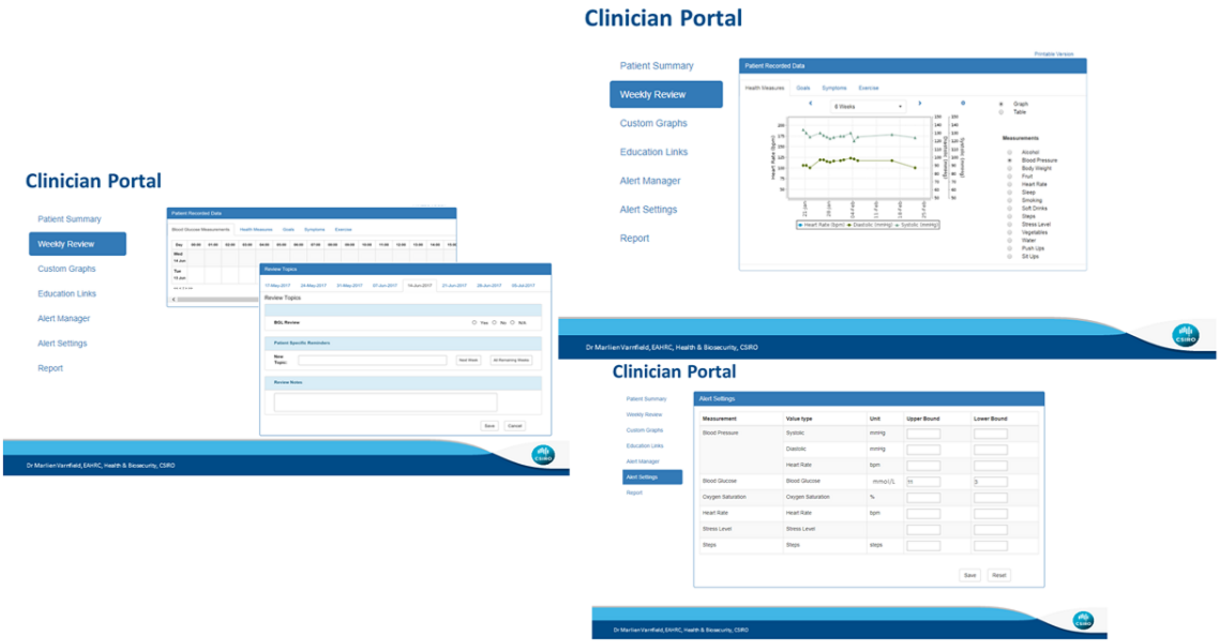
31 Jul. 2018 10:00:00 am
Cardiac rehabilitation can assist you in improving the quality of your life after a heart problem.

Web Portal

The web portal enables mentors to provide patients with real-time individualized coaching, feedback, support, and instructions related to cardiovascular symptoms. Data gathered from the MoTER-ACS smartphone app are linked to the web portal, and after reviewing patients’ physiological data and health measures, mentors can provide their feedback and support through phone calls or FaceTime (Figure 8).

The MoTER-ACS platform ensures user privacy by registering a user account into a database on a secure server. The web portal server is hosted in an Amazon Web Services Virtual Private Cloud managed by the Commonwealth Scientific and Industrial Research Organisation (CSIRO). The Android and iOS apps are publicly available via Google Play and the App Store for patients to download and register. However, smartphone- and web-based app functionalities are contingent on authentication with the secure web server in order to prevent unauthorized use of the MoTER-ACS platform, and all data transmission is encrypted.

Figure 8. MoTER-ACS web portal screenshot. MoTER-ACS: Mobile Technology–Enabled Rehabilitation for Patients With Acute Coronary Syndrome.



Measurement	Value type	Unit	Upper Bound	Lower Bound
Blood Pressure	Systolic	mmHg		
	Diastolic	mmHg		
	Heart Rate	bpm		
Blood Glucose	Blood Glucose	mmol/L	11	3
Oxygen Saturation	Oxygen Saturation	%		
Heart Rate	Heart Rate	bpm		
Stress Level	Stress Level			
Steps	Steps	steps		

Pilot Testing

MoTER-ACS provides an alternative model of care to a postdischarge clinic, and it is essential to determine how the intervention compares with traditional programs before it can be implemented as routine practice. Therefore, a pilot randomized controlled trial is planned to test the feasibility of the MoTER-ACS platform with a small group (N=54) of patients with ACS for 12 weeks. The trial will aim to test aspects of the platform and intervention content design, such as educational instructions, self-management interventions, and health care providers' support. The pilot study will be conducted based on the adopted framework and assess several outcome measures, such as participants' compliance and adherence to the intervention, modifiable cardiovascular risk factors, medication adherence, and health status. Based on the results of the pilot randomized controlled trial, patient survey, and focus group discussion (as a poststudy activity), the MoTER-ACS intervention will be modified and a subsequent larger randomized controlled trial with longer follow-up will be conducted to observe patients over a longer period.

Discussion

Principal Findings

This paper outlines the development and extension of an evidence- and theory-based mHealth intervention to support patients with ACS. Advanced sensor and smartphone technologies overcome common accessibility barriers that limit patient support. mHealth interventions can potentially meet the needs of patients who are unable or unwilling to attend outpatient clinics following discharge from the hospital [28]. MoTER-ACS provides educational instructions, self-management interventions, and health care providers' support. We aimed to close the gap between hospital- and home-based postdischarge programs by proposing an alternative model of health care delivery.

The incorporation of social and behavioral theory in apps is crucial in developing complex interventions [29]. However, there is a lack of theory-based mHealth interventions, even though some theories of behavior change are validated in evidenced-based interventions for primary and secondary prevention [30,31]. To develop the MoTER-ACS intervention, the self-efficacy theory (drawn from the social cognitive theory) was adopted to reinforce behavior change, including problem solving, goal setting, action planning, e-diaries, self-monitoring, educational instructions, role modeling, and health care providers' persuasion.

The prestudy surveys provided information on patients' perceived learning needs and health-related literacy, which was required for developing mHealth educational interventions. Prior to developing educational interventions, assessing and evaluating patients' learning needs is essential, and it is evidenced that adopting a structured educational plan increases the likelihood of a successful recovery [32]. Furthermore, engaging patients in the development phase and identifying their needs and preferences empowers their enthusiasm for learning and therefore assists health care providers in developing patient-centered interventions.

The focus group discussion with health care professionals resulted in useful feedback regarding content and features for the development of the MoTER-ACS intervention. The educational materials and self-management interventions identified from the prestudy surveys and systematic reviews [15,28] were incorporated into the MoTER-ACS platform to engage patients with the intervention and improve their health outcomes after discharge from the hospital. The MoTER-ACS app contains multimedia educational instructions and tools to support self-management. Educational materials with diet information could influence patient knowledge and provide guidance to follow healthy diet instructions, tips, and recipes.

Previous research has demonstrated the benefits of tracking health behaviors through mHealth. Self-monitoring, goal setting, and feedback are recommended for tracking because they are likely to increase patients' engagement in their personalized care and offer health care providers assessments of their patients' daily activity patterns [11]. A recent systematic review and meta-analysis of remotely delivered interventions using self-monitoring and feedback demonstrated a significant effect on patient behavior change [33]. Therefore, we considered symptom monitoring, goal setting, and feedback from health care providers within the MoTER-ACS intervention.

A systematic review of mHealth pain management showed that mobile apps are beneficial for patients, particularly those in outpatient clinics, and that both patients and health care providers were satisfied with apps that provided pain management tools [34]. Furthermore, previous research has shown that a smartphone pain diary facilitates gathering more accurate and complete pain ratings [35]. Hence, a body pain map for the MoTER-ACS app was developed to identify pain and its characteristics.

Motivational messages are delivered as part of the MoTER-ACS intervention. It is evidenced that 4 SMS text messages per week are effective in improving patients' health outcomes [36]. Additionally, personally tailored and interactive interventions seem to be more efficacious, especially when users can choose when to receive messages [36].

Although mHealth can potentially improve patients' self-management, adoption of such technologies by adults 50 years or older is limited by age-related barriers. Higher age, physical decline, comorbidities, and low health literacy have been identified as mHealth engagement barriers [37]. Despite older adults' interest in using mHealth, current evidence reports that their usage and adoption of such interventions are inconsistent [38]. Therefore, facilitators and barriers potentially influencing older adults' acceptance of mHealth must be considered [37]. In addition to providing flexible tools to engage patients via written, verbal, or video interactions, there is a need to consider how individuals without advanced technical skills will interact with the app or participate in mHealth interventions [39]. Previous research has shown that most participants expect in-person training on the use of the mHealth app in addition to on-demand online help, phone support, or support from family and friends [40-42]. Therefore, in-person training on using the mHealth intervention and the involvement of family and caregivers could be solutions to improve engagement with the MoTER-ACS intervention.

Strengths and Limitations

Well-designed educational strategies grounded in theory and contemporary evidence are crucial in the development of complex interventions [15]. We used a theoretical framework to form the important constructs of the MoTER-ACS intervention and fully integrated it into all parts of the platform. We conducted the prestudy surveys to enhance the intervention design before embarking on feasibility testing. However, the patient survey was limited by the recruitment of a small sample of 30 patients with ACS from a metropolitan hospital. This limits the generalizability of the study results compared with sampling from other settings, including regional and remote areas. A larger sample size and nonurban hospital may provide precise information about patients' learning needs and preferences; however, the results of this study are consistent with previous research that examined cardiovascular patients' educational needs [43].

The MoTER-ACS platform provides comprehensive educational and self-management interventions based on the secondary prevention components recommended by national guidelines. The MoTER-ACS app is designed for both Android and iOS smartphones. Moreover, the flexible platform architecture enables rapid integration of new smartphones and wearable and nonwearable sensor capabilities as they become available.

Learning to use a smartphone and health-measuring devices represents a potential barrier for older adults [37]. Although we aimed to design the MoTER-ACS app to be user-friendly, a dedicated training module is required to familiarize patients with the technology. However, some patients may not be able to overcome the technological barriers. Therefore, involving family and caregivers could be a solution to overcome these barriers [44].

Conclusion

The MoTER-ACS platform extends the capabilities of the previous MoTER platform and provides an alternative model of care for postdischarge follow-up of patients with ACS. If proven effective, this research would enable clinicians to overcome the accessibility barriers of traditional hospital-based programs by providing mHealth follow-up. Mobile technology provides an ideal platform for the delivery of health care services and could easily be applied to the prevention and management of other chronic diseases.

Acknowledgments

The authors would like to acknowledge the generous support received from the staff of the Department of Cardiology at Prince Charles Hospital, Queensland, Australia. The authors would also like to acknowledge the contribution by Professor Manny Noakes, former director of the CSIRO food and nutrition department, to the heart-healthy foods package in the MoTER-ACS app.

This study was conducted as part of a PhD project for NB and funded by the Australian e-Health Research Centre of the Commonwealth Scientific and Industrial Research Organisation. The project has received funding from the Prince Charles Hospital Foundation to conduct the pilot study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Patient Participants’ characteristics.
[DOCX File , 14 KB - formative_v4i12e17542_app1.docx]

Multimedia Appendix 2
Patient Participants’ responses- CPLNI.
[DOCX File , 12 KB - formative_v4i12e17542_app2.docx]

Multimedia Appendix 3
Healthcare providers’ characteristics.
[DOCX File , 14 KB - formative_v4i12e17542_app3.docx]

Multimedia Appendix 4
Cardiologists' survey.
[DOCX File , 153 KB - formative_v4i12e17542_app4.docx]

References

1. National Heart Foundation of Australia, The Cardiac Society of Australia and New Zealand. Guidelines for the management of acute coronary syndromes 2006. Med J Aust 2006 Apr 17;184(S8):27. [doi: 10.5694/j.1326-5377.2006.tb00304.x]
2. Australia's health 2018. Australian Institute of Health and Welfare. 2018. URL: https://www.aihw.gov.au/getmedia/7c42913d-295f-4bc9-9c24-4e44eff4a04a/aihw-aus-221.pdf [accessed 2020-12-01]
3. Chew D, French J, Briffa TG, Hammett CJ, Ellis CJ, Ranasinghe I, et al. Acute coronary syndrome care across Australia and New Zealand: the SNAPSHOT ACS study. Med J Aust 2013 Aug 05;199(3):185-191. [doi: 10.5694/mja12.11854] [Medline: 23909541]
4. Beswick A, Rees K, Griebsch I, Taylor F, Burke M, West R, et al. Provision, uptake and cost of cardiac rehabilitation programmes: improving services to under-represented groups. Health Technol Assess 2004 Oct;8(41):1-152. [doi: 10.3310/hta8410] [Medline: 15461879]
5. Neubeck L, Freedman SB, Clark AM, Briffa T, Bauman A, Redfern J. Participating in cardiac rehabilitation: a systematic review and meta-synthesis of qualitative data. Eur J Prev Cardiolog 2011 May 05;19(3):494-503. [doi: 10.1177/1741826711409326] [Medline: 22779092]
6. Kreindler SA. Watching your wait: evidence-informed strategies for reducing health care wait times. Qual Manag Health Care 2008;17(2):128-135. [doi: 10.1097/01.QMH.0000316990.48673.9f] [Medline: 18425026]
7. Collins F. How to fulfill the true promise of "mHealth": Mobile devices have the potential to become powerful medical tools. Sci Am 2012 Jul;307(1):16. [doi: 10.1038/scientificamerican0712-16] [Medline: 22779258]
8. Bashi N, Karunanithi M, Fatehi F, Ding H, Walters D. Remote Monitoring of Patients With Heart Failure: An Overview of Systematic Reviews. J Med Internet Res 2017 Jan 20;19(1):e18 [FREE Full text] [doi: 10.2196/jmir.6571] [Medline: 28108430]
9. Blasco A, Carmona M, Fernández-Lozano I, Salvador CH, Pascual M, Sagredo PG, et al. Evaluation of a Telemedicine Service for the Secondary Prevention of Coronary Artery Disease. J Cardiopulm Rehabil Prev 2012;32(1):25-31. [doi: 10.1097/hcr.0b013e3182343aa7] [Medline: 22113368]
10. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile Phone-Based Telemonitoring for Heart Failure Management: A Randomized Controlled Trial. J Med Internet Res 2012 Feb 16;14(1):e153 [FREE Full text] [doi: 10.2196/jmir.1909] [Medline: 26084979]
11. Chiauzzi E, Rodarte C, DasMahapatra P. Patient-centered activity monitoring in the self-management of chronic health conditions. BMC Med 2015 Apr 9;13(1):77. [doi: 10.1186/s12916-015-0319-2] [Medline: 25889598]
12. Wang J, Wang Y, Wei C, Yao N, Yuan A, Shan Y, et al. Smartphone Interventions for Long-Term Health Management of Chronic Diseases: An Integrative Review. Telemedicine and e-Health 2014 Jun;20(6):570-583. [doi: 10.1089/tmj.2013.0243] [Medline: 24787747]
13. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, Implementing, and Evaluating Mobile Health Technologies for Managing Chronic Conditions in Older Adults: A Scoping Review. JMIR mHealth uHealth 2016 Jun 09;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
14. Varnfield M, Karunanithi M, Lee C, Honeyman E, Arnold D, Ding H, et al. Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients: results from a randomised controlled trial. Heart 2014 Jun 27;100(22):1770-1779. [doi: 10.1136/heartjnl-2014-305783] [Medline: 24973083]

15. Bashi N, Fatehi F, Fallah M, Walters D, Karunanithi M. Self-Management Education Through mHealth: Review of Strategies and Structures. JMIR Mhealth Uhealth 2018 Oct 19;6(10):e10771 [FREE Full text] [doi: 10.2196/10771] [Medline: 30341042]
16. Guidelines for the management of absolute cardiovascular disease risk. National Vascular Disease Prevention Alliance. 2012. URL: https://www.heartfoundation.org.au/getmedia/4342a70f-4487-496e-bbb0-dae33a47fcb2/Absolute-CVD-Risk-Full-Guidelines_2.pdf [accessed 2020-11-24]
17. Bandura A. Self-efficacy: Toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191-215. [doi: 10.1037/0033-295x.84.2.191] [Medline: 847061]
18. Bashi N, Windsor C, Douglas C. Evaluating a Web-Based Self-Management Intervention in Heart Failure Patients: A Pilot Study. JMIR Res Protoc 2016 Jun 20;5(2):e116 [FREE Full text] [doi: 10.2196/resprot.5093] [Medline: 27324213]
19. Bandura A. Self-Efficacy Beliefs of Adolescents. Charlotte, NC: Information Age Publishing; 2006.
20. Galdeano L, Furuya RK, Rodrigues MA, Dantas RAS, Rossi LA. Reliability of the Cardiac Patients Learning Needs Inventory (CPLNI) for use in Portugal. J Clin Nurs 2014 Jun;23(11-12):1532-1540. [doi: 10.1111/j.1365-2702.2012.04158.x] [Medline: 22805321]
21. Gerard PS, Peterson LM. Learning needs of cardiac patients. Cardiovasc Nurs 1984;20(2):7-11. [Medline: 6570812]
22. Mattson CC, Rawson K, Hughes JW, Waechter D, Rosneck J. Health literacy predicts cardiac knowledge gains in cardiac rehabilitation participants. Health Educ J 2014 Feb 19;74(1):96-102. [doi: 10.1177/0017896914522029]
23. Hoover RS, Koerber AL. Using NVivo to Answer the Challenges of Qualitative Research in Professional Communication: Benefits and Best Practices Tutorial. IEEE Trans Profess Commun 2011 Mar;54(1):68-82. [doi: 10.1109/tpc.2009.2036896]
24. Bashi N, Hassanzadeh H, Varnfield M, Wee Y, Walters D, Karunanithi M. Multidisciplinary Smartphone-Based Interventions to Empower Patients With Acute Coronary Syndromes: Qualitative Study on Health Care Providers' Perspectives. JMIR Cardio 2018 Oct 31;2(2):e10183 [FREE Full text] [doi: 10.2196/10183] [Medline: 31758781]
25. Introduction to Program Evaluation for Public Health Programs: A Self-Study Guide. US Department of Health and Human Services Centers for Disease Control and Prevention. 2011. URL: https://www.cdc.gov/eval/guide/index.htm [accessed 2020-12-01]
26. Ahn J, Park J, Kim C. Effects of an individualised nutritional education and support programme on dietary habits, nutritional knowledge and nutritional status of older adults living alone. J Clin Nurs 2017 Oct 27;27(9-10):2142-2151. [doi: 10.1111/jocn.14068]
27. Park LG, Howie-Esquivel J, Dracup K. A quantitative systematic review of the efficacy of mobile phone interventions to improve medication adherence. J Adv Nurs 2014 Apr 01;70(9):1932-1953. [doi: 10.1111/jan.12400] [Medline: 24689978]
28. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth Chronic Disease Management on Treatment Adherence and Patient Outcomes: A Systematic Review. J Med Internet Res 2015 Feb 24;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
29. Glanz K, Bishop DB. The Role of Behavioral Science Theory in Development and Implementation of Public Health Interventions. Annu Rev Public Health 2010 Mar;31(1):399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
30. Ajzen I, Fishbein M. Understanding attitudes predicting social behavior. London, England: Pearson; 1980.
31. Jacobs MA, Graham AL. Iterative development and evaluation methods of mHealth behavior change interventions. Curr Opinion Psychol 2016 Jun;9:33-37. [doi: 10.1016/j.copsyc.2015.09.001]
32. Beagley L. Educating Patients: Understanding Barriers, Learning Styles, and Teaching Techniques. J Perianesth Nurs 2011 Oct;26(5):331-337. [doi: 10.1016/j.jopan.2011.06.002] [Medline: 21939886]
33. Teasdale N, Elhussein A, Butcher F, Piernas C, Cowburn G, Hartmann-Boyce J, et al. Systematic review and meta-analysis of remotely delivered interventions using self-monitoring or tailored feedback to change dietary behavior. Am J Clin Nutr 2018 Feb 01;107(2):247-256 [FREE Full text] [doi: 10.1093/ajcn/nqx048] [Medline: 29529158]
34. Thurnheer SE, Gravestock I, Pichierri G, Steurer J, Burgstaller JM. Benefits of Mobile Apps in Pain Management: Systematic Review. JMIR Mhealth Uhealth 2018 Oct 22;6(10):e11231 [FREE Full text] [doi: 10.2196/11231] [Medline: 30348633]
35. Garcia-Palacios A, Herrero R, Belmonte M, Castilla D, Guixeres J, Molinari G, et al. Ecological momentary assessment for chronic pain in fibromyalgia using a smartphone: A randomized crossover study. Eur J Pain 2013 Nov 22;18(6):862-872. [doi: 10.1002/j.1532-2149.2013.00425.x] [Medline: 24921074]
36. Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, et al. Effect of Lifestyle-Focused Text Messaging on Risk Factor Modification in Patients With Coronary Heart Disease. JAMA 2015 Sep 22;314(12):1255-1263. [doi: 10.1001/jama.2015.10945] [Medline: 26393848]
37. Wildenbos G, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: A literature based framework (MOLD-US). Int J Med Inform 2018 Jun;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
38. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: An extension of the UTAUT model. Int J Med Inform 2017 May;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]
39. Martin T. Assessing mHealth: opportunities and barriers to patient engagement. J Health Care Poor Underserved 2012 Aug;23(3):935-941. [doi: 10.1353/hpu.2012.0087] [Medline: 24212144]

40. Beatty AL, Magnusson SL, Fortney JC, Sayre GG, Whooley MA. VA FitHeart, a Mobile App for Cardiac Rehabilitation: Usability Study. JMIR Hum Factors 2018 Jan 15;5(1):e3 [FREE Full text] [doi: 10.2196/humanfactors.8017] [Medline: 29335235]
41. Nguyen T, Irizarry C, Garrett R, Downing A. Access to mobile communications by older people. Australas J Ageing 2014 Apr 17;34(2):E7-E12. [doi: 10.1111/ajag.12149] [Medline: 24750499]
42. Bashi N, Windsor C, Douglas C. Evaluating a Web-Based Self-Management Intervention in Heart Failure Patients: A Pilot Study. JMIR Res Protoc 2016 Jun 20;5(2):e116 [FREE Full text] [doi: 10.2196/resprot.5093] [Medline: 27324213]
43. Boyde M, Tuckett A, Peters R, Thompson DR, Turner C, Stewart S. Learning style and learning needs of heart failure patients (The Need2Know-HF patient study). Eur J Cardiovasc Nurs 2009 Dec;8(5):316-322. [doi: 10.1016/j.ejcnurse.2009.05.003] [Medline: 19520614]
44. Evans R. IMS Institute for Healthcare Informatics. Nurs Standard 2014 Apr 09;28(32):30-30. [doi: 10.7748/ns2014.04.28.32.30.s36]

Abbreviations

ACS: acute coronary syndrome
BP: blood pressure
BW: body weight
CPLNI: Cardiac Patients Learning Needs Inventory
CSIRO: Commonwealth Scientific and Industrial Research Organisation
CVD: cardiovascular disease
HF: heart failure
mHealth: mobile health
MI: myocardial infarction
MoTER: Mobile Technology–Enabled Rehabilitation
MoTER-ACS: Mobile Technology–Enabled Rehabilitation for Patients With Acute Coronary Syndrome
S-TOFHLA: Short Test of Functional Health Literacy in Adults

Edited by G Eysenbach; submitted 18.12.19; peer-reviewed by M Gonzalez-Garcia,, E Ross, R Kadri; comments to author 14.02.20; revised version received 27.05.20; accepted 03.10.20; published 18.12.20.

Please cite as:
Bashi N, Varnfield M, Karunanithi M
A Smartphone App for Patients With Acute Coronary Syndrome (MoTER-ACS): User-Centered Design Approach
JMIR Form Res 2020;4(12):e17542
URL: http://formative.jmir.org/2020/12/e17542/
doi:10.2196/17542
PMID:33337339

Original Paper

Stochastic Channel-Based Federated Learning With Neural Network Pruning for Medical Data Privacy Preservation: Model Development and Experimental Validation

Rulin Shao[1], BA; Hongyu He[2], BA; Ziwei Chen[3], PhD; Hui Liu[4], BA; Dianbo Liu[5], PhD

[1]Department of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China

[2]Department of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China

[3]Beijing Jiaotong University, Beijing, China

[4]Department of Mathematics, Mianyang Vocational College, Mianyang, China

[5]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, United States

Corresponding Author:
Dianbo Liu, PhD
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
77 Massachusetts Ave
Cambridge, MA, 02139
United States
Phone: 1 6177101859
Email: dianbo@mit.edu

Abstract

Background: Artificial neural networks have achieved unprecedented success in the medical domain. This success depends on the availability of massive and representative datasets. However, data collection is often prevented by privacy concerns, and people want to take control over their sensitive information during both the training and using processes.

Objective: To address security and privacy issues, we propose a privacy-preserving method for the analysis of distributed medical data. The proposed method, termed stochastic channel-based federated learning (SCBFL), enables participants to train a high-performance model cooperatively and in a distributed manner without sharing their inputs.

Methods: We designed, implemented, and evaluated a channel-based update algorithm for a central server in a distributed system. The update algorithm will select the channels with regard to the most active features in a training loop, and then upload them as learned information from local datasets. A pruning process, which serves as a model accelerator, was further applied to the algorithm based on the validation set.

Results: We constructed a distributed system consisting of 5 clients and 1 server. Our trials showed that the SCBFL method can achieve an area under the receiver operating characteristic curve (AUC-ROC) of 0.9776 and an area under the precision-recall curve (AUC-PR) of 0.9695 with only 10% of channels shared with the server. Compared with the federated averaging algorithm, the proposed SCBFL method achieved a 0.05388 higher AUC-ROC and 0.09695 higher AUC-PR. In addition, our experiment showed that 57% of the time is saved by the pruning process with only a reduction of 0.0047 in AUC-ROC performance and a reduction of 0.0068 in AUC-PR performance.

Conclusions: In this experiment, our model demonstrated better performance and a higher saturating speed than the federated averaging method, which reveals all of the parameters of local models to the server. The saturation rate of performance could be promoted by introducing a pruning process and further improvement could be achieved by tuning the pruning rate.

(JMIR Form Res 2020;4(12):e17265) doi:10.2196/17265

KEYWORDS

federated learning; differential privacy preserving; neural network pruning; health care; privacy; medical data; machine learning; neural network

Introduction

Medical Data Privacy

Medical data analysis in health care brings many benefits and holds great promise for transforming the field. With the help of a wide range of health care networks, health care organizations are now able to analyze a vast volume of data with great variety and velocity to support decision making [1-3]. In addition, automated machine-learning algorithms could facilitate patients and physicians to make better informed choices by providing empirical estimates based on gigabytes of data [4]. Apart from decision support, medical data analysis could also promote analytical capability for patterns of use, analysis of unstructured data, predictive capability, and traceability [3].

However, health care data security and privacy issues have raised broad ethical and legal concerns in recent years given the sensitive nature of health information [5]. Health care research often involves studies of a large amount of data collected from various sources such as pharmacies, insurance companies, government agencies, and research institutions. For instance, to discover new drugs or assess a new therapy, a research institute may need clinical records provided by hospitals' autonomous databases [6]. This direct sharing of medical data is likely to violate individual privacy and expose data owners to the threat of illegal data collection [7].

To address this data privacy concern, different countries have enacted different legislations and policies [8,9], which impose limitations on data collection and utilization for health care research. Over the years, many traditional methods for privacy preserving have been proposed, such as deidentification [10,11], a hybrid execution model [12], and identity-based anonymization [13]. However, as pointed out by several authors, these methods alone could not guarantee the anonymity and security of medical data [14-16]. Recently developed machine-learning methods require considerable data to acquire models with sufficient accuracy [17]. To leverage massive and diverse datasets and promote machine-learning models, the issue of balancing privacy and regulatory requirements has to be addressed [18].

Federated Learning

In conventional deep learning, all training data are shared with a central server that performs the analysis. Having no control over this process, the clients that contribute the data may have to upload their sensitive information to the server without a guarantee of its security of privacy. Furthermore, the learned model is generally not directly available to the client so that they have to reveal the inputs to the cloud when using the model [19], risking privacy leakage in both the training and using processes. Federated learning can address this problem by introducing some algorithmic techniques that distribute the learning process to local devices so that the clients could keep their data private and obtain a local model for future use.

Federated optimization has been studied by Konečný et al [20,21] for distributed optimization in machine learning. This work introduced a setting for distributed optimization in which none of the distinctive assumptions [21] is satisfied, making federated learning a feasible alternative to other methods. The proposed framework is different from conventional distributed machine learning [22-27] owing to the huge number of clients, extremely unbalanced/nonindependent and identically distributed data obtainable for each client, and poor network connections [28]. To address the latter constraint, Konečný et al [28] proposed two approaches to reduce the uplink communication costs: structured updates and sketched updates. McMahan et al [29,30] advocated for federated stochastic gradient descent (SGD) and federated averaging algorithms as feasible approaches for the federated learning of neural networks based on iterative model averaging. As an alternative to protecting a single data point's contribution to a learning a model [31], Geyer et al [32] proposed an algorithm for client-sided federated optimization to hide the specific contributions of individual clients during training. Further, methods to strengthen the reliability of federated learning, such as secure aggregation [33], essentially need synchronization on a rigid set of devices so that only a simple summation of the updates from users is consumed by the server side of the algorithm [34]. Applications based on federated learning algorithms have been proposed in several domains, ranging from content suggestions [35] to next-word prediction [36]. Bagdasaryan et al [37] focused on the vulnerability of federated learning. This work showed that the federated learning algorithm is vulnerable to a model-poisoning attack, which is different from poisoning attacks that target only the training data.

Besides the direct leakage of privacy mentioned above, participants in the distributed system may indirectly reveal some information about sensitive data via the weights uploaded to the server in the training process.

To address both direct and indirect privacy leakage of health care data, we developed the stochastic channel-based federated learning (SCBFL) method, which enables local participants to manipulate their data confidentially while benefitting the model's performance from the server with only a small proportion of the locally trained gradients revealed stochastically to the central model.

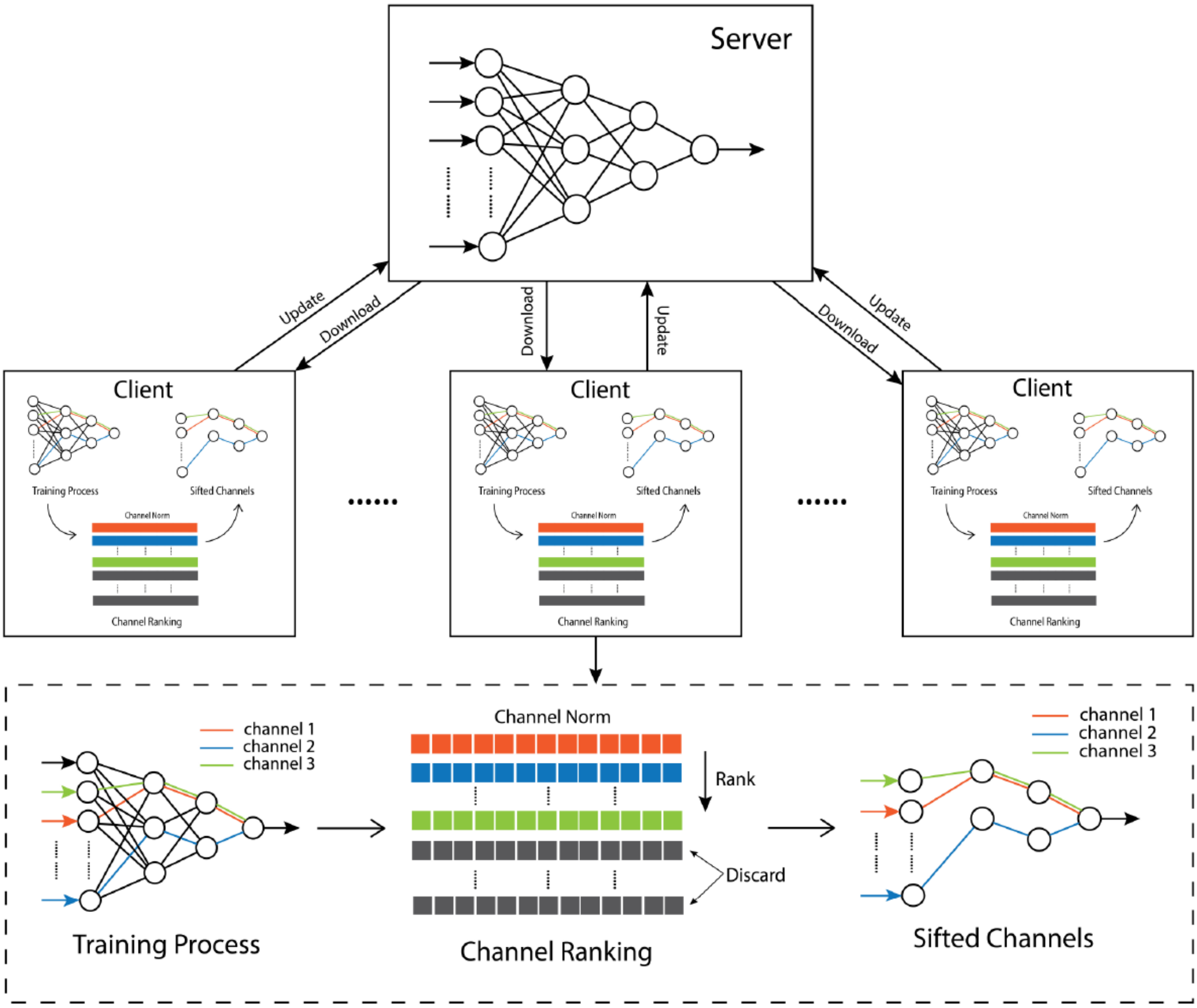
Methods

Principle of the SCBFL Approach

Based on the observation that different features do not contribute equally to the training process and that the importance of each feature may vary from one dataset to another, SCBFL (Figure 1) was developed as a privacy-preserving approach that seizes the most vital information from the local training results only by uploading a small fraction of gradients stochastically. The intuition behind this method is that the biological neural circuit follows the law of use and disuse, and the strongest neurons for an individual are those that constitute an active circuit in the learning process, suggesting that the neurons in one artificial neural network are not independent throughout a specific training process. Thus, we could consider the collaborative effect of neurons in each channel (similar to a biological neural circuit) when selecting parameters for a server update: if a channel of neurons changes substantially in a training loop, we can assume that it is a strong neural circuit in the network, corresponding to a sensitive feature in the input sets, whereas the neural channels showing little change in one training loop should be regarded as deteriorated channels, whose information could be kept private with minimal effect on the server's final performance. Choosing the channels with the most substantial variation enables SCBFL to only upload a small percentage of the gradients in each training loop while achieving comparable accuracy to the federated averaging method without uploading the integrate local weights to the server, as will be demonstrated in the Results section.

The update algorithm plays an essential role in SCBFL. In each global loop, SCBFL computes the norms of channels in gradients that result from the local training process, calculates the α-percentile of the norms, and then sifts out the channels with greater variation in the gradients than the percentile, in which α is the update rate set by the local participant. The sifted parameters are then used for the server update.

Figure 1. Schematic of the stochastic channel-based federated learning (SCBF) model.



To facilitate the description of the algorithm, suppose there are N features as input and an L-layer deep neural network is conducted with m_1, m_2,…, m_L neurons in each layer. For convenience, we denote $m_0=N$ as the input dimension, and we denote the weight matrix as $W=[W_1, W_2,\ldots W_L]$ and the bias matrix as $B=[B_1, B_2,\ldots B_L]$. The shapes of the weight matrix and bias matrix could be expressed as follows:

Where $l=1,2,\ldots L$, and m_l is the number of neurons in the lth layer.

The update algorithm includes five steps:

1. Train the local model: The local models are trained separately on their own datasets, and each model results in a gradient matrix showing the change in the weight matrix during each training loop. The gradient matrix G has the same shape as the weight matrix W. Since the influence from the bias matrix is negligible compared to that of the weight matrix, the changes in bias are omitted for the sake of efficiency.

2. Compute channel norms: Considering that a channel must go through a neuron in each layer and correlates to an

L-dimensional vector comprising the index of these neurons, the results of a channel's norm could be saved in an *L*-dimensional tensor *T*, each element of which equals a channel norm. The shape of *T* should be:

In addition, is the *i*th channel, in which is the index of the tensor that correlates the neurons this channel goes through in each layer. The Euclidean norm of each channel is calculated by

and is saved in the *L*-dimensional tensor *T*:

3. Sort norms: Given a fixed upload rate α (also referred to as "update rate" in this paper), we could straighten the gradient tensor to a vector and sort it, computing the α-quantile q_α as a threshold for the channel selection.

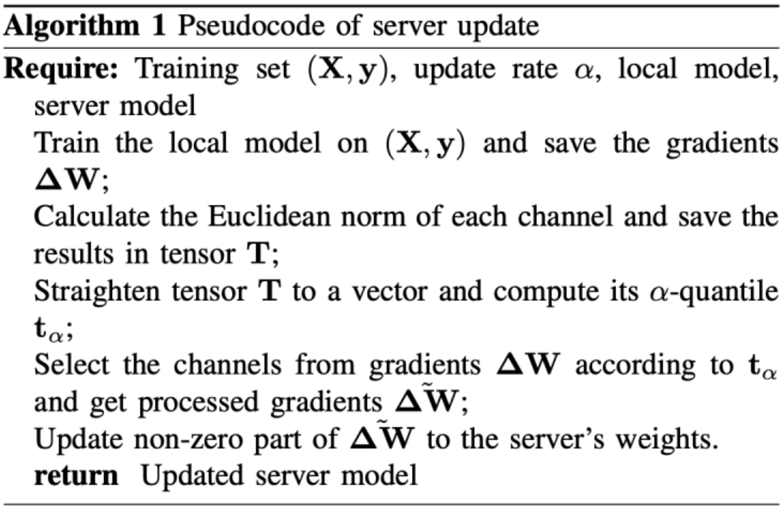
4. Process gradients: There are two ways to process the gradients. With negative selection, the channels with norms below the α-quantile are discarded and the remaining parameters are selected for the update. With positive selection, the channels with a norm above q_α are selected and the remaining parameters are set to zeros.

In our preliminary trials, both selection methods worked well. Considering that different neural channels may include the same neurons, positive selection tends to behave better than negative selection due to the preference to upload more parameters with the same update rate. Taking positive selection as an example, for each element $T_{t1,\ t2,\ldots,tL}$ in tensor *T*, which corresponds to a specific channel, the gradients are processed with respect to the rank of this channel's norm, as shown in the following form:

5. Update server: Finally, the processed gradient matrix G˜ is uploaded to the server, and the server will update its parameters by adding gradients G˜ to its original weights (Figure 2).

The server update algorithm is executed every global loop, and our experiment showed that with only 10% of local channels revealed, the server could have comparative performances to those of the federated averaging methods with higher speed to reach saturation. Before the next training loop begins, the local model downloads the server's latest weights. The download rate was set to 100% since we suppose that the server weights could be shared publicly, which could be adjusted according to the application scenarios.

Figure 2. Pseudocode of the server update.



```
Algorithm 1 Pseudocode of server update
Require: Training set $(\mathbf{X}, \mathbf{y})$, update rate $\alpha$, local model, server model
  Train the local model on $(\mathbf{X}, \mathbf{y})$ and save the gradients $\Delta\mathbf{W}$;
  Calculate the Euclidean norm of each channel and save the results in tensor $\mathbf{T}$;
  Straighten tensor $\mathbf{T}$ to a vector and compute its $\alpha$-quantile $\mathbf{t}_\alpha$;
  Select the channels from gradients $\Delta\mathbf{W}$ according to $\mathbf{t}_\alpha$ and get processed gradients $\Delta\tilde{\mathbf{W}}$;
  Update non-zero part of $\Delta\tilde{\mathbf{W}}$ to the server's weights.
  return Updated server model
```

Pruning Process

Training a model with privacy-preserving methods could be time-consuming, especially when the training sets are enormous. To address this problem, we introduced a neural network pruning process to SCBFL that could prune off the redundant nodes in the neural network based on the validation set, thus saving a substantial amount of time. This work is done circularly in the first several global loops until the distributed system reaches a suitable scale, so that SCBFL with pruning (SCBFLwP) learns from the datasets more efficiently.

Neural network pruning (Figure 3) is not a novel concept. Yang [38] proposed a method to prune connections based on the magnitude of weights. He et al [39] used a channel-pruning method to accelerate a deep convolutional neural network. Han et al [40] introduced a growing-and-pruning approach for a fuzzy neural network. Moreover, Srinivas [41] proposed a systematic method to prune one neuron at a time, addressing the problem of pruning parameters in a trained neural network model.

Given the fact that each neural network has a computation process consisting of multiplication, addition, and activation, neurons whose output consists mostly of zeros may have little effect on the output of subsequent layers, not to mention the final results [31]. Removing these redundant nodes from the model will do little harm to the accuracy of the network but will save abundant execution time.

Average percentage of zeros (APoZ) [42], which measures the percentage of zeros in the activations of a neuron under rectified linear unit (ReLU) mapping, is used to evaluate the redundancy of neurons in the network. is denoted as the output of the *c*th neuron in the *i*th layer. Let *M* denote the output dimension and *N* denote the total quantity of validation examples. of the *c*th neuron in the *i*th layer is then defined as:

where f(·)=1 if true and f(·)=0 if false.

SCBFLwP (Figure 4) then decides which neurons will be pruned according to APoZ using validation sets: those having the highest APoZ will be pruned, the number of which is a fixed percentage of the total number of neurons left in each global loop.

Figure 3. Neural network pruning.

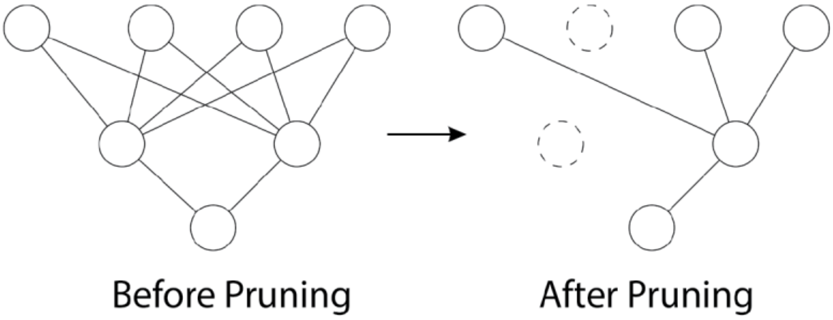
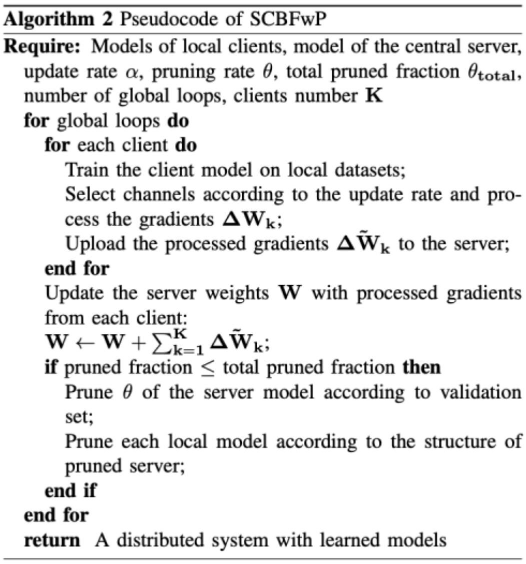


Figure 4. Pseudocode of stochastic channel-based federated learning with pruning (SCBFwP).



```
Algorithm 2 Pseudocode of SCBFwP
Require: Models of local clients, model of the central server,
  update rate $\alpha$, pruning rate $\theta$, total pruned fraction $\theta_{\mathbf{total}}$,
  number of global loops, clients number $\mathbf{K}$
  for global loops do
    for each client do
      Train the client model on local datasets;
      Select channels according to the update rate and process the gradients $\mathbf{\Delta W_k}$;
      Upload the processed gradients $\mathbf{\Delta \tilde{W}_k}$ to the server;
    end for
    Update the server weights $\mathbf{W}$ with processed gradients from each client:
    $\mathbf{W} \leftarrow \mathbf{W} + \sum_{\mathbf{k=1}}^{\mathbf{K}} \mathbf{\Delta \tilde{W}_k}$;
    if pruned fraction $\leq$ total pruned fraction then
      Prune $\theta$ of the server model according to validation set;
      Prune each local model according to the structure of pruned server;
    end if
  end for
  return A distributed system with learned models
```

Distributed Learning Setting

We propose a privacy-preserving federated learning method based on the neural network. Federated learning could be executed on a distributed system such as a mobile device to achieve collaborative deep-learning goals with little risk of privacy leaks. Each device trains its model on the local dataset for several epochs in each global loop and only stochastically uploads a small percentage of the model weights to the server to achieve good performance in the server without sharing the local data or the overall model weights.

In our trial, we implemented a distributed system with 5 clients contributing to one server. Preliminary experiments were conducted to determine the proper structure for the proposed model. Through manual tuning, we found that the model achieves the best performance with high efficiency using 3 layers. Therefore, for each local client, we constructed an artificial neural network for binary prediction of mortality with 3 fully connected layers including 64, 32, and 1 neuron in the corresponding layers using ReLU activation at hidden layers and sigmoid activation at the output layer. We also added a dropout layer between the second and third hidden layers to reduce overfitting.

Regarding the parameters of communication between server and clients, the download rate was set to 100% for each client model, supposing the parameters of the server are shared publicly. The update rate was set to 30% for both the channel-based federated learning method and distributed selective SGD method. To enhance the influence of the latest update parameters, we chose 0.8 as the decay rate. As for the training process, we trained each model for 100 global loops and 5 epochs in each loop with the batch size set to 32.

We used the SGD algorithm to optimize our neural networks. Concerning the configuration of SGD, the learning rate is a hyperparameter that controls how much to adjust the model in response to the estimated error each time when the model weights are updated. Our experiments on testing models with various learning rates suggested that the proper learning rate was around 0.01 to guarantee both good performance and stable results.

In addition to the configuration of the model, the importance of performance measurement has long been recognized. With respect to the assessment of a classification model, area under the receiver operating characteristic curve (AUC-ROC) and area under the precision recall curve (AUC-PR) are reliable metrics: the higher the AUC value, the better the model is at distinguishing between patients in terms of mortality and survival.

Dataset for the Experiment

The data used in our experiment were provided by hospitals, comprising 30,760 admissions with status information represented by alive or dead. To explore the relationship between mortality and admissions, we developed a model that takes the medications as inputs and predictions of binary mortality as the output. The cohort was managed for 2917 different medications in total. Information on whether a patient took each of the medications after admission was adopted as a binary input feature. We used 60% of the dataset for training, 10% as the validation set, and 30% as the test set. The training set was equally divided into five parts as local training sets.

Statistical Analysis

The performances of models were evaluated by the AUC-ROC and AUC-PR, which are both typically used for measuring the performance of a classifier. To better understand the ROC curve, the concept of a confusion matrix first needs to be introduced.

A confusion matrix is a table consisting of four different combinations of prediction and ground truth, which are the true positive (TP), false positive (FP), false negative (FN), and true negative (TN). TP means that the model predicts the sample as positive and it is in fact positive. The value of TP can be calculated by counting the number of correct positive predictions. The other three parameters can be interpreted in a similar manner. With the help of a confusion matrix, more performance indicators can be defined, including the true positive rate (TPR), also known as recall and sensitivity, the false positive rate (FPR), and precision, which are calculated as follows:

TPR=TP/TP+FN **(8)**

FPR=FP/TN+FP **(9)**

Precision=TP/TP+FP **(10)**

The ROC curve is plotted with TPR against FPR at various classification thresholds, where TPR is on the y-axis and FPR is on the x-axis. Lowering the classification threshold means that the model will predict more samples as positive, thus increasing both FPR and TPR. As an alternative to ROC, the PR curve is plotted with precision against recall (TPR) at various classification thresholds. When datasets are imbalanced or skewed, the PR curve is a preferred alternative to the ROC curve. Both curves provide a visualization of model performance at different thresholds, and the AUC measures the entire two-dimensional area under these curves, providing an aggregate measurement of performance across all possible thresholds. Ranging in value from 0 to 1, AUC-ROC and AUC-PR can be interpreted as the possibility that the model ranks a positive sample more highly than a negative sample.

Results

Performance of the SCBFL Model

The update rate controls the number of selected channels whose nonzero part is uploaded to the server in each global loop, playing a vital role in affecting the final performance. To choose a suitable update rate for our distributed system, we implemented SCBFL models with different update rates ranging from 10% to 100%. Neural network pruning was used in this step to accelerate the training process. The performances are plotted in the first row of Figure 5. The result showed that even with 10% of the channels uploaded to the server, the SCBFL model achieved an AUC-ROC of 0.9776 and an AUC-PR of 0.9695, which outperformed the model that shared all of the parameters with the server. In addition, using a wide range of upload rates only led to a 0.01319 amplitude change in AUC-ROC and a 0.02739 amplitude change in AUC-PR, which facilitated the configuration process with stably high performance.

Figure 5. Performances of stochastic channel-based federated learning (SCBF) models. The top two graphs show the performances of SCBF with pruning (SCBFwP) using different update rates, and the bottom two graphs show the comparison between the performances of SCBF and federated averaging (FA). The left column shows the area under the receiving operating characteristic curve (AUC-ROC) and the right column shows the area under the precision-recall curve (AUC-PR) as performance metrics.

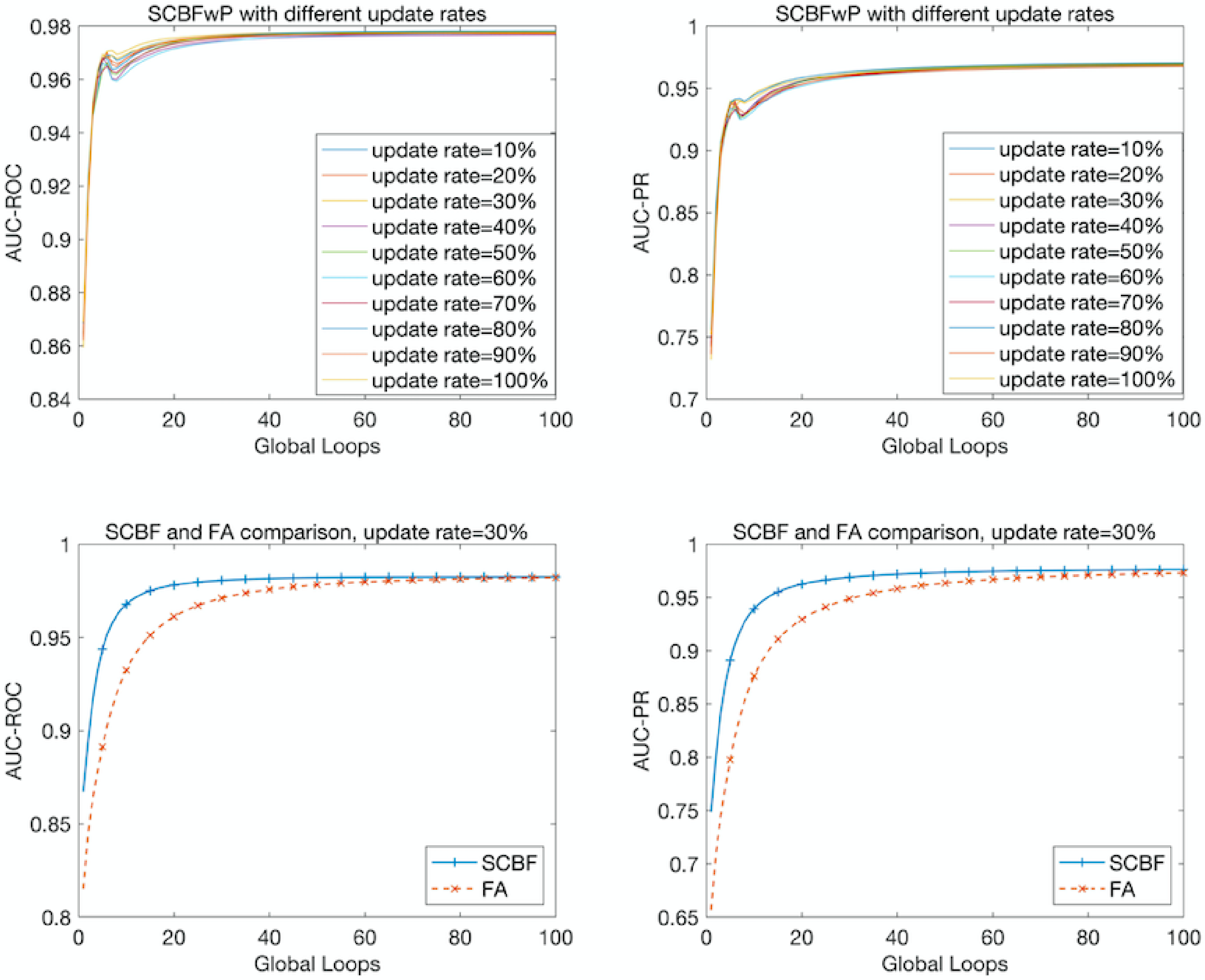


Table 1 compares the effectiveness of the SCBFL method with that of federated averaging, which is widely used in distributed systems and implements the federated learning by averaging the gradients obtained from local training processes [36]. For this comparison, we set the update rate to 30% for SCBFL and conducted both methods for federated learning on the same datasets for 100 global loops without pruning. As shown in Figure 5, our model reached saturation at the 20th global loop, which was faster than that obtained with federated average, which reached saturation at the 60th global loop. The performance of SCBFL consistently exceeded that of federal averaging. In the 4th global loop, SCBFL achieved a 0.05388 higher AUC-ROC and 0.09695 higher AUC-PR than those of federated averaging. After 100 global loops, the AUC-ROC and AUC-PR of SCBFL was 0.0033 and 0.0032 higher than that of federated averaging, respectively.

Table 1. Saturated performances of stochastic channel-based federated learning compared with federated averaging.

Method	AUC-ROC[a]	AUC-PR[b]
SCBFL[c]	0.9825	0.9763
Federated averaging	0.9821	0.9731

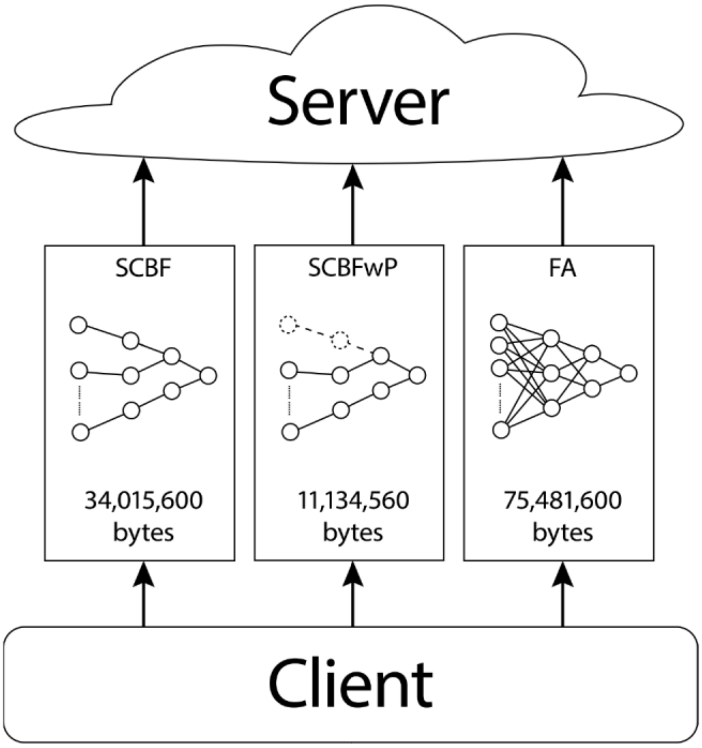
[a] AUC-ROC: area under the receiver operating characteristic curve.

[b] AUC-PR: area under the precision-recall curve.

[c] SCBFL: stochastic channel-based federated learning.

As shown in Figure 6, when the upload rate for the channels was set to 30%, 45% of the parameters were uploaded to the server using positive selection. With half of the parameters unrevealed to the server, the model achieved better performance and higher saturating speed.

Figure 6. Trans-information for upload processes using different methods. The stochastic channel-based federated learning with pruning (SCBFwP) method could save 85% of the trans-information compared to federated averaging (FA), and stochastic channel-based federated learning (SCBF) could save 55% compared with FA.



Performance of SCBFLwP

To speed up the training process and reduce the size of the neural network, we conducted network pruning for several loops after pretraining the model. In our trials, we set the pruning rate for each global loop to 10%, which represents the proportion of neurons to be pruned in the training loop. The total proportion of neurons to be pruned in the first several loops was set to 47%, which determines the final scale of the pruned model. Table 2 summarizes the performance of the SCBFLwP method for different update rates. Figure 7 compares the AUC-ROC and AUC-PR values of the SCBFL and federated averaging models with and without pruning.

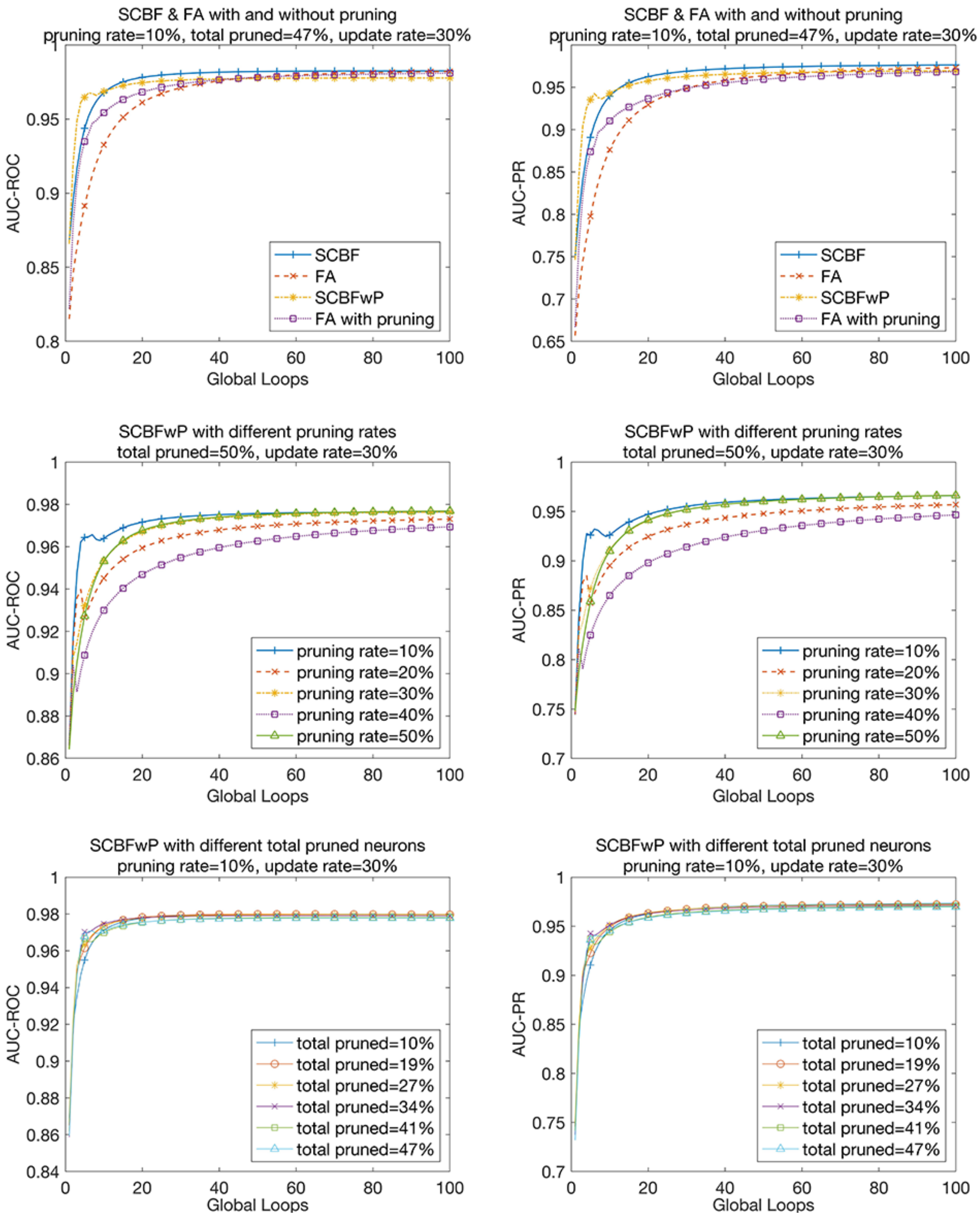
Table 2. Saturated performances of stochastic channel-based federated learning with pruning with different update rates.

Update rate	AUC-ROC[a]	AUC-PR[b]
10%	0.9776	0.9695
20%	0.9772	0.9686
30%	0.9777	0.9697
40%	0.9768	0.9604
50%	0.9780	0.9695
60%	0.9774	0.9682
70%	0.9774	0.9688
80%	0.9781	0.9703
90%	0.9774	0.9676
100%	0.9775	0.9685

[a] AUC-ROC: area under the receiver operating characteristic curve.

[b] AUC-PR: area under the precision-recall curve.

Figure 7. Performance of stochastic channel-based federated learning with pruning (SCBFwP). The top two graphs show the comparison between stochastic channel-based federated learning (SCBF) and federating averaging (FA) with and without pruning. The middle graphs show the performance of SCBFwP with different pruning rates. The performances of SCBFwP with different numbers of total pruned neurons are shown in the bottom two graphs. The left column shows the area under the receiver operating characteristic curve (AUC-ROC) and the right column shows the area under the precision-recall curve (AUC-PR) as performance metrics.



As shown in Table 3, the AUC-ROC for SCBFL with pruning was reduced by 0.0048 and the AUC-PR was reduced by 0.006814. There was a reduction of 0.0012 in AUC-ROC and of 0.0047 in AUC-PR for the federated averaging method compared to federated averaging with pruning.

Table 3. Saturated performances of stochastic channel-based federated learning (SCBFL) and federated averaging (FA) with and without pruning.

Methods	AUC-ROC[a]	AUC-PR[b]
FA	0.9821	0.9731
SCBFL	0.9825	0.9763
FAwP[c]	0.9809	0.9683
SCBFLwP[d]	0.9776	0.9694

[a] AUC-ROC: area under the receiver operating characteristic curve.

[b] AUC-PR: area under the precision-recall curve.

[c] FAwP: federating averaging with pruning.

[d] SCBFLwP: stochastic channel-based federated learning with pruning.

Moreover, the best performance was achieved by SCBFL after 100 loops of training with an AUC-ROC of 0.9825 and an AUC-PR of 0.9763 (Table 3). The highest evaluation in the first 5 loops was obtained by the SCBFL model with pruning.

To assess the stability of our model with the pruning rate and total pruned fraction (also called the total pruned rate), we executed the models of SCBFLwP controlling the variate. First, we fixed the total pruned fraction to 50% and ran the programs with different pruning rates ranging from 10% to 50% (Table 4). As shown in Figure 7, with an increase in the pruning rate, the final performance improved and saturated more quickly under most circumstances. However, there were also exceptions with higher performances at a 10% pruning rate for both AUC-ROC and AUC-PR, and lower performances at a 40% pruning rate for AUC-PR. In the graphs in the bottom row of Figure 7, we fixed the pruning rate to 10% and executed pruning for different times ranging from 1 to 6 (Table 5). The total pruned fractions were calculated and are annotated in the corresponding labels.

Table 4. Saturated performances of stochastic channel-based federated learning with pruning when the total pruned proportion was fixed and the pruning rate for each training loop changed.

Pruning rate/loop	AUC-ROC[a]	AUC-PR[b]
10%	0.9765	0.9661
20%	0.9730	0.9568
30%	0.9763	0.9662
40%	0.9693	0.9465
50%	0.9769	0.9663

[a] AUC-ROC: area under the receiver operating characteristic curve.

[b] AUC-PR: area under the precision-recall curve.

Table 5. Saturated performances of stochastic channel-based federated learning with pruning when the pruning rate for each training loop was fixed and the total pruned proportion changed.

Total pruned proportion	AUC-ROC[a]	AUC-PR[b]
10%	0.9769	0.9731
19%	0.9797	0.9722
27%	0.9795	0.9725
34%	0.9789	0.9714
41%	0.9781	0.9703
47%	0.9778	0.9697

[a] AUC-ROC: area under the receiver operating characteristic curve.

[b] AUC-PR: area under the precision-recall curve.

As shown in Figure 6, the SCBFLwP could save 85% of the trans-information compared to federated averaging. For SCBFL, when the upload rate for channels was set to 30%, 45% of the parameters were uploaded to the server using positive selection.

Running Time

SCBFL preserves the privacy of data by adding a channel-based upload algorithm, which will lead to an increased burden of calculations when applied to a complex neural network. However, this problem could be addressed by introducing a pruning process in several global loops. To illustrate this, the times consumed by SCBFL and federated averaging before and after pruning described in the last section are compared in Table 6. The pruning process could reduce 57% of the time for SCBFL and 48% of the time consumed by federated averaging.

Table 6. Time consumed by stochastic channel-based federated learning (SCBFL) and federated averaging with and without pruning.

Methods	Time (seconds)
Federated averaging	8679
Federated averaging with pruning	4508
SCBFL	19,696
SCBFL with pruning	8469

Table 7 shows that models with lower update rates tended to consume less time than those with larger update rates, indicating that choosing a lower rate for the update could better preserve the privacy as well as save time.

Table 7. Time consumed by stochastic channel-based federated learning with pruning with different update rates.

Update rate	Time (seconds)
10%	8339
20%	8545
30%	8469
40%	7987
50%	8359
60%	12,577
70%	9278
80%	11,462
90%	13,169
100%	13,030

Table 8 shows that different pruning rates for each global loop can equally save time. In addition, the model will consume more time if the number of pruned neurons is too small due to the executing time of the pruning process. With a fixed pruning rate, the time consumed by the model tended to decrease by reducing the model size.

Table 8. Time consumed by stochastic channel-based federated learning with pruning using different pruning rates for each loop or different total pruned proportions.

Pruning parameter	Time (seconds)
Pruning rate/loop	
10%	11,144
20%	8561
30%	11,852
40%	8389
50%	12,000
Total pruned	
10%	25,755
19%	22,717
27%	17,579
34%	15,909
41%	8050
47%	8469

Discussion

Principal Results

The proposed SCBFL method computes the norms of channels in gradients resulting from the local training output after each global loop, calculates the α-percentile of the channel norms, and then sifts out the channels that have greater variation in the gradients compared with the percentile for the server update. In this method, the server seizes the information from the uploading channels with the biggest variation, achieving comparative performance to the state-of-the-art method (federated averaging), which has to convey the entire local weights to the server when updating. Figure 1 shows the relationship between the server and clients, and demonstrates the process of the server update. This confirms the intuition behind SCBFL: the importance of a feature differs when training on different datasets, and thus important information can be extracted from the channels through which features with the greatest variation pass. We could infer that less than 10% of the channels contain the most fundamental information and that ignoring the remaining information does little harm to the learning of models.

It is important to train a small-scaled deep-learning model with high processing speed. The results showed that network pruning could speed up the training process and accelerate convergence while maintaining higher performance. As expected, pruning 47% of the neurons from the network decreased the final performance due to the simplified model structure. The reduction in performance is negligible in many application situations but the acceleration in both saturating and training speed is quite beneficial, as discussed further below. Overall, these results demonstrate that SCBFL is a reliable choice for federated learning, and that the SCBFLwP method might be a better choice when a quicker saturating speed is desired.

The graphs in the first row of Figure 7 show an obvious decline in the performance of SCBFLwP, which indicates an overpruned phenomenon for our trials. This indicates a tradeoff between time efficiency and the final accuracy. However, by tuning the pruning rate for each global loop and the total pruned rate of the model, we could achieve better performance. This is because if only the redundant neurons are pruned, the model could promote its learning efficiency without retaining useless information.

Figure 7 also shows that the performance of SCBFL improved when the times of pruning were reduced. The results with a fixed pruning rate were more stable than those with a fixed total pruned rate, indicating that more attention should be paid to the selection of the pruning rate for each step when building models, and it is stable for a SCBFL model to adjust the times of neural network pruning. Therefore, after choosing a suitable pruning rate, we could appropriately increase the loops in which the model was pruned to shorten the execution time with little effect on the final performance.

Differential Privacy Preservation

Differential privacy [23,43-46], as a strong criterion for privacy-preserving, is defined when the probability of a given output does not primarily depend on the involvement of a data point in the inputs [19]. This is useful because conventional deep learning has raised substantial privacy concerns, which may prevent a company from collecting data for model training. A model-inversion attack may extract parts of the training data through a deep-learning network, as demonstrated by Fredrikson et al [47]. One might attempt to reduce the risk of privacy leakage by adding noise to the parameters that result from the training process. However, it is hard to achieve a balance between performance and privacy preservation since stronger noise offers protection for privacy as well as worse performance. Therefore, we have been seeking methods that can help to preserve local privacy during the training process.

To address this issue, the SCBFL method realizes the function of differential privacy preservation by protecting the two sources of potential privacy leakages from federated learning: the actual values of uploaded gradients from the local participants and the mechanism by which these gradients are chosen [19]. By setting a threshold to select the parameters of gradients channel-wise, the actual values uploaded to the server are stored in a sparse tensor that is processed from SGD, a stochastic training process that has already been used for many privacy-preserving cases [48,49]. In addition, the participant could independently choose the update rate for their models, thus making it hard to track the selection of the channels used for the update, especially when they are trained individually using different datasets.

Limitations

We proposed an update algorithm that plays a vital role in SCBFL. This algorithm involves calculation of the α-percentile of the norms and searching for channels with greater variation in the gradients than the percentile. Given the massive size of input features, the model structure has to be extended to reach high performance. However, as is the case for most deep-learning models with complex structures, the time complexity will increase with the expansion of model size. Although the neural network pruning method has been introduced to reduce the executing time, the performance of the model will slightly decrease because of the simplified model structure. Moreover, differential privacy could be further conducted with our models to evaluate the privacy-preserving ability quantitatively.

Comparison With Prior Work

A large and representative dataset is usually required to train a neural network model. The dataset may contain sensitive information and the models should not expose the private information. Conventional methods that rely on a centrally trained model have a higher risk of privacy leakage. In conventional deep learning, the owners of the data cannot control the learning objective and have no idea of what can be inferred from their data. The federated averaging method represented progress in this regard by using iterative model averaging. Nevertheless, this approach still involves the exposure of all model parameters. The proposed SCBFL method, which improves server performance by uploading only a proportion of gradients, could address both direct and indirect privacy leakage concerns. In addition, an inverse-model attack, which extracts information from the uploaded parameters, could be hindered by the stochastic nature of our upload algorithm taking advantage of SGD. We found that even with only 10% of the channels uploaded to the server, the SCBFL model achieved an AUC-ROC of 0.9776 and an AUC-PR of 0.9695, outperforming the models that share all parameters with the server. As shown in our trials, after 100 global loops, the AUC-ROC and AUC-PR of SCBFL was 0.0033 and 0.0032 higher than that of federated averaging, respectively. Therefore, we could conclude that our method achieves comparative performance to the federated averaging method but with a higher saturating speed.

Conclusions

We proposed a privacy-preserving approach for distributed systems whose models are trained based on any type of neural network. Our methodology involved development of a channel-based update algorithm for the server, enabling the system to achieve state-of-the-art performance without forcing the participants to reveal their inputs or model weights to the server. Addressing both direct and indirect privacy leakage concerns, our model uploads a fraction of channels in the gradients from local models to the server and could achieve better performance with only 10% of the channels uploaded, thereby reducing the redundancy of gradients while preserving privacy. Inverse-model attack, which analyzes information from the uploaded parameters, could be obstructed by the stochastic nature of our upload algorithm taking advantage of SGD. Moreover, we introduced a neural pruning process to the model, which could accelerate the training process and saturating speed with little sacrifice to the final performance. Experimental validation showed that neural network pruning could efficiently speed up the training process as well as the saturation of performance. Moreover, better performance was achieved when tuning the pruning proportion to cut off the redundant neurons in several training loops.

Acknowledgments

ZC received financial support from the “Fundamental Research Funds for the Central Universities” (2019RC032). The authors would like to thank the Department of Biomedical Informatics of Harvard Medical School for provision of the dataset.

Authors' Contributions

RS wrote and revised the manuscript, conceived of the method, implemented the models, and conducted the methodological analysis. DL conceived of the methodology, provided scientific oversight, and revised the manuscript. HH performed the experiments, and wrote and revised the manuscript. HL performed the experiments and revised the manuscript.

Conflicts of Interest

None declared.

References

1. Watson HJ. Tutorial: Big Data Analytics: Concepts, Technologies, and Applications. Commun Assoc Inf Syst 2014;34. [doi: 10.17705/1cais.03465]
2. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2014;2:3 [FREE Full text] [doi: 10.1186/2047-2501-2-3] [Medline: 25825667]
3. Wang Y, Kung L, Byrd TA. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. Technol Forecast Soc Change 2018 Jan;126:3-13. [doi: 10.1016/j.techfore.2015.12.019]
4. Peterson ED. Machine Learning, Predictive Analytics, and Clinical Practice: Can the Past Inform the Present? JAMA 2019 Nov 22. [doi: 10.1001/jama.2019.17831] [Medline: 31755902]
5. Adam N, White T, Shafiq B, Vaidya J, He X. Privacy preserving integration of health care data. AMIA Annu Symp Proc 2007 Oct 11:1-5 [FREE Full text] [Medline: 18693786]
6. Bertino E, Ooi B, Yang Y, Deng R. Privacy and ownership preserving of outsourced medical data. : IEEE; 2005 Apr 18 Presented at: 21st International Conference on Data Engineering (ICDE'05); 2005; Tokoyo, Japan p. 521-532. [doi: 10.1109/ICDE.2005.111]
7. Abouelmehdi K, Beni-Hessane A, Khaloufi H. Big healthcare data: preserving security and privacy. J Big Data 2018 Jan 9;5(1):1. [doi: 10.1186/s40537-017-0110-7]
8. Craig T, Ludloff M. Privacy and big data: the players, regulators, and stakeholders. Sebastopol, California: O'Reilly Media, Inc; Sep 2011.
9. Jensen M. Challenges of Privacy Protection in Big Data Analytics. : IEEE; 2013 Sep 16 Presented at: 2013 IEEE International Congress on Big Data; June 27-July 2, 2013; Santa Clara, CA p. 235-238. [doi: 10.1109/bigdata.congress.2013.39]
10. Li N, Li T, Venkatasubramanian S. t-Closeness: Privacy Beyond k-Anonymity and l-Diversity. : IEEE; 2007 Jun 4 Presented at: 2007 IEEE 23rd International Conference on Data Engineering; April 15-20, 2007; Istanbul, Turkey p. 106-115. [doi: 10.1109/icde.2007.367856]
11. Machanavajjhala A, Gehrke J, Kifer D, Venkitasubramaniam M. L-diversity: privacy beyond k-anonymity. : IEEE; 2006 Apr 24 Presented at: 22nd International Conference on Data Engineering (ICDE'06); April 3-7, 2006; Atlanta, GA p. 24. [doi: 10.1109/ICDE.2006.1]

12. Jain P, Gyanchandani M, Khare N. Big data privacy: a technological perspective and review. J Big Data 2016 Nov 26;3(1):25. [doi: 10.1186/s40537-016-0059-y]
13. Sedayao J, Bhardwaj R, Gorade N. Making Big Data, Privacy, and Anonymization Work Together in the Enterprise: Experiences and Issues. : IEEE; 2014 Sep 25 Presented at: 2014 IEEE International Congress on Big Data; June 27-July 2, 2014; Anchorage, AK p. 601-607. [doi: 10.1109/bigdata.congress.2014.92]
14. Hundepool A, Willenborg L. Chapter 5: μ- and τ-ARGUS: Software for Statistical Disclosure Control. In: Record Linkage Techniques 1997: Proceedings of an International Workshop and Exposition. Washington, DC: National Academies Press; 1997.
15. Samarati P. Protecting respondents identities in microdata release. IEEE Trans Knowl Data Eng 2001 Dec;13(6):1010-1027. [doi: 10.1109/69.971193]
16. Sweeney L. In: Lin TY, Qian S, editors. Datafly: a system for providing anonymity in medical data. Boston, MA: Springer; 1998:356-381.
17. Mohassel P, Zhang Y. SecureML: A System for Scalable Privacy-Preserving Machine Learning. : IEEE; 2017 Jun 26 Presented at: 2017 IEEE Symposium on Security and Privacy (SP); May 22-26, 2017; San Jose, CA p. 17-38. [doi: 10.1109/sp.2017.12]
18. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. N Engl J Med 2019 Apr 04;380(14):1347-1358. [doi: 10.1056/NEJMra1814259] [Medline: 30943338]
19. Shokri R, Shmatikov V. Privacy-Preserving Deep Learning. : Association for Computing Machinery; 2015 Presented at: Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security; October 2015; New York, NY p. 1310-1321. [doi: 10.1145/2810103.2813687]
20. Konečný J, McMahan B, Ramage D. Federated Optimization:Distributed Optimization Beyond the Datacenter. arXiv. 2015 Nov 11. URL: http://arxiv.org/abs/1511.03575 [accessed 2019-11-15] [WebCite Cache ID https://arxiv.org/abs/1511.03575]
21. Konečný J, McMahan H, Ramage D, Richtárik P. Federated Optimization: Distributed Machine Learning for On-Device Intelligence. arXiv. 2016 Oct 8. URL: http://arxiv.org/abs/1610.02527 [accessed 2019-11-15]
22. Dean J, Corrado G, Monga R, Chen K, Devin M, Mao M, et al. Large Scale Distributed Deep Networks. Red Hook, NY: Curran Associates Inc; 2012 Dec Presented at: 25th International Conference on Neural Information Processing Systems; December 2012; Lake Tahoe, Nevada p. 1223-1231 URL: https://dl.acm.org/doi/10.5555/2999134.2999271
23. Sarwate AD, Chaudhuri K. Signal Processing and Machine Learning with Differential Privacy: Algorithms and challenges for continuous data. IEEE Signal Process Mag 2013 Sep 01;30(5):86-94 [FREE Full text] [doi: 10.1109/MSP.2013.2259911] [Medline: 24737929]
24. Ma C, Konečný J, Jaggi M, Smith V, Jordan MI, Richtárik P, et al. Distributed optimization with arbitrary local solvers. Optim Method Soft 2017 Feb 31;32(4):813-848. [doi: 10.1080/10556788.2016.1278445]
25. Shamir O, Srebro N, Zhang T. Communication-Efficient Distributed Optimization using an Approximate Newton-type Method. 2014 Jan Presented at: 31st International Conference on Machine Learning (ICML 2014); June 21-26, 2014; Beijing, China p. 1000-1008 URL: http://proceedings.mlr.press/v32/shamir14.pdf
26. Zhang Y, Xiao L. DiSCO: Distributed optimization for self-concordant empirical loss. 2015 Jun 1 Presented at: 32nd International Conference on Machine Learning (ICML 2015); July 6-11, 2015; Lille, France p. 362-370 URL: http://proceedings.mlr.press/v37/zhangb15.pdf
27. Chilimbi T, Suzue Y, Apacible J, Kalyanaraman K. Project Adam: Building an Efficient and Scalable Deep Learning Training System. 2014 Presented at: 11th USENIX Symposium on Operating Systems Design and Implementation (OSDI 14); October 6-8, 2014; Broomfield, CO p. 571-582 URL: https://www.usenix.org/conference/osdi14/technical-sessions/presentation/chilimbi
28. Konečný J, McMahan H, Yu F, Richtárik P, Suresh A, Bacon D. Federated Learning: Strategies for Improving Communication Efficiency. arXiv.: NIPS Workshop on Private Multi-Party Machine Learning (2016); 2017 Oct 30. URL: https://arxiv.org/abs/1610.05492 [accessed 2019-11-15]
29. McMahan H, Moore E, Ramage D, Hampson S. Communication-Efficient Learning of Deep Networks from Decentralized Data. 2017 Presented at: 20th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017; May 9-11, 2017; Fort Lauderdale, FL p. 1273-1282 URL: http://proceedings.mlr.press/v54/mcmahan17a/mcmahan17a.pdf
30. McMahan H, Moore E, Ramage D, Arcas B. Federated Learning of Deep Networks using Model Averaging. arXiv. 2016 Feb 17. URL: https://uk.arxiv.org/pdf/1602.05629v1.pdf [accessed 2019-11-15]
31. Abadi M, Chu A, Goodfellow I, McMahan H, Mironov I, Talwar K, et al. Deep Learning with Differential Privacy. 2016 Oct 24 Presented at: 2016 ACM SIGSAC Conference on Computer and Communications Security; October 24-28, 2016; Hofburg Palace, Vienna, Austria p. 308-318 URL: https://dl.acm.org/doi/10.1145/2976749.2978318 [doi: 10.1145/2976749.2978318]
32. Geyer R, Klein T, Nabi M. Differentially Private Federated Learning: A Client Level Perspective. arXiv. 2018 May 1. URL: https://arxiv.org/abs/1712.07557 [accessed 2019-11-15]
33. Bonawitz K, Ivanov V, Kreuter B, Marcedone A, McMahan H, Patel S, et al. Practical Secure Aggregation for Privacy-Preserving Machine Learning. New York, NY: Association for Computing Machinery; 2017 Oct 30 Presented at:

2017 ACM SIGSAC Conference on Computer and Communications Security; October 30-November 3, 2017; Dallas, TX p. 1175-1191 URL: https://dl.acm.org/doi/abs/10.1145/3133956.3133982 [doi: 10.1145/3133956.3133982]

34. Bonawitz K, Eichner H, Grieskamp W, Huba D, Ingerman A, Ivanov V, et al. Towards Federated Learning at Scale: System Design. 2019 Mar 22 Presented at: 2nd SysML Conference 2019; March 31-April 2, 2019; Palo Alto, CA URL: https://arxiv.org/abs/1902.01046
35. Yang T, Andrew G, Eichner H, Sun H, Li W, Kong N, et al. Applied Federated Learning: Improving Google Keyboard Query Suggestions. arXiv. 2018 Dec 7. URL: https://arxiv.org/abs/1812.02903 [accessed 2019-11-15]
36. Hard A, Rao K, Mathews R, Ramaswamy S, Beaufays F, Augenstein S, et al. Federated Learning for Mobile Keyboard Prediction. arXiv. 2019 Feb 28. URL: https://arxiv.org/abs/1811.03604 [accessed 2019-11-15]
37. Bagdasaryan E, Veit A, Hua Y, Estrin D, Shmatikov V. How To Backdoor Federated Learning. : PMLR; 2019 Aug 6 Presented at: 23rd International Conference on Artificial Intelligence and Statistics (AISTATS 2020); August 26-28, 2020; Online p. 2938-2948 URL: http://proceedings.mlr.press/v108/bagdasaryan20a.html
38. Yang T, Chen Y, Sze V. Designing Energy-Efficient Convolutional Neural Networks Using Energy-Aware Pruning. : IEEE; 2017 Nov 9 Presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); July 21-26, 2017; Honolulu, HI p. 6071-6079 URL: https://ieeexplore.ieee.org/document/8100126 [doi: 10.1109/CVPR.2017.643]
39. He Y, Zhang X, Sun J. Channel Pruning for Accelerating Very Deep Neural Networks. : IEEE; 2017 Dec 25 Presented at: 2017 IEEE International Conference on Computer Vision (ICCV); October 22-29, 2017; Venice, Italy p. 1398-1406 URL: https://ieeexplore.ieee.org/document/8237417 [doi: 10.1109/iccv.2017.155]
40. Han H, Qiao J. A Self-Organizing Fuzzy Neural Network Based on a Growing-and-Pruning Algorithm. IEEE Trans Fuzzy Syst 2010 Dec;18(6):1129-1143. [doi: 10.1109/TFUZZ.2010.2070841]
41. Srinivas S, Babu R. Data-free parameter pruning for Deep Neural Networks. : BMVC Press; 2015 Jul 22 Presented at: 26th British Machine Vision Conference (BMVC 2015); September 7-10, 2015; Swansea, UK URL: https://arxiv.org/abs/1507.06149 [doi: 10.5244/c.29.31]
42. Hu H, Peng R. Network Trimming: A Data-Driven Neuron Pruning Approach towards Efficient Deep Architectures. arXiv. 2016 Jul 12. URL: https://arxiv.org/abs/1607.03250 [accessed 2019-11-15]
43. Dwork C, Roth A. The Algorithmic Foundations of Differential Privacy. FNT in Theoretical Computer Science 2014;9(3-4):211-407. [doi: 10.1561/0400000042]
44. Dwork C, McSherry F, Nissim K, Smith A. Calibrating Noise to Sensitivity in Private Data Analysis. In: Halevi S, Rabin T, editors. Theory of Cryptography. Berlin, Heidelberg: Springer; 2006:265-284.
45. Dwork C, Naor M, Pitassi T, Rothblum G. Differential privacy under continual observation. New York, NY: Association for Computing Machinery; 2010 Presented at: 42nd ACM symposium on Theory of computing; June 5-8, 2010; Cambridge, MA p. 715-724. [doi: 10.1145/1806689.1806787]
46. Dwork C, Rothblum G, Vadhan S. Boosting and Differential Privacy. : IEEE; 2010 Presented at: 2010 IEEE 51st Annual Symposium on Foundations of Computer Science; October 23-26, 2010; Las Vegas, NV p. 51-60 URL: https://ieeexplore.ieee.org/document/5670947 [doi: 10.1109/focs.2010.12]
47. Fredrikson M, Jha S, Ristenpart T. Model Inversion Attacks that Exploit Confidence Information and Basic Countermeasures. New York, NY: Association for Computing Machinery; 2015 Presented at: 22nd ACM SIGSAC Conference on Computer and Communications Security; October 12-16, 2015; Denver, CO p. 1322-1333 URL: https://dl.acm.org/doi/10.1145/2810103.2813677 [doi: 10.1145/2810103.2813677]
48. Song S, Chaudhuri K, Sarwate A. Stochastic gradient descent with differentially private updates. : IEEE; 2013 Presented at: 2013 IEEE Global Conference on Signal and Information Processing; December 3-5, 2013; Austin, TX p. 245-248 URL: https://ieeexplore.ieee.org/document/6736861 [doi: 10.1109/globalsip.2013.6736861]
49. Bassily R, Smith A, Thakurta A. Private Empirical Risk Minimization: Efficient Algorithms and Tight Error Bounds. : IEEE; 2014 Presented at: 2014 IEEE 55th Annual Symposium on Foundations of Computer Science; October 18-21, 2014; Philadelphia, PA p. 464-473 URL: https://ieeexplore.ieee.org/document/6979031 [doi: 10.1109/focs.2014.56]

Abbreviations

APoZ: average percentage of zeros
AUC-PR: area under the precision-recall curve
AUC-ROC: area under the receiver operating characteristic curve
FN: false negative
FP: false positive
FPR: false positive rate
ReLU: rectified linear unit
SCBFL: stochastic channel-based federated learning
SCBFLwP: stochastic channel-based federated learning with pruning
SGD: stochastic gradient descent
TN: true negative

TP: true positive
TPR: true positive rate

Edited by C Lovis; submitted 30.11.19; peer-reviewed by B Sheng, K Pradeep, M Pradhan; comments to author 08.04.20; revised version received 15.07.20; accepted 22.07.20; published 22.12.20.

Please cite as:
Shao R, He H, Chen Z, Liu H, Liu D
Stochastic Channel-Based Federated Learning With Neural Network Pruning for Medical Data Privacy Preservation: Model Development and Experimental Validation
JMIR Form Res 2020;4(12):e17265
URL: http://formative.jmir.org/2020/12/e17265/
doi:10.2196/17265
PMID:33350391

Original Paper

Reasons People Who Use Opioids Do Not Accept or Carry No-Cost Naloxone: Qualitative Interview Study

Alex S Bennett[1,2], PhD; Robert Freeman[3], MA; Don C Des Jarlais[1,2,4], PhD; Ian David Aronson[1,2], PhD

[1]Department of Social and Behavioral Sciences, School of Global Public Health, New York University, New York, NY, United States
[2]Center for Drug Use and HIV Research, School of Global Public Health, New York University, New York, NY, United States
[3]School of Social Work, New York University, New York, NY, United States
[4]Department of Epidemiology, School of Global Public Health, New York University, New York City, NY, United States

Corresponding Author:
Alex S Bennett, PhD
Department of Social and Behavioral Sciences
School of Global Public Health
New York University
665 Broadway, 11th Floor
New York, NY, 10012
United States
Phone: 1 917 470 7000
Email: asb19@nyu.edu

Abstract

Background: Many people use opioids and are at risk of overdose. Naloxone is an opioid antagonist used to counter the effects of opioid overdose. There is an increased availability of naloxone in New York City; however, many who use opioids decline no-cost naloxone even when offered. Others may have the medication but opt not to carry it and report that they would be reluctant to administer it if they were to witness an overdose.

Objective: We aim to better understand why people who use opioids may be reluctant to accept, carry, and administer naloxone, and to inform the development of messaging content that addresses barriers to its acceptance and use.

Methods: We conducted formative qualitative interviews with 20 people who use opioids who are 18 years and older in New York City. Participants were recruited via key informants and chain referral.

Results: Participants cited 4 main barriers that may impede rates of naloxone acceptance, possession, and use: (1) stigma related to substance use, (2) indifference toward overdose, (3) fear of negative consequences of carrying naloxone, and (4) fear of misrecognizing the need for naloxone. Participants also offered suggestions about messaging content to tackle the identified barriers, including messages designed to normalize naloxone possession and use, encourage shared responsibility for community health, and elicit empathy for people who use drugs. Taken together, participants' narratives hold implications for the following potential messaging content: (1) naloxone is short-acting, and withdrawal sickness does not have to be long-lasting; (2) it is critical to accurately identify an opioid-involved overdose; (3) anyone can overdose; (4) naloxone cannot do harm; and (5) the prompt administration of the medication can help ensure that someone can enjoy another day. Finally, participants suggested that messaging should also debunk myths and stereotypes about people who use drugs more generally; people who use opioids who reverse overdoses should be framed as lay public health advocates and not just "others" to be managed with stigmatizing practices and language.

Conclusions: It must be made a public health priority to get naloxone to people who use opioids who are best positioned to reverse an overdose, and to increase the likelihood that they will carry naloxone and use it when needed. Developing, tailoring, and deploying messages to address stigma, indifference toward overdose, fear and trepidation about reversing an overdose, and fear of police involvement may help alleviate fears among some people who are reluctant to obtain naloxone and use the medication on someone in an overdose situation.

(JMIR Form Res 2020;4(12):e22411) doi:10.2196/22411

KEYWORDS

overdose; opioids; naloxone; people who use opioids; messaging; harm reduction; public health intervention

Introduction

The United States is in a public health crisis involving opioid-related morbidity and mortality; overdose rates are at epidemic proportions across the country [1,2]. In response, take-home naloxone (THN) has emerged as a critical medical technology to reverse opioid-related overdose that can be used safely in community settings [3-5]. As of the late 1990s, community-based organizations—and then state and local health departments—developed overdose education and naloxone distribution (OEND) programs in response to skyrocketing rates of unintentional overdose mortality. These programs were established to equip people who use drugs, their friends, and family members with naloxone and overdose recognition, reversal, and response skills [5-8]. The size and distribution modalities of programs vary by state and city; however, these programs all typically provide no-cost naloxone to people who use opioids, along with training in how to identify an overdose, conduct rescue breathing and cardiopulmonary resuscitation (CPR), and administer naloxone [9,10]. Overdose training and naloxone are also provided to staff at many drug treatment programs and syringe service programs (SSPs), and to first responders such as law enforcement, Emergency Medical Services (EMS), and fire department personnel [11,12]. These efforts have resulted in greater access to naloxone for many individuals who might not otherwise receive the medication [13-15], and evaluations of THN programs have consistently found that these programs effectively reduce overdose mortality and lead to few adverse events [16-19].

Surveys conducted from 2013 to 2019 show that the number of SSPs with OEND programs grew from 55% to 94% in response to a dramatic rise in opioid-related overdose fatalities [20]. In pace with this expansion of naloxone distribution at SSPs, a growing body of research on naloxone access [21] and experiences emerged [22-27]. This period saw a great increase in access to OEND at SSPs and more opportunities for both people who use drugs and those who do not to obtain naloxone. However, barriers to widespread access remain, and some people who use opioids still do not carry the medication, even if they were trained and given naloxone at no cost. In a recent national survey of attitudes toward naloxone among the general population, half of the respondents endorsed that “naloxone is only necessary for people who abuse opioids,” and 51% of respondents endorsed that “having naloxone available enables more drug use among people who abuse opioids” [28]. These contested beliefs about the place and role of naloxone—who should have it and the behavioral impacts of possession and use—permeate across social networks of both people who use opioids and people who do not use drugs (or do not identify as people who use drugs) [22-27]. The belief that naloxone promotes risky drug use and that only “drug abusers” are susceptible to overdose (rather than users of prescription opioids or occasional users) diminishes enthusiasm for scaling up OEND and, thus, prioritizes other traditional supply-and-demand reduction approaches to tackling the overdose crisis.

Moreover, even among people who use opioids who are trained in OEND, there is an observed lack of acceptance of, and willingness to carry, naloxone. Among 353 Baltimore adults who reported lifetime heroin use, 90% (318/353) reported naloxone awareness, and over two-thirds (224/353, 69%) reported ever receiving take-home naloxone [29]. Of the 224 individuals who had ever received naloxone, one-third reported that they never (83/224, 37%) or rarely/sometimes (84/224, 38%) carried the medication, and only 25% (57/224) reported that they always carried the medication [29]. Commonly cited reasons for not carrying or using naloxone included fear that a person may become violent or aggressive after being revived, or that police will threaten bystanders at an overdose event, or that they had insufficient overdose training [30]. In addition to gaps in naloxone access, some people who use opioids decline the medication even when it is offered. In a small pilot study conducted in New York City, 6 of 10 participants who identified as actively using opioids accepted a THN kit when offered it free-of-charge while visiting an SSP [31]. A substantial minority (4/10), however, declined it [31]. Similarly, among 472 veterans regularly using opioids in New York City who were offered free naloxone in street-based community settings, about one-fourth (110/472) declined the free naloxone kit when offered [32].

As the above research shows, a substantial number of people who use opioids and are at risk of overdose nevertheless decline no-cost THN, opt not to carry it, and report that they would be reluctant to administer it if they were to witness an overdose. However, people who use opioids and other drugs can act as critical and effective first responders to overdose [29,33,34]; therefore, there must be minimal barriers impeding their access and use of the medication. Technology-based messaging may be one avenue to increase naloxone uptake and use. Video and text messages have been used effectively by our study team members to increase HIV/HCV testing among high-risk populations and train people who use opioids to administer naloxone [31].

To inform the development of message content addressing barriers to naloxone access, possession, and use, we recruited 20 people who use opioids in New York City from June 2019 to August 2019 to participate in in-depth qualitative interviews. The participants reported barriers they and others in their social networks have encountered in the acceptance, carrying, and use of naloxone, and they offered suggestions for overcoming these identified barriers.

Methods

People who use opioids in New York City were recruited via chain referral and key informants affiliated with several SSPs located in the Bronx and Manhattan. Interviews were conducted by 2 experienced qualitative researchers in semiprivate settings in public parks and public spaces from August 2019 to October 2019. Interviewers traveled to the communities where the participants lived to conduct the interviews, representing 4 of New York City’s 5 boroughs. The semistructured interviews explored why people decline naloxone kits when offered, why people might be reluctant to use naloxone to reverse an overdose, and what types of messaging content could increase uptake, possession, and the likelihood of using naloxone to reverse an overdose. Participants were asked about their own experiences obtaining, carrying, and using naloxone, and they were asked about the experiences of others. Verbal consent was obtained and interviews lasted approximately 30 minutes, for which participants received $20 in cash. All interviews were digitally audio-recorded and transcribed for analysis. The institutional review board of the authors' home institution approved all consent documents, procedures, and the interview guide.

Interviews were conducted until thematic saturation was reached (N=20). A combination of a priori and emergent code categories were used by 2 of the authors to analyze the interviews. A deductive approach to exploring the a priori topic of naloxone possession was combined with an inductive regard for the emergent themes suggested by the participants. Thereby, the project used an iterative process of reading, coding, and discussing transcripts to achieve a flexible approach to illuminating some of the forces and processes that underlie decisions about obtaining, possessing, and using naloxone [35]. The interviewers jointly coded each transcript using the Dedoose software platform (version 4.12; SocioCultural Research Consultants) and frequently met to discuss emerging themes.

Results

Participant Characteristics

Interviews were conducted with 20 people who use opioids. Participants were predominately male (12/20, 60%), the mean age was 37 (range 23-55) years, and the majority (16/20, 80%) injected their opioids. Of the 20 participants, 75% (15/20) identified as White, 20% (4/20) identified as Black, 10% (2/20) identified as White Hispanic, and 5% (1/20) identified as Asian. All participants had been trained in naloxone use and overdose reversal; 6 of the 20 participants were not carrying naloxone at the time of the interview.

Stigma Related to Substance Use

The majority of participants cited drug-related—and particularly, opioid-related—stigma as one of the primary reasons people decline to carry or administer naloxone, even if they had previously accepted a naloxone kit and related training. Participants suggested that accepting or carrying naloxone would likely cause concerns for individuals who are particularly sensitive to being perceived as an illicit substance user or an addict. Several participants specifically cited the blue bag in which no-cost naloxone is distributed in New York as being nearly synonymous with illicit substance use and accompanying stigmas.

> *I know that some people are afraid . . . of (other) people discovering their naloxone, and that will out them as an opioid user, and I do think that those little blue bags are very recognizable. So maybe just getting rid of the blue bag would help. . . .You know, people are much less likely to recognize what it is if they don't see that, right?* [36-year-old White, non-Hispanic woman]

Other participants stressed the importance of normalizing the carrying or administering of naloxone in general.

> *When I do outreach (at a local SSP), it's like, some people will just be, like, "Oh, what you think we use, drugs?" . . . They're not comfortable with being identified as a user or whatever. I try my best to, kind of, change that perception because it's like, anybody should be able to carry Narcan. Like, it's usually not the person that's falling out that's Narcanning themselves. . . . I try to address it with basically just saying, you know . . . there's a lot of overdoses in the area, and we're giving this out to give you the opportunity to save a person's life.* [27-year-old White, non-Hispanic woman]

Even interviewees who were not carrying naloxone at the time of the interview (citing various reasons for not carrying it that particular day) recommended a number of steps they felt should be taken to overcome the stigma associated with carrying naloxone. These ranged from opting not to carry naloxone in the blue bag to delinking opioid use from naloxone and normalizing it in general.

> *I'm interested in seeing naloxone normalized. I just try to make them feel more comfortable, like (when I give people naloxone and tell them) "you guys don't have to carry it in a blue pouch if you're embarrassed. You can carry it in your pocket, you know."* [33-year-old White, Hispanic woman]

> *I would emphasize that, you know, that there's no reason in this day and age for anybody to not carry naloxone, and everybody should know this and be carrying naloxone. It doesn't mean anything other than you're trying to, you're willing to save people's lives, or give somebody else your naloxone so they can save it. . . .The message needs to get out that, I think, the naloxone is ubiquitous. You know, just because somebody has it doesn't mean they're a drug user. I mean, there's that double-edged thing of, like, well, we don't want people to be stigmatized, but at the same time, maybe people do want, you know (to identify as a drug user). So how do, how do you find ways to talk about this shit from all the different angles? How do you have the nonstigmatizing, nonjudgmental discussion?* [46-year-old White, non-Hispanic man]

In nearly all of the interviews, individuals suggested that in addition to humanizing substance users in general, one particularly effective strategy for combatting substance use–related stigma surrounding naloxone might be to develop messaging and practices that actively seek to delink naloxone with active, illicit opioid use.

Indifference Toward Overdose

Many participants also speculated that community members, referring to both people who do not use drugs and those who do, might decline to carry or administer naloxone because the lives of people who use illicit substances (and particularly opioids) are often devalued. Therefore, following this line of reasoning, some individuals simply lack the care and concern to become involved in a potential overdose situation or seek out OEND training in the first place.

> *Some people are just selfish and self-centered, you know? And they, and they, you know, if it ain't doesn't affect them or somebody they know, they don't give a damn.* [52-year-old Black, non-Hispanic man]

Based on this, participants suggested that future interventions intended to encourage individuals to carry or administer naloxone include content designed to elicit empathy and foster a shared sense of responsibility for preventing overdose-related fatalities. The point that participants thought an intervention’s messaging should make is perhaps best summed up in the simple message suggested by one participant: “You could save a life or two, you know?” [53-year-old South Asian man]. One of the strategies frequently recommended to increase community empathy was sharing personal stories from individuals who have reversed life-threatening opioid-related overdoses.

> *I feel good because . . . I can save a life, so that means something to me, you know? I would want somebody to do that for me, so, like, you know, there should be more people that, who want to do this ‘cause . . . it's not like you Superwoman or nothing, but it's somebody that is still gonna be around that you can see, that you can be like, “Wow, I saved that person.”* [36-year-old White, non-Hispanic woman]

Participants also noted that it was helpful to remind people that even individuals with whom someone is close to might be using opioids in a way that is not always apparent to others.

> *So, this would be, all right, say if a family member, a close family member of yours or very close friend, is dealing with this same situation, you wouldn't know, but you would like to know . . . you would like to know if they were doing it (using drugs) so I can have this with me, just in case you are doing it, I can save your life as well. For me and many others.* [23-year-old Black, non-Hispanic man]

Participants frequently suggested that one possibly effective strategy might be to emphasize that anyone could know someone who is at risk for a potentially fatal opioid-related overdose.

> *Because sometimes people don't know about something and they don't care. But once you sit down and you speak with somebody that knows what they're talking about, it opens up their mind, you know? It opens up their eyes, too. A relative could be overdosed, you know?* [53-year-old South Asian man]

> *I’ve explained that it's, you know, a really good thing to do, to give people another chance at life and stuff, you know? And that you don't have to be using drugs to carry naloxone, and if anything, I ask people to think about maybe some other people that they might know in their own life might be using drugs . . . there's a lot of overdoses in the area.* [42-year-old White, non-Hispanic man]

Fear of Negative Consequences of Carrying Naloxone

Several participants expressed concerns about potential legal problems related to administering naloxone or being associated with an opioid-related overdose. Moreover, participants noted that law enforcement may mischaracterize naloxone possession as evidence that a person on parole is using drugs or associating with drug users, and that people in homeless shelters could be forcibly removed if they are found with naloxone. To address these issues, participants recommended that interventions include content designed to clarify existing Good Samaritan legislation to help people who use opioids understand their rights and address related anxieties.

> *I know there's this thing now (Good Samaritan Laws), and you can't get arrested or something like that. 'Cause I know a lot of people get scared with that. And they don't call 911. . . . You want to save the person. You don't want . . . a death under your belt.* [31-year-old White, Hispanic woman]

For many participants, this was especially important considering the likelihood that medical professionals might not reach an individual experiencing an overdose until it results in an overdose-related fatality.

> *Once you learn how to use it, you shouldn't be afraid to use it because you're not gonna get charged if someone . . . let's say someone's OD'ing and you come to their aid until the EMTs get there, you might have saved someone's life right there, you know what I mean?* [53-year-old South Asian man]

Fears of Misrecognizing the Need for Naloxone

Many participants also expressed concern that they or others might not be able to accurately recognize an opioid overdose and, as a result, could administer naloxone to an individual who is otherwise intoxicated (eg, unconscious after using alcohol or benzodiazepines) or who is simply homeless and sleeping in public. Importantly, these concerns strongly discouraged people in our sample who reported they would otherwise administer, or at least consistently carry, naloxone.

> *If you're on the train and someone's really, like, nodding out, do you want to bother them and ruin their high? . . . you don't really know . . . are they drinking? Extremely drunk? And then you don't know, like, what kind of reaction you're going to get from people, ‘cause there's a certain degree of, like, mental health (problems).* [55-year-old White, non-Hispanic man]

Indeed, several participants suggested that misrecognizing an overdose could lead to serious negative consequences. Participants expressed that in addition to concerns regarding when naloxone use is actually warranted, there is always the possibility that a person given a dose of naloxone will immediately experience opioid withdrawal and become angry or even violent.

> *They don't want that naloxone to come in there and take that opiate out because it's going to make them sick, and then they don't know where their next dollar is coming from. And then someone comes along because, like, you're nodding out a little too hard or you seem like you're discombobulated, to the point where they're, like, almost dead. But if you try to get them naloxone, they'll, they'll fight you to the death not to, not to give them that, um, you know, injection*

> *or the nasal spray, because they don't want to lose that high.* [41-year-old White, non-Hispanic man]

Potentially negative financial consequences for someone who has had naloxone administered was also a frequently noted concern.

> *So, being as an addict, who wanna be sick? You know what I mean? When you gotta have, find money to buy more. You know what I'm saying? It's not like people, it's not like you're gonna give it to me. I'm sick and I'm broke. You know what I mean? I, I, I gotta find ways to get more.* [54-year-old Black, non-Hispanic man]

Participants underscored the role of experience in identifying, responding, and communicating safety through messaging about overdose. Participants discussed several popular myths related to reversing an overdose. In several instances, myths did, in fact, correspond to an opioid overdose reversal. However, when myths correspond to someone's experience managing someone else's overdose (eg, the person awakens, or regains consciousness, or becomes relatively alert), this can perpetuate myths that spread through peoples' social networks. For example, a slap, or a yell, or just a lift up may be all that is needed to prevent an overdose from becoming more serious and requiring naloxone. Participants were aware of the myths about overdose reversal and commented on the misinformation in their communities, including potentially dangerous and not scientifically proven overdose reversal methods such as hitting a person or injecting them with cocaine or milk.

> *I was revived by my roommate . . . I don't know, awakened by, um, his slapping me and telling me that I was blue. But it didn't require naloxone to revive me. And he had no drug-using experience . . . to try and bring the person back, right? Which was like a lot of the folklore around, you know: throw somebody in an ice bath, beat the bottom of their feet with sticks. All these, you know, kinds of techniques that just got around on the street.* [39-year-old White, non-Hispanic woman]

Participants also expressed doubts regarding the severity of potentially negative physical and emotional responses to administered naloxone, and the level of certainty that an individual will respond negatively at all.

> *Like, just how much . . . it is sort of the myth that's out there, that you're gonna make someone very sick. I can't pinpoint . . . it's just out there, it has always been out there. People sort of take it as a given. I think it's important to know that it only lasts 90 minutes. That's important to know, too, so that if you do use it, it's not like someone's gonna necessarily be in massive withdrawal for 12 hours.* [27-year-old White, non-Hispanic woman]

Discussion

Principal Findings

Participants cited 4 main barriers that may impede rates of naloxone acceptance, possession, and use: (1) stigma related to substance use, (2) indifference toward overdose, (3) fear of the negative consequences of carrying naloxone, and (4) fear of misrecognizing the need for naloxone. Some of these barriers have been identified by other researchers, including the fear of precipitating withdrawal sickness, stigma, and the fear of arrest [12,24,36,37]. Our study participants cited stigma toward drug users and concerns of being outed as a drug user as influencing naloxone-related practices, including decisions about whether or not to carry or prominently display naloxone when carrying it. Relatedly, participants cited the public's general indifference toward overdose and a lack of altruism as a barrier to naloxone access and use. In part, this may be associated with a perception on the part of people who use drugs of "acceptable" or "unacceptable" drug-use behaviors. For example, Bowles et al [38] found that people who frequently overdosed were often shunned by their drug-using social networks; they were considered a liability that placed others who used drugs at risk. Here, ongoing criminalization has created an atmosphere so toxic for people who use drugs that the only means of group safety and preservation is to distance from riskier drug users. However, this practice exacerbates the risk for those who are perhaps in need of the most support. It is clear that we need messaging that emphasizes naloxone as a medical technology promoting community health needs to reach community members at large. At the same time, tailored messaging is needed for people who use drugs who may themselves judge, avoid, and distance themselves from others who use drugs.

The barriers voiced by participants may not be immediately addressed simply through messaging, as many of the harms perpetrated by our dominant supply-reduction policy approach evolved over decades. For instance, stigma directed toward people who use drugs in a US context has deep historical roots [39,40]. However, what we learn from people who use opioids through this research and other efforts can help us develop tailored messaging content to overcome these barriers. Normalizing naloxone could have life-saving benefits, and destigmatizing people who use drugs could benefit their health in a broad range of communities. These messages may help move people incrementally toward regularly carrying naloxone, even if they are highly reluctant to do so initially. Many participants expressed pride when showing off their naloxone kits, strapped to belts, in bags, or on chains around necks, which we interpret as a testament to the potential for naloxone possession to be a marker of community pride and compassion for others. As Wagner et al [26] point out, reversing or witnessing an overdose can be a traumatic, cathartic, humbling, or empowering experience. Accordingly, efforts to normalize naloxone could complement participants' messaging recommendations to emphasize shared responsibility for community health and elicit empathy for people who use drugs. Messaging could also debunk myths and stereotypes about people who use drugs more generally and help frame people who use opioids and reverse overdoses as lay public health advocates, and not just "others" to be managed with stigmatizing practices and language.

Of particular concern is the participants' expressed trepidations about potentially misrecognizing an overdose, inadvertently precipitating withdrawal, and creating an undue financial burden on some people who use opioids, which are fears that have also been found in other research [30]. One approach to overcoming these barriers could be OEND booster messaging, deployed electronically and virtually, about protections when 911 is called, and how to identify an overdose. These messages could be coupled with other health promotion and risk-reduction messaging targeting people who use drugs. Additional resources could be allocated to community-based programs to develop and broadly distribute print and web-based messaging about best overdose-response practices and resources. Standard OEND trainings provide information on how to distinguish an overdose from a “nod” by encouraging bystanders to call out to the person if an overdose is suspected, rub the sternum or pinch an earlobe to see if there is a response, or check for breathing before administering naloxone. This information presented in OEND training could be delivered to people who use opioids on an ongoing basis, with messaging covering each point in the standard OEND curriculum. Both community-based and virtual refresher trainings could help accomplish this. Taken together, participants’ narratives hold implications for the following potential messaging content: (1) naloxone is short-acting, and withdrawal sickness does not have to be long-lasting; (2) it is critical to accurately identify an opioid-involved overdose; (3) anyone can overdose; (4) naloxone is safe to use; and (5) the prompt administration of the medication can help ensure that someone can enjoy another day.

As we grapple with the novel coronavirus, we need to remain especially vigilant. Rates of overdose continue to be high. The long-term impacts of the virus are as yet unknown, and social distancing and isolation are placing more people who use drugs at risk for a potentially fatal overdose [41-43]. Thus, compounded by the state of the COVID-19 pandemic, getting naloxone to people who use opioids is vital.

Limitations

The findings may be unique to the population of people who use opioids in New York City, where there is robust naloxone distribution. Our findings may or may not generalize to other cities, where the overdose rate and degree of stigmatization may be different and Good Samaritan legislation may or may not be actively followed. Finally, given the nature of qualitative research, which involves small samples and nonprobabilistic sampling methods, the findings are not intended to be generalized to the broader population, suggesting the need for additional cross-sectional, longitudinal, and comparative investigations.

Conclusions

Getting naloxone to those who are best positioned to reverse an overdose, and increasing the likelihood they will use it when needed, should continue to be a public health priority. However, the barriers identified by participants regarding naloxone access, possession, and use are considerable. Our study found a common reluctance to administer naloxone because of stigma, apathy, concerns about precipitating withdrawal, misrecognizing a good “high” as an overdose, and fears of police. These underscore the importance of distributing this proven, life-saving medication and of creating positive, acceptable messaging to ensure people use it when it is needed most.

Acknowledgments

The authors would like to thank the National Institute on Drug Abuse for the funding (#3P30DA011041-22S1; P30DA011041; P30DA029926; DA R01DA046653; DA R01 003574) to conduct this research. Points of view expressed in this paper do not necessarily represent the official position of the US Government, the National Institutes of Health, the National Institute on Drug Abuse, or New York University. The authors express their deep appreciation to all who participated in our study.

Conflicts of Interest

None declared.

References

1. Rudd RA, Aleshire N, Zibbell JE, Gladden RM. Increases in drug and opioid overdose deaths--United States, 2000-2014. MMWR Morb Mortal Wkly Rep 2016 Jan 01;64(50-51):1378-1382. [doi: 10.15585/mmwr.mm6450a3] [Medline: 26720857]
2. Hedegaard H, Miniño AM, Warner M. Drug overdose deaths in the United States, 1999-2018. NCHS Data Brief 2020 Jan(356):1-8 [FREE Full text] [Medline: 32487285]
3. McDonald R, Campbell ND, Strang J. Twenty years of take-home naloxone for the prevention of overdose deaths from heroin and other opioids-conception and maturation. Drug Alcohol Depend 2017 Sep 01;178:176-187. [doi: 10.1016/j.drugalcdep.2017.05.001] [Medline: 28654870]
4. Strang J, Darke S, Hall W, Farrell M, Ali R. Heroin overdose: the case for take-home naloxone. BMJ 1996 Jun 08;312(7044):1435-1436 [FREE Full text] [doi: 10.1136/bmj.312.7044.1435] [Medline: 8664611]
5. Wheeler E, Jones TS, Gilbert MK, Davidson PJ, Centers for Disease Control and Prevention (CDC). Opioid overdose prevention programs providing naloxone to laypersons - United States, 2014. MMWR Morb Mortal Wkly Rep 2015 Jun 19;64(23):631-635 [FREE Full text] [Medline: 26086633]
6. Clark AK, Wilder CM, Winstanley EL. A systematic review of community opioid overdose prevention and naloxone distribution programs. J Addict Med 2014;8(3):153-163. [doi: 10.1097/ADM.0000000000000034] [Medline: 24874759]
7. Kim D, Irwin KS, Khoshnood K. Expanded access to naloxone: options for critical response to the epidemic of opioid overdose mortality. Am J Public Health 2009 Mar;99(3):402-407. [doi: 10.2105/AJPH.2008.136937] [Medline: 19150908]

8. Compton WM, Volkow ND, Throckmorton DC, Lurie P. Expanded access to opioid overdose intervention: research, practice, and policy needs. Ann Intern Med 2013 Jan 01;158(1):65-66. [doi: 10.7326/0003-4819-158-1-201301010-00013] [Medline: 23277902]
9. Bennett AS, Bell A, Tomedi L, Hulsey EG, Kral AH. Characteristics of an overdose prevention, response, and naloxone distribution program in Pittsburgh and Allegheny County, Pennsylvania. J Urban Health 2011 Dec;88(6):1020-1030 [FREE Full text] [doi: 10.1007/s11524-011-9600-7] [Medline: 21773877]
10. Substance use: community management of opioid overdose. World Health Organization. 2014. URL: https://apps.who.int/iris/bitstream/handle/10665/137462/9789241548816_eng.pdf
11. Rando J, Broering D, Olson JE, Marco C, Evans SB. Intranasal naloxone administration by police first responders is associated with decreased opioid overdose deaths. Am J Emerg Med 2015 Sep;33(9):1201-1204. [doi: 10.1016/j.ajem.2015.05.022] [Medline: 26095132]
12. Bessen S, Metcalf SA, Saunders EC, Moore SK, Meier A, McLeman B, et al. Barriers to naloxone use and acceptance among opioid users, first responders, and emergency department providers in New Hampshire, USA. Int J Drug Policy 2019 Dec;74:144-151 [FREE Full text] [doi: 10.1016/j.drugpo.2019.09.008] [Medline: 31590090]
13. Lenton SR, Dietze PM, Degenhardt L, Darke S, Butler TG. Naloxone for administration by peers in cases of heroin overdose. Med J Aust 2009 Oct 19;191(8):469. [Medline: 19835548]
14. Maxwell S, Bigg D, Stanczykiewicz K, Carlberg-Racich S. Prescribing naloxone to actively injecting heroin users: a program to reduce heroin overdose deaths. J Addict Dis 2006;25(3):89-96. [doi: 10.1300/J069v25n03_11] [Medline: 16956873]
15. Walley AY, Xuan Z, Hackman HH, Quinn E, Doe-Simkins M, Sorensen-Alawad A, et al. Opioid overdose rates and implementation of overdose education and nasal naloxone distribution in Massachusetts: interrupted time series analysis. BMJ 2013 Jan 30;346:f174 [FREE Full text] [doi: 10.1136/bmj.f174] [Medline: 23372174]
16. Giglio RE, Li G, DiMaggio CJ. Effectiveness of bystander naloxone administration and overdose education programs: a meta-analysis. Inj Epidemiol 2015 Dec;2(1):10 [FREE Full text] [doi: 10.1186/s40621-015-0041-8] [Medline: 27747742]
17. McDonald R, Strang J. Are take-home naloxone programmes effective? Systematic review utilizing application of the Bradford Hill criteria. Addiction 2016 Jul;111(7):1177-1187 [FREE Full text] [doi: 10.1111/add.13326] [Medline: 27028542]
18. Doe-Simkins M, Walley AY, Epstein A, Moyer P. Saved by the nose: bystander-administered intranasal naloxone hydrochloride for opioid overdose. Am J Public Health 2009 May;99(5):788-791. [doi: 10.2105/AJPH.2008.146647] [Medline: 19363214]
19. Wagner KD, Valente TW, Casanova M, Partovi SM, Mendenhall BM, Hundley JH, et al. Evaluation of an overdose prevention and response training programme for injection drug users in the Skid Row area of Los Angeles, CA. Int J Drug Policy 2010 May;21(3):186-193 [FREE Full text] [doi: 10.1016/j.drugpo.2009.01.003] [Medline: 19268564]
20. Lambdin BH, Bluthenthal RN, Wenger LD, Wheeler E, Garner B, Lakosky P, et al. Overdose education and naloxone distribution within syringe service programs - United States, 2019. MMWR Morb Mortal Wkly Rep 2020 Aug 21;69(33):1117-1121. [doi: 10.15585/mmwr.mm6933a2] [Medline: 32817603]
21. Bennett A, Elliott L, Wolfson-Stofko B. Commentary on Madah-Amiri et al. (2019): Beyond saturation. Addiction 2019 Jan;114(1):101-102. [doi: 10.1111/add.14499] [Medline: 30520182]
22. Heavey SC, Chang YP, Vest BM, Collins RL, Wieczorek W, Homish GG. 'I have it just in case' - Naloxone access and changes in opioid use behaviours. Int J Drug Policy 2018 Jan;51:27-35. [doi: 10.1016/j.drugpo.2017.09.015] [Medline: 29156400]
23. Kerensky T, Walley AY. Opioid overdose prevention and naloxone rescue kits: what we know and what we don't know. Addict Sci Clin Pract 2017 Jan 07;12(1):4 [FREE Full text] [doi: 10.1186/s13722-016-0068-3] [Medline: 28061909]
24. Kirane H, Ketteringham M, Bereket S, Dima R, Basta A, Mendoza S, et al. Awareness and attitudes toward intranasal naloxone rescue for opioid overdose prevention. J Subst Abuse Treat 2016 Oct;69:44-49. [doi: 10.1016/j.jsat.2016.07.005] [Medline: 27568509]
25. Lankenau SE, Wagner KD, Silva K, Kecojevic A, Iverson E, McNeely M, et al. Injection drug users trained by overdose prevention programs: responses to witnessed overdoses. J Community Health 2013 Feb;38(1):133-141 [FREE Full text] [doi: 10.1007/s10900-012-9591-7] [Medline: 22847602]
26. Wagner KD, Davidson PJ, Iverson E, Washburn R, Burke E, Kral AH, et al. "I felt like a superhero": the experience of responding to drug overdose among individuals trained in overdose prevention. Int J Drug Policy 2014 Jan;25(1):157-165 [FREE Full text] [doi: 10.1016/j.drugpo.2013.07.003] [Medline: 23932166]
27. Jones JD, Campbell A, Metz VE, Comer SD. No evidence of compensatory drug use risk behavior among heroin users after receiving take-home naloxone. Addict Behav 2017 Aug;71:104-106 [FREE Full text] [doi: 10.1016/j.addbeh.2017.03.008] [Medline: 28325710]
28. Smith JO, Malinowski SS, Ballou JM. Public perceptions of naloxone use in the outpatient setting. Ment Health Clin 2019 Jul;9(4):275-279 [FREE Full text] [doi: 10.9740/mhc.2019.07.275] [Medline: 31293847]
29. Tobin K, Clyde C, Davey-Rothwell M, Latkin C. Awareness and access to naloxone necessary but not sufficient: examining gaps in the naloxone cascade. Int J Drug Policy 2018 Sep;59:94-97 [FREE Full text] [doi: 10.1016/j.drugpo.2018.07.003] [Medline: 30075401]

30. Dayton L, Gicquelais RE, Tobin K, Davey-Rothwell M, Falade-Nwulia O, Kong X, et al. More than just availability: who has access and who administers take-home naloxone in Baltimore, MD. PLoS One 2019;14(11):e0224686 [FREE Full text] [doi: 10.1371/journal.pone.0224686] [Medline: 31697736]
31. Aronson ID, Bennett A, Marsch LA, Bania TC. Mobile technology to increase HIV/HCV testing and overdose prevention/response among people who inject drugs. Front Public Health 2017;5:217. [doi: 10.3389/fpubh.2017.00217] [Medline: 28879174]
32. Bennett AS, Elliott L, Wolfson-Stofko B, Almenana R. Developing a peer delivery model for naloxone distribution and increasing opioid safety knowledge among veterans. 2016 Presented at: Harm Reduction Coalition; October 3-6, 2016; San Diego, CA.
33. Bennett AS, Bell A, Doe-Simkins M, Elliott L, Pouget E, Davis C. From peers to lay bystanders: findings from a decade of naloxone distribution in Pittsburgh, PA. J Psychoactive Drugs 2018;50(3):240-246. [doi: 10.1080/02791072.2018.1430409] [Medline: 29424656]
34. Strang J, Best D, Man L, Noble A, Gossop M. Peer-initiated overdose resuscitation: fellow drug users could be mobilised to implement resuscitation. Int J Drug Policy 2000 Dec 01;11(6):437-445. [doi: 10.1016/s0955-3959(00)00070-0] [Medline: 11099924]
35. Miles MB, Huberman AM. Qualitative Data Analysis: An Expanded Sourcebook 2nd ed. Thousand Oaks, CA: Sage; 1994.
36. Bowles JM, Lankenau SE. "I Gotta Go With Modern Technology, So I'm Gonna Give 'em the Narcan": the diffusion of innovations and an opioid overdose prevention program. Qual Health Res 2019 Feb;29(3):345-356. [doi: 10.1177/1049732318800289] [Medline: 30311841]
37. Wagner KD, Harding RW, Kelley R, Labus B, Verdugo SR, Copulsky E, et al. Post-overdose interventions triggered by calling 911: centering the perspectives of people who use drugs (PWUDs). PLoS One 2019;14(10):e0223823 [FREE Full text] [doi: 10.1371/journal.pone.0223823] [Medline: 31622401]
38. Bowles J, Smith L, Verdugo S, Wagner K, Davidson P. "Generally, you get 86'ed because you're a liability": an application of Integrated Threat Theory to frequently witnessed overdoses and social distancing responses. Soc Sci Med 2020 Sep;260:113190. [doi: 10.1016/j.socscimed.2020.113190] [Medline: 32673794]
39. Acker CJ. Creating the American Junkie: Addiction Research in the Classic Era of Narcotic Control. Baltimore: Johns Hopkins University Press; 2002.
40. Courtwright DT. Dark Paradise: A History of Opiate Addiction in America. Cambridge, MA: Harvard University Press; 2001.
41. Glick SN, Prohaska SM, LaKosky PA, Juarez AM, Corcorran MA, Des Jarlais DC. The impact of COVID-19 on syringe services programs in the United States. AIDS Behav 2020 Sep;24(9):2466-2468 [FREE Full text] [doi: 10.1007/s10461-020-02886-2] [Medline: 32333209]
42. Volkow ND. Collision of the COVID-19 and addiction epidemics. Ann Intern Med 2020 Jul 07;173(1):61-62 [FREE Full text] [doi: 10.7326/M20-1212] [Medline: 32240293]
43. Wakeman SE, Green TC, Rich J. An overdose surge will compound the COVID-19 pandemic if urgent action is not taken. Nat Med 2020 Jun;26(6):819-820. [doi: 10.1038/s41591-020-0898-0] [Medline: 32555514]

Abbreviations

OEND: overdose education and naloxone distribution programs
SSP: syringe service program
THN: take-home naloxone

Edited by G Eysenbach; submitted 10.07.20; peer-reviewed by S Moore, C Freiermuth; comments to author 31.07.20; revised version received 25.09.20; accepted 03.12.20; published 23.12.20.

Please cite as:
Bennett AS, Freeman R, Des Jarlais DC, Aronson ID
Reasons People Who Use Opioids Do Not Accept or Carry No-Cost Naloxone: Qualitative Interview Study
JMIR Form Res 2020;4(12):e22411
URL: http://formative.jmir.org/2020/12/e22411/
doi:10.2196/22411
PMID:33355094

Original Paper

Grocery Delivery of Healthy Foods to Pregnant Young Women With Low Incomes: Feasibility and Acceptability Mixed Methods Study

Ione Locher[1*], MS; Marika Waselewski[1*], MPH; Kendrin Sonneville[1], ScD, RD; Ken Resnicow[1], PhD; Tammy Chang[1*], MD, MPH, MS

University of Michigan, Ann Arbor, MI, United States
*these authors contributed equally

Corresponding Author:
Tammy Chang, MD, MPH, MS
University of Michigan
2800 Plymouth Road
Ann Arbor, MI, 48109
United States
Phone: 1 734 730 4636
Email: tachang@med.umich.edu

Abstract

Background: Poor maternal diets increase the risk of excess gestational weight gain which can contribute to serious intergenerational morbidity for both the mother and infant. Pregnant young women with low incomes have disproportionately high rates of inadequate fruit and vegetable consumption as well as excess weight gains during pregnancy.

Objective: Our aim was to describe the feasibility and acceptability of Special Delivery, a longitudinal nutrition intervention that delivers healthy foods to pregnant youth (aged 14-24 years) with low incomes.

Methods: The Special Delivery pilot study, conducted in Michigan, enrolled pregnant young women with low incomes. Study participants were sent twice-monthly grocery deliveries consisting of US $35 worth of healthy foods, primarily fruits and vegetables. Between grocery deliveries, participants received daily SMS text message prompts to confirm receipt of delivery and document diet and weight. Program feasibility was assessed by the number of grocery orders placed, delivered, and confirmed by participants. Qualitative interviews and SMS text message data were used to determine acceptability by assessing participants' perspectives on grocery delivery, participants' perspectives on dietary impact of the program, and foods consumed by participants.

Results: A total of 27 participants were enrolled in the pilot study. The mean age was 20.3 years (SD 2.0), and 59.3% (16/27) were African American or Black. During the pilot, 263 deliveries were sent with 98.5% (259/263) successful deliveries and 89.4% (235/263) deliveries confirmed by participants. Participants reported that grocery delivery was convenient; that delivered foods were high quality; and that the program improved their diet, increased access to healthy foods, and promoted healthy habits during pregnancy.

Conclusions: A grocery delivery–based weight gain and nutrition intervention is both feasible and acceptable among low-income pregnant youth. Grocery deliveries were successfully completed and participants were willing and able to receive grocery deliveries, eat the healthy foods that were delivered, and communicate via SMS text message with study coordinators. The Special Delivery program warrants further evaluation for efficacy in promoting healthy weight gain for low-income youth during pregnancy.

(JMIR Form Res 2020;4(12):e21602) doi:10.2196/21602

KEYWORDS

pregnancy; adolescent; young adult; female; gestational weight gain; diet; food preferences; text messaging; feasibility studies

Introduction

Pregnancy is a crucial time for women to make healthy dietary choices. Maternal consumption of fruit and vegetables is associated with healthy pregnancy weight gain and reduced risk of poor pregnancy outcomes such as miscarriage, gestational hypertension, and gestational diabetes [1-3]. Meanwhile, dietary patterns that result in excess gestational weight gain (defined as weight gain above the National Academy of Medicine gestational weight gain guidelines [4]) can have long-term consequences on lifetime weight gain and health outcomes for pregnant woman and their children [5]. Excess weight gain in pregnancy can also lead to complications during birth and permanently impacts fetal genetic programming, which determines risk for chronic disease among infants [6-9]. Adolescent pregnant women are at significantly higher risk of excess gestational weight compared to older women [4,10,11]. Women with low incomes are at additional risk for both low fruit and vegetable consumption and excess gestational weight gain [12,13].

The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) is a federally (United States) funded program that provides financial support for the purchase of healthy foods to pregnant women with low incomes and their children, to encourage healthy eating during and after pregnancy. However, many young mothers are unable to adequately access these healthy foods due to logistical barriers such as lack of transportation, little experience in grocery shopping, and difficulty navigating WIC benefit redemption requirements [14-17]. As convenience and availability of healthy foods are among the most powerful factors driving weight gain and diet behavior [18-20], innovative approaches that minimize these logistical barriers should be examined. Grocery delivery represents a well-established, inexpensive, and convenient service that can significantly increase availability of healthy foods [21]. Grocery delivery has already been demonstrated to improve access to healthy food in underresourced areas, such as urban food deserts [22]. In addition to providing convenient access to healthy foods, grocery delivery of healthy foods has been shown to improve diet quality [21,23,24]. Home delivery of non sugar-sweetened beverages to adolescents was associated with significant reductions in adolescents’ intake of sugar-sweetened beverages, and the impact persisted 2 years after the intervention [25].

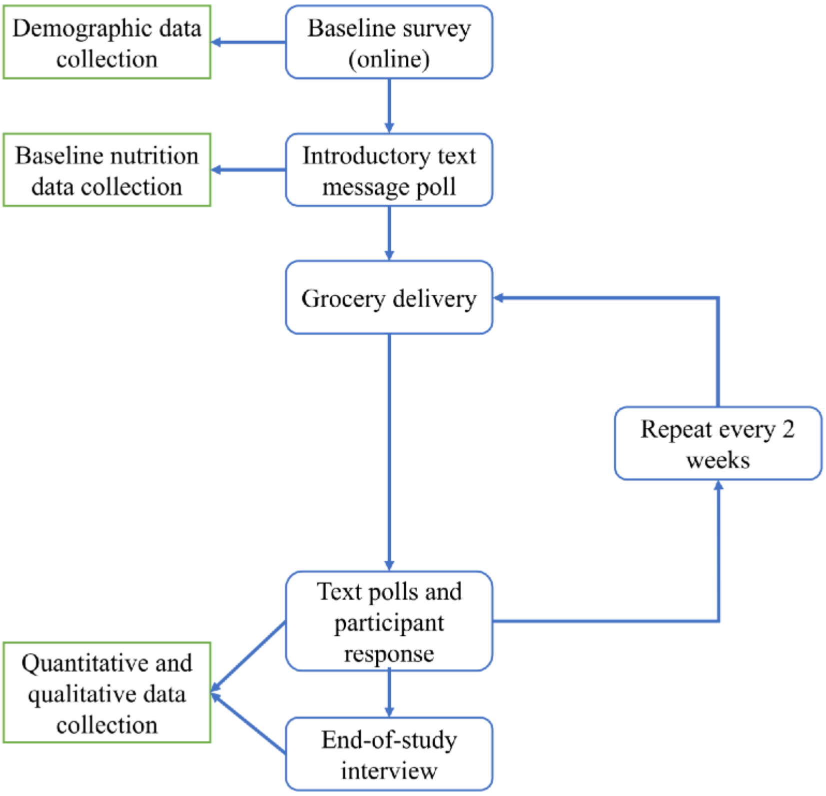
No studies to date have assessed the impact of providing home grocery delivery to pregnant young women. To address the critical need for increased access to healthy food for pregnant adolescents with low incomes, we developed an intervention that uses twice-monthly home delivery of groceries to provide healthy food options to participating pregnant young women with low incomes in southeast Michigan. In this paper, we describe our pilot study to assess feasibility and acceptability of this intervention.

Methods

Overview

Special Delivery is a grocery delivery program that was offered to WIC enrollees who were youths to improve healthy weight gain during pregnancy. We leveraged a low-cost online service to provide healthy foods directly to youths' homes throughout their pregnancy. Data were collected by phone, primarily via text messaging, allowing for participants to quickly and easily contribute their perspectives in an accessible medium. During the study, real-time participant feedback was used to improve data collection tools and delivery processes. An outline of the Special Delivery study process is depicted in Figure 1. This study was approved by the University of Michigan institutional review board (HUM00140840).

Figure 1. Outline of the special delivery process.



Recruitment

Youth were referred to Special Delivery by local WIC offices in Genesee, Washtenaw, and Wayne counties. Eligibility criteria were age 14-24 years, enrollment in WIC, first pregnancy, gestational age <24 weeks, low-risk singleton pregnancy, fluent in written and spoken English, access to a phone with text messaging capabilities, and home address within the radius of the delivery service. Low-risk pregnancy was determined based on self-reported absence of high-risk conditions such as pre-eclampsia, gestational diabetes, serious mental health conditions, or any other condition requiring specialty care. If youth were eligible, consent was obtained; parental consent was waived for minors as approved by the institutional review board both to protect the confidentiality of pregnant minors and because the study was of minimal risk to participants.

At enrollment, participants completed an intake assessment either online or over the phone with study coordinators to confirm study eligibility and collect participant demographic data including date of birth, phone number, home address, estimated date of delivery, prepregnancy weight, current weight, height, and socioeconomic variables. Food insecurity was assessed for all participants via a youth-validated 2-question scale [26,27]. All eligible individuals who completed the assessment online were called by a study coordinator to confirm participation, review study procedures, create a food delivery schedule, and collect baseline food preferences.

Intervention Period

The intervention period of the Special Delivery program lasted from participant enrollment until infant delivery; depending on gestational age at enrollment, the intervention period ranged from 6-8 months. During the intervention, participants received twice-monthly grocery deliveries of healthy foods and daily text prompts for data collection. Participants were encouraged to send several types of text messages including photos to confirm grocery deliveries, photos of foods eaten by participants, and text responses to food frequency surveys. Phone and text message-based communication was used as it is low burden to youth, who commonly own mobile devices and communicate via text [28,29]. Textizen (Vox Metropolis Inc), a secure online platform with response-based automated texting that simulates an active conversation, was used to collect all text message data.

Grocery Delivery

Twice-monthly grocery deliveries were made to Special Delivery participants until the end of their pregnancy (completion of the program) or unenrollment from the program. Each delivery contained US $35 worth of food and consisted of primarily WIC-approved foods for pregnant women such as seasonal fresh fruits and vegetables; frozen fruits and vegetables; and a variety of healthy snacks, including yogurt, cheese, whole grain cereal, whole grain waffles, popcorn, granola, coconut water, and flavored unsweetened water (examples are shown in Figure 2). Each delivery was approximately 75% fruits and vegetables with 3-4 healthy snack items. Participants were asked to share their food preferences at enrollment and again via text message 2 days prior to each grocery delivery. Deliveries were tailored to accommodate participant choice whenever possible.

Figure 2. Example photos confirming receipt of typical grocery deliveries.

Participants were asked to confirm receipt of each delivery by either texting a photo or calling the study coordinator. Specifically, confirmation could be completed by response to automated texts on the day of grocery delivery, response to automated texts on subsequent days before the next delivery, or directly with study coordinators by phone or text. Examples of photo confirmations submitted via text message are shown in Figure 2.

Groceries were delivered via the delivery service Shipt, a web- and app-based company that contracts with freelance shoppers to shop and deliver orders of groceries to home addresses. Shipt facilitates grocery delivery within a defined area (specified by zip codes) at any time the contracted grocery store is open for business. Deliveries are ordered for a specific timeframe and arrive within a 1-hour window of the selected time. The Special Delivery team managed its own Shipt account to coordinate deliveries and acted as the primary contact for questions related to substitutions or difficulties completing a delivery. This account was donated by Shipt for the purposes of this research project; however, a yearly subscription typically costs US $99. At enrollment, participants were asked to provide any details that could help Shipt shoppers in locating their address to successfully complete deliveries, including buzzer information, passcodes for gated communities, unofficial street names, or even colors of houses.

Data Collection

To assess feasibility and acceptability of grocery delivery, daily data collection polls associated with each grocery delivery were sent via text message to participants and daily surveys prompted participants to confirm receipt of grocery deliveries and share examples of foods eaten to provide real-time data. Participants were encouraged to send photos of any foods consumed, regardless of whether they were from their grocery delivery. Additionally, all participants were contacted around their estimated infant due date for an end-of-study interview. Interviews were conducted over the phone by study coordinators in a semistructured fashion. Using an inductive approach, interview questions elicited feedback from participants about their perception of grocery deliveries, food quality, consumption of food sent, dietary impact of deliveries, and use of text message communication.

Participant Incentives

Throughout the study, participants received incentives: US $10 for completion of the intake assessment, $10 for completion of end-of-study interviews, $1 for each day they responded to automated text surveys, and a $3 bonus for answering all text surveys in 1 month. Incentives were sent as Amazon gift cards via weblinks texted directly to the participants’ phones.

Data Analysis

Using a mixed methods study design, we evaluated feasibility and acceptability with 2-way text message–based data collection and qualitative interviews (Table 1). Outcomes used to assess feasibility included quantification of the number of deliveries ordered by study coordinators, successfully delivered by Shipt shoppers, and confirmed by recipients. Acceptability outcomes were assessed by reviewing end-of-study interviews for perceptions of the grocery delivery process, delivered foods, and impact of grocery delivery on diet. Interview transcripts were reviewed by 2 researchers to identify themes in participant perceptions and feedback. Program acceptability was measured by frequency of positive perceptions reported by participants in interviews.

Participant food photos were analyzed by identifying the foods depicted in the image and categorizing them as either being consistent with groceries included in deliveries or not. A high frequency of foods that were consistent with foods in the deliveries was considered to be support for program acceptability.

Table 1. Program feasibility and acceptability measures.

Feasibility	Acceptability
Number of grocery orders placed	Participant perception of delivery process and delivered foods
Number of successful deliveries	Participant perception of impact on health
Number of confirmed deliveries	Foods eaten by participants

Results

Study Population

Enrollment in the Special Delivery pilot study ran from January 2019 to November 2019, and 27 participants were enrolled in the program, of which 24 participants completed their intervention periods as defined by the end of their pregnancy, and 21 completed the end-of-study interview (Figure 3). During the study, 1 participant did not confirm receipt of any grocery deliveries and was presumed lost to follow-up, resulting in stopped deliveries, 1 participant withdrew from enrollment prior to the end of her pregnancy because she did not have reliable access to an address for home delivery, and 1 participant withdrew from enrollment prior to the end of her pregnancy due to preference not to receive grocery deliveries. Participants who participated until their end of pregnancy received groceries for an average of 4.8 months (range 1.8-7.2).

Demographic characteristics of enrolled participants are displayed in Table 2. At enrollment, participants were between 17 to 23 years of age (mean 20.3, SD 2.0), and the mean gestational age was 16.3 weeks (SD 6.5). A majority of participants identified as non-Hispanic Black (16/27, 59%) and had attained a high-school level education or less (20/27, 74%); 56% of participants (15/27) reported a prepregnancy weight classified as normal (BMI between 18.5 and 24.9 kg/m^2), and 30% of participants (8/27) reported BMI values classified as obese.

Figure 3. Participant enrollment and follow-up flowchart.

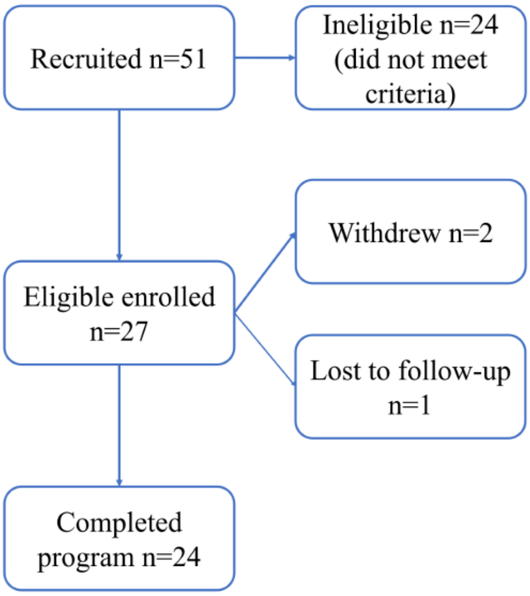


Table 2. Baseline demographic characteristics.

Characteristic	Value (n=27)
Age (years), mean (SD)	20.3 (2.0)
Gestational age at enrollment (weeks), mean (SD)	16.3 (6.5)
Race/ethnicity, n (%)	
Non-Hispanic White	4 (15)
Non-Hispanic Black	16 (59)
Non-Hispanic other	5 (19)
Hispanic	2 (7)
BMI classification (kg/m^2), n (%)	
Underweight (<18.5)	1 (4)
Normal weight (18.5-24.9)	15 (56)
Overweight (25-29.9)	3 (11)
Obese (>30.0)	8 (30)
Education level, n (%)	
Some high school	8 (30)
High school graduate	12 (44)
Some college	4 (15)
Associate’s degree	1 (4)
Bachelor’s degree	2 (7)
Food insecure (n=26), n (%)	
Yes	14 (54)
No	12 (46)

Program Feasibility

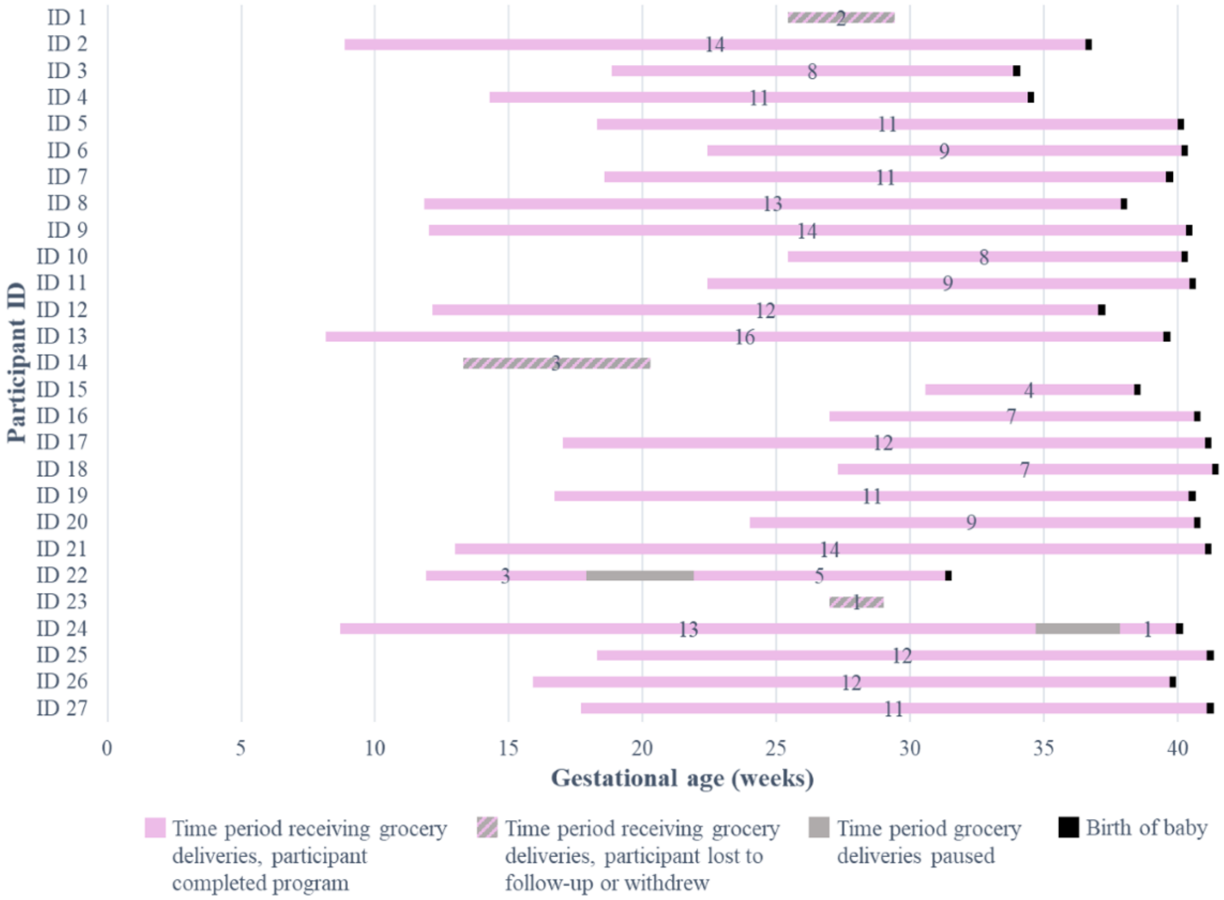
The Special Delivery program staff ordered a total of 263 deliveries during the pilot study. Of these, 259 deliveries (98.5%) of deliveries were considered successfully delivered. Only 1 delivery was reported as missing by the recipient and had to be reordered (<1%). In addition, 3 grocery orders (1.1%) were delivered later than the scheduled delivery hour, though they were delivered within 1 day.

Throughout the pilot study, Shipt shoppers communicated the outcome of every grocery delivery to study coordinators via text messaging. Shipt shoppers were self-directed and required relatively little contact from study coordinators to successfully complete grocery deliveries. Occasionally, Shipt shoppers were not able to hand off grocery deliveries directly to recipients because of difficulty locating addresses, restricted access to gated communities, or no answer at the front door. In these cases, Shipt shoppers discussed alternative delivery options with study coordinators, including asking for directions, leaving groceries at leasing offices or, as a last resort, at the front door of a house or apartment building. Study coordinators conveyed information between Shipt shoppers and the program participants receiving grocery deliveries.

Participants received deliveries twice monthly, and on average, participants received 9.7 deliveries in total (range 1-16 deliveries). The number of grocery orders delivered to each participant during the Special Delivery pilot is displayed in Figure 4. over the course of the pilot study, 1 participant paused deliveries for 4 weeks due to a shut-off in phone service (because no alternate contact method was listed) but was able to restore her phone service and continue to successfully receive and confirm grocery deliveries. In addition, 1 participant experienced an interruption in regular deliveries for three weeks due to limitations in the availability of grocery delivery services during the first few weeks of the Michigan stay-at-home order during the COVID-19 pandemic [30].

Participants confirmed receipt of 235 of the 263 deliveries (89.4%) by text, photo, or phone call. The low number of grocery deliveries reported missing by participants and the consistent communication from Shipt shoppers about the outcome of every grocery delivery both support the high fidelity of grocery delivery.

Figure 4. Each participant's intervention period with number of grocery deliveries received indicated.



Program Acceptability

Participants were asked about their perceptions of the Special Delivery process, including grocery delivery and the quality of delivered foods, during qualitative end-of-study interviews (Table 3). Nearly all participants who completed an end-of-study interview (19/21) reported that home delivery of groceries was convenient. Some participants also noted specific aspects of inconvenience (3/21). However, these were related to study procedures rather than the delivery of groceries (for example, 1 participant described the inconvenience of contacting study coordinators to change the day or time of delivery). Finally, 1 participant noted she felt that she had to change her clothes to be more presentable when answering her front door to receive her groceries.

Most participants (20/21) reported that delivered foods, including fruits and vegetables, were high quality, and 8 participants specifically described the freshness of fruits and vegetables as a positive aspect. Notably, 2 participants who reported foods as high quality also observed occasional instances when fruits and vegetables went bad within days of delivery, and 1 participant reported that the delivered foods were consistently low quality for the same reason. Despite these exceptions, perceptions of the Special Delivery process were strongly positive.

Participants were also asked about their perceptions of the impact of Special Delivery on health. Nearly all participants reported that the program helped them to have a healthy diet (20/21). Many participants reported improved access to healthy foods as a result of the Special Delivery program (15/21). In addition, most participants reported that the Special Delivery program helped them to build healthy habits (13/21). Examples of these healthy habits include substituting unhealthy food for healthy alternatives, cooking at home, tracking intake of healthy foods, maintaining adequate hydration, and shopping for more fruits and vegetables at the grocery store. Some participants also described trying new healthy foods as a result of the Special Delivery program (12/21).

All participants (21/21) recommended continuing or expanding the program to other pregnant women, and some even noted friends or family members who hoped to participate in a similar program. Not every participant gave an explanation about why the program should be continued, but many cited the positive impact on diet as a primary reason that other pregnant women could benefit from the program.

Participants submitted a weekly average of 1.9 photos and 5.4 text messages during the program. Responses to daily text message polls included text descriptions as well as photos, which provided qualitative evidence of foods consumed during program participation (Figure 5). A majority of the photos that were submitted (591/821 71%) showed foods consistent with those included in grocery deliveries. This suggests that participants ate the delivered foods and supports the acceptability of the program.

Table 3. Categories of end-of-study interview feedback and representative quotations.

Categories	Representative quotations
Delivery process	
Special Delivery is convenient	"It was convenient, like, you guys came when I actually needed something. And I liked it. I liked it a lot" "It's just easy, someone comes to your door and then you have your groceries"
Special Delivery foods were high-quality	"Everything that I got was fresh. I never got anything that was, you know, like spoiled or really close to the expiration date. So the shoppers did a really good job picking out, you know, the freshest food they could find"
Impact on diet	
Special Delivery helped me have a healthy diet	"I would pack myself a lunch and most of it was from the delivery" "It kinda reminded me to eat more and eat healthier throughout the day" "Every time we got a salad, we ate those that day"
Special Delivery improved my access to healthy foods	"Some days I wouldn't eat at all, but since I had groceries in the fridge, like fruit, vegetables, I would just grab like some and take it for a snack" "It was just more healthy options in the house" "Being able to actually have access…just made it that much easier"
Special Delivery helped me build healthy habits	"With this, it was like okay, I'm craving junk food, but maybe I'll have an apple or an orange" "A lot of the things that I got were things that I don't usually get at the store. So now when I go grocery shopping, those are things that I'm picking up"
Special Delivery helped me try new foods	"I was able to try different stuff that I don't normally get" "Let me just try it, let me eat it, so it don't go to waste"
Special Delivery could help other pregnant women	"It made me eat healthy stuff, so I guess it can make somebody else eat some healthy stuff too" "I think it's good for all pregnant women" "It's a little hard when you're craving something bad for you. But just, just having that convenience is very helpful. So I think I would definitely continue it and have other moms experience it as well"

Figure 5. Example photos of delivered foods submitted by study participants in response to text message surveys.

Discussion

Principal Results

Grocery delivery represents a well-established and inexpensive service that removes logistical barriers to obtaining healthy food but is underused by those who may need it most—young pregnant women. In this Special Delivery pilot study, we demonstrated the feasibility and acceptability of a grocery delivery–based weight gain and nutrition intervention among pregnant youth with low incomes. Grocery deliveries were successfully completed and confirmed by participants. Participants reported strong positive perceptions of the grocery delivery process and of the impact of delivered foods on health. Qualitative photo evidence submitted by participants demonstrated foods consistent with the healthy foods that were delivered. These results suggest the feasibility and acceptability of the Special Delivery program; participants were willing and able to receive grocery deliveries and eat the healthy foods that were delivered.

Comparison With Prior Work

Our finding that grocery orders can be placed and delivered is consistent with prior studies [31,32] that demonstrated the feasibility of 1-time grocery delivery for residents of urban food deserts and of online grocery shopping for individuals with low income receiving food stamps. Though we limited inclusion to those residing in the delivery region of a single grocery delivery service (Shipt), the feasibility of this intervention could likely be extrapolated to most geographic areas since grocery delivery services continue to expand. Interview feedback and photos of the foods that were eaten demonstrated the high acceptability of grocery delivery to our participant cohort. Participants reported that the grocery delivery service was easy to utilize, convenient, and helped improve their diet quality, which is consistent with the findings of related studies of grocery delivery implementations as nutrition interventions [33]. Our acceptability findings differ from those from older studies [21,34,35] that concerns about the freshness of perishable goods were a major barrier to online grocery shopping acceptability for low-income populations. Instead, nearly all participants in our study reported that high-quality perishable foods and fresh fruits and vegetables were delivered.

The finding of acceptability from our pilot was also consistent with recent findings that a majority of grocery shoppers would be willing to order groceries online [36]. Our research provides new evidence that grocery delivery of healthy foods is a feasible and acceptable intervention for pregnant young women with low incomes.

Limitations

Our assessment is not without limitations. We enrolled only a small cohort and have not yet explored the scalability of our methods. However, throughout our pilot, we focused on streamlining processes with the intention of creating a scalable method. For example, we now collect delivery instructions from participants during enrollment and provide participant contact information directly to Shipt deliverers to mimic real-life use of their service. Though participants did not order their own groceries, we are encouraged by the ease of the Shipt user interface and the potential for future participants to order on their own.

Conclusions

Grocery delivery is an inexpensive service that has the potential to increase access to healthy foods for those who face significant logistical barriers to obtaining healthy foods. In our study, grocery delivery of healthy foods was found to be feasible and acceptable to pregnant young women. Our findings suggest that large-scale evaluation of the impact on gestational weight gain and overall quality of diet could further extend the potential benefits of grocery delivery to low-income young women and other vulnerable populations.

Acknowledgments

Research reported in this publication was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number TL1TR002242 (awarded to IL). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The research is also supported by National Institutes of Health, Eunice Kennedy Shriver National Institute of Child Health & Human Development (Speaking Their Language: Using Social Media and Texting to Create an Adolescent-Centered Approach to Healthy Weight Gain During Pregnancy, 1K23HD083527-01A1; principal investigator: TC).

Conflicts of Interest

None declared.

References

1. Maconochie N, Doyle P, Prior S, Simmons R. Risk factors for first trimester miscarriage--results from a UK-population-based case-control study. BJOG 2007 Feb;114(2):170-186 [FREE Full text] [doi: 10.1111/j.1471-0528.2006.01193.x] [Medline: 17305901]
2. Ikem E, Halldorsson TI, Birgisdóttir BE, Rasmussen MA, Olsen SF, Maslova E. Dietary patterns and the risk of pregnancy-associated hypertension in the Danish National Birth Cohort: a prospective longitudinal study. BJOG 2019 Apr;126(5):663-673 [FREE Full text] [doi: 10.1111/1471-0528.15593] [Medline: 30675768]
3. Shin D, Lee KW, Song WO. Dietary patterns during pregnancy are associated with risk of gestational diabetes mellitus. Nutrients 2015 Nov 12;7(11):9369-9382 [FREE Full text] [doi: 10.3390/nu7115472] [Medline: 26569302]
4. Weight gain during pregnancy: reexamining the guidelines. Institute of Medicine and National Research Council. Washington DC: National Academies Press; 2010. URL: https://www.ncbi.nlm.nih.gov/books/NBK32813/pdf/Bookshelf_NBK32813.pdf [accessed 2020-12-14]
5. Joseph NP, Hunkali KB, Wilson B, Morgan E, Cross M, Freund KM. Pre-pregnancy body mass index among pregnant adolescents: gestational weight gain and long-term post partum weight retention. J Pediatr Adolesc Gynecol 2008 Aug;21(4):195-200. [doi: 10.1016/j.jpag.2007.08.006] [Medline: 18656073]
6. Aliyu MH, Luke S, Kristensen S, Alio AP, Salihu HM. Joint effect of obesity and teenage pregnancy on the risk of preeclampsia: a population-based study. J Adolesc Health 2010 Jan;46(1):77-82. [doi: 10.1016/j.jadohealth.2009.06.006] [Medline: 20123261]
7. Harper LM, Chang JJ, Macones GA. Adolescent pregnancy and gestational weight gain: do the Institute of Medicine recommendations apply? Am J Obstet Gynecol 2011 Aug;205(2):140.e1-140.e8 [FREE Full text] [doi: 10.1016/j.ajog.2011.03.053] [Medline: 21620365]
8. Siega-Riz AM, Viswanathan M, Moos M, Deierlein A, Mumford S, Knaack J, et al. A systematic review of outcomes of maternal weight gain according to the Institute of Medicine recommendations: birthweight, fetal growth, and postpartum weight retention. Am J Obstet Gynecol 2009 Oct;201(4):339.e1-339.14. [doi: 10.1016/j.ajog.2009.07.002] [Medline: 19788965]
9. Oken E, Rifas-Shiman SL, Field AE, Frazier AL, Gillman MW. Maternal gestational weight gain and offspring weight in adolescence. Obstet Gynecol 2008 Nov;112(5):999-1006 [FREE Full text] [doi: 10.1097/AOG.0b013e31818a5d50] [Medline: 18978098]
10. Morrison L, DeJonckheere M, Nichols LP, Smith DG, Plegue MA, McKee K, et al. Knowledge, behaviors, and social factors that influence pregnancy weight gain among youth ages 16-24 years. J Pediatr Adolesc Gynecol 2020 Feb;33(1):64-71 [FREE Full text] [doi: 10.1016/j.jpag.2019.10.001] [Medline: 31606389]
11. Chang T, Moniz MH, Plegue MA, Sen A, Davis MM, Villamor E, et al. Characteristics of women age 15-24 at risk for excess weight gain during pregnancy. PLoS One 2017;12(3):e0173790 [FREE Full text] [doi: 10.1371/journal.pone.0173790] [Medline: 28291802]
12. Mardis A, Anand R. A look at the diet of pregnant women (Insight 17 April 2000). Family Economics and Nutrition Review 2001;13(2):55-58.
13. Nunnery D, Ammerman A, Dharod J. Predictors and outcomes of excess gestational weight gain among low-income pregnant women. Health Care Women Int 2018 Jan;39(1):19-33 [FREE Full text] [doi: 10.1080/07399332.2017.1391263] [Medline: 29068777]
14. Panzera AD, Bryant CA, Hawkins F, Goff R, Napier A, Schneider T, et al. Mapping a WIC mother's journey. Social Marketing Quarterly 2017 Feb 21;23(2):137-154. [doi: 10.1177/1524500417692526]
15. Liu CH, Liu H. Concerns and structural barriers associated with WIC participation among WIC-eligible women. Public Health Nurs 2016 Dec;33(5):395-402 [FREE Full text] [doi: 10.1111/phn.12259] [Medline: 26956356]
16. Cole N. National survey of WIC participants. Office of Analysis Nutrition Evaluation. Alexandria, VA; 2001. URL: https://fns-prod.azureedge.net/sites/default/files/WICSurvey.pdf [accessed 2020-05-10]
17. Pooler J, Gleason SF. Comparison of WIC benefit redemptions in Michigan indicates higher utilization among Arab American families. J Nutr Educ Behav 2014;46(3 Suppl):S45-S52. [doi: 10.1016/j.jneb.2014.02.019] [Medline: 24809996]
18. Wise NJ. Pregnant adolescents, beliefs about healthy eating, factors that influence food choices, and nutrition education preferences. J Midwifery Womens Health 2015;60(4):410-418. [doi: 10.1111/jmwh.12275] [Medline: 26255801]

19. Neumark-Sztainer D, Story M, Perry C, Casey MA. Factors influencing food choices of adolescents: findings from focus-group discussions with adolescents. J Am Diet Assoc 1999 Aug;99(8):929-937. [doi: 10.1016/S0002-8223(99)00222-9] [Medline: 10450307]
20. Thomas P. Determinants of food choice and prospects for modifying food attitudes and behavior. In: Improving America's Diet and Health: From Recommendations to Action. Washington DC: National Academies Press; 1991.
21. Jilcott Pitts SB, Ng SW, Blitstein JL, Gustafson A, Niculescu M. Online grocery shopping: promise and pitfalls for healthier food and beverage purchases. Public Health Nutr 2018 Dec;21(18):3360-3376. [doi: 10.1017/S1368980018002409] [Medline: 30338751]
22. Brandt EJ, Silvestri DM, Mande JR, Holland ML, Ross JS. Availability of grocery delivery to food deserts in states participating in the online purchase pilot. JAMA Netw Open 2019 Dec 02;2(12):e1916444 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.16444] [Medline: 31790562]
23. Wing RR, Jeffery RW. Food provision as a strategy to promote weight loss. Obes Res 2001 Nov;9 Suppl 4:271S-275S [FREE Full text] [doi: 10.1038/oby.2001.130] [Medline: 11707553]
24. Jeffery RW, Wing RR, Thorson C, Burton LR, Raether C, Harvey J, et al. Strengthening behavioral interventions for weight loss: a randomized trial of food provision and monetary incentives. J Consult Clin Psychol 1993 Dec;61(6):1038-1045. [doi: 10.1037//0022-006x.61.6.1038] [Medline: 8113481]
25. Ebbeling CB, Feldman HA, Osganian SK, Chomitz VR, Ellenbogen SJ, Ludwig DS. Effects of decreasing sugar-sweetened beverage consumption on body weight in adolescents: a randomized, controlled pilot study. Pediatrics 2006 Mar;117(3):673-680. [doi: 10.1542/peds.2005-0983] [Medline: 16510646]
26. Council on Community Pediatrics, Committee on Nutrition. Promoting food security for all children. Pediatrics 2015 Nov;136(5):e1431-e1438 [FREE Full text] [doi: 10.1542/peds.2015-3301] [Medline: 26498462]
27. Hager ER, Quigg AM, Black MM, Coleman SM, Heeren T, Rose-Jacobs R, et al. Development and validity of a 2-item screen to identify families at risk for food insecurity. Pediatrics 2010 Jul;126(1):e26-e32. [doi: 10.1542/peds.2009-3146] [Medline: 20595453]
28. Anderson M. For teens, phone calls are reserved for closer relationships. Pew Research Center. 2015 Aug 15. URL: https://www.pewresearch.org/fact-tank/2015/08/17/for-teens-phone-calls-are-reserved-for-closer-relationships/ [accessed 2020-05-10]
29. Pew RC. Mobile fact sheet. Pew Research Center: Internet, Science & Tech. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/
30. Mervosh S, Lu D, Swales V. See which states and cities have told residents to stay at home June 17, 2020. The New York Times. URL: https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html [accessed 2020-05-10]
31. Appelhans BM, Lynch EB, Martin MA, Nackers LM, Cail V, Woodrick N. Feasibility and acceptability of Internet grocery service in an urban food desert, Chicago, 2011-2012. Prev Chronic Dis 2013 May 02;10:E67 [FREE Full text] [doi: 10.5888/pcd10.120299] [Medline: 23639762]
32. Lagisetty P, Flamm L, Rak S, Landgraf J, Heisler M, Forman J. A multi-stakeholder evaluation of the Baltimore City virtual supermarket program. BMC Public Health 2017 Oct 23;17(1):837 [FREE Full text] [doi: 10.1186/s12889-017-4864-9] [Medline: 29061141]
33. Gorin AA, Raynor HA, Niemeier HM, Wing RR. Home grocery delivery improves the household food environments of behavioral weight loss participants: results of an 8-week pilot study. Int J Behav Nutr Phys Act 2007 Nov 14;4:58 [FREE Full text] [doi: 10.1186/1479-5868-4-58] [Medline: 18001469]
34. Martinez O, Tagliaferro B, Rodriguez N, Athens J, Abrams C, Elbel B. EBT Payment for online grocery orders: a mixed-methods study to understand its uptake among SNAP recipients and the barriers to and motivators for its use. J Nutr Educ Behav 2018 Apr;50(4):396-402.e1. [doi: 10.1016/j.jneb.2017.10.003] [Medline: 29187304]
35. Rogus S, Guthrie JF, Niculescu M, Mancino L. Online grocery shopping knowledge, attitudes, and behaviors among SNAP participants. J Nutr Educ Behav 2020 May;52(5):539-545. [doi: 10.1016/j.jneb.2019.11.015] [Medline: 31870741]
36. Growing grocery online: innovation in digital commerce for CPG. The Nielsen Company. 2016 Jul 11. URL: https://www.nielsen.com/us/en/insights/article/2016/growing-grocery-online-innovation-in-digital-commerce-for-cpg/ [accessed 2020-04-20]

Abbreviations

WIC: Special Supplemental Nutrition Program for Women, Infants, and Children

Edited by G Eysenbach; submitted 18.06.20; peer-reviewed by G Kolenic, B Bais; comments to author 28.07.20; revised version received 31.07.20; accepted 28.10.20; published 24.12.20.

Please cite as:
Locher I, Waselewski M, Sonneville K, Resnicow K, Chang T
Grocery Delivery of Healthy Foods to Pregnant Young Women With Low Incomes: Feasibility and Acceptability Mixed Methods Study
JMIR Form Res 2020;4(12):e21602
URL: http://formative.jmir.org/2020/12/e21602/
doi:10.2196/21602
PMID:33361055

Original Paper

Automated Categorization of Systemic Disease and Duration From Electronic Medical Record System Data Using Finite-State Machine Modeling: Prospective Validation Study

Gumpili Sai Prashanthi[1], MTech; Ayush Deva[2], BTech; Ranganath Vadapalli[1], BSc; Anthony Vipin Das[1], FRCS

[1]Department of eyeSmart EMR & AEye, LV Prasad Eye Institute, Hyderabad, Telangana, India

[2]International Institute of Information Technology, Hyderabad , Telangana, India

Corresponding Author:
Anthony Vipin Das, FRCS
Department of eyeSmart EMR & AEye
LV Prasad Eye Institute
Kallam Anji Reddy Campus, L V Prasad Marg
Hyderabad, Telangana
India
Phone: 91 9885071960
Email: vipin@lvpei.org

Abstract

Background: One of the major challenges in the health care sector is that approximately 80% of generated data remains unstructured and unused. Since it is difficult to handle unstructured data from electronic medical record systems, it tends to be neglected for analyses in most hospitals and medical centers. Therefore, there is a need to analyze unstructured big data in health care systems so that we can optimally utilize and unearth all unexploited information from it.

Objective: In this study, we aimed to extract a list of diseases and associated keywords along with the corresponding time durations from an indigenously developed electronic medical record system and describe the possibility of analytics from the acquired datasets.

Methods: We propose a novel, finite-state machine to sequentially detect and cluster disease names from patients' medical history. We defined 3 states in the finite-state machine and transition matrix, which depend on the identified keyword. In addition, we also defined a state-change action matrix, which is essentially an action associated with each transition. The dataset used in this study was obtained from an indigenously developed electronic medical record system called eyeSmart that was implemented across a large, multitier ophthalmology network in India. The dataset included patients' past medical history and contained records of 10,000 distinct patients.

Results: We extracted disease names and associated keywords by using the finite-state machine with an accuracy of 95%, sensitivity of 94.9%, and positive predictive value of 100%. For the extraction of the duration of disease, the machine's accuracy was 93%, sensitivity was 92.9%, and the positive predictive value was 100%.

Conclusions: We demonstrated that the finite-state machine we developed in this study can be used to accurately identify disease names, associated keywords, and time durations from a large cohort of patient records obtained using an electronic medical record system.

(JMIR Form Res 2020;4(12):e24490) doi:10.2196/24490

KEYWORDS

electronic health records; data analysis; machine learning; algorithms; ophthalmology

Introduction

Electronic medical record (EMR) systems have been increasingly replacing paper-based records; using these systems has advantages such as increased efficiency and standardized quality, thereby enabling accurate clinical documentation [1]. Research that is dependent on reviewing paper records is not only cumbersome but also prone to human errors. The amount of time taken to retrieve and analyze large volumes of data from EMR systems is minimal compared to the manual process. Moreover, the adoption of EMR systems has led to the availability of diverse sources of clinical information, such as demographic data, history of diagnosis, prescriptions, and laboratory test results, which have established EMR systems as a treasure trove for large-scale analysis of health data. As a result, to obtain meaningful insights, there is a need to extract useful information and patterns from the rapidly growing volumes of data.

In general, 3 types of data are available as EMRs: structured, semistructured, and unstructured data [2]. Fixed-mode databases contain basic information and are usually used to store structured data. Unstructured data includes reports; records regarding surgery, medical history, and discharge; and clinical notes. One of the major challenges in the health care sector is that approximately 80% of the data remains unstructured and unused after it has been generated [3]. Since it is difficult to handle this sort of unstructured data obtained from EMRs, it tends to be neglected for analysis in most hospitals or medical centers [4]. Therefore, there is a need to analyze unstructured big data in health care systems so that we can optimally utilize the data and unearth all possible unexploited information from it.

The aim of this study was to extract a list of mentioned diseases and associated keywords, along with time durations, from the indigenously developed EMR system eyeSmart, which has been implemented across a large multitier ophthalmology network in India. We also aimed to describe the possibility of analytics from the datasets thus acquired.

Methods

Data Extraction

We retrieved systemic disease information of a subset of patients who presented to a large multitier ophthalmology network in India between August 2010 and December 2019 by using eyeSmart EMR system [1]. The dataset analyzed included the past medical history of patients and contained 10,000 records of distinct patients. From the given plaintext data about the medical history of the patients, we retrieved the names of systemic disease(s) from a fixed set of known disease names (Textbox 1) documented in the patients’ past medical history column and the duration of the disease, using Python and the techniques mentioned below.

Textbox 1. Search terms used to retrieve disease names and other associated keywords from the dataset.

Disease names (keyword) and their associated keywords:

- Diabetes mellitus
    - DM
    - Insulin
    - FBS
    - PPBS
    - IDDM
- Hypertension
    - HTN
- Asthma
- Acid peptic disease
    - Gastric
    - Ulcer
- Hypothyroidism
- Hyperthyroidism
- Rheumatoid arthritis
    - RA
- Allergy
- Tuberculosis
- Sinusitis
- Arthritis
    - Joint pain
- Coronary artery disease
    - CAD
- Cholesterol
- Migraine
- Cancer
- Paralysis
- Spondylitis
- Hepatitis
- Epilepsy
    - Fits
    - Seizures
- Malaria

Data Availability

The dataset analyzed during the current study is not publicly available as it contains confidential patient information, but it can be made available from the corresponding author on reasonable request.

Ethical Approval

This study was approved by the Institutional Review Board of LV Prasad Eye Institute, Hyderabad (Ethics Ref. No. LEC BHR-R-09-20-497), and all procedures were in accordance with the tenets of the Declaration of Helsinki. All data were fully anonymized prior to access by the study group.

Assumptions About the Data

We made the following assumptions about the data in order to set some baselines for the information retrieval task:

1. The names of systemic diseases were spelled correctly.
2. Duration of the disease (if it exists) always followed the name of the disease and did not precede the documented disease.
3. If a disease name (D1) was followed by another disease name (D2) without any duration tag in between, then the duration for D1 was assumed to be missing and the next duration tag encountered would be associated with D2.

Given the unstructured plaintext data about systemic diseases and their durations, the following 2 steps were used to extract useful information and convert it into a structured data format.

Tag Identification

This involved the identification of disease names and duration tags in the extracted data. The task was to identify the presence of one or more of the enlisted diseases (Textbox 1) in the plaintext data. Since we assumed that the disease names were spelled correctly, we used string matching in Python to check if any disease names were present. Similarly, to identify the duration tags, we used regular expressions in Python to identify both (1) the value of duration (ie, a number) and (2) the unit of duration (ie, day, week, month, or year)

Clustering

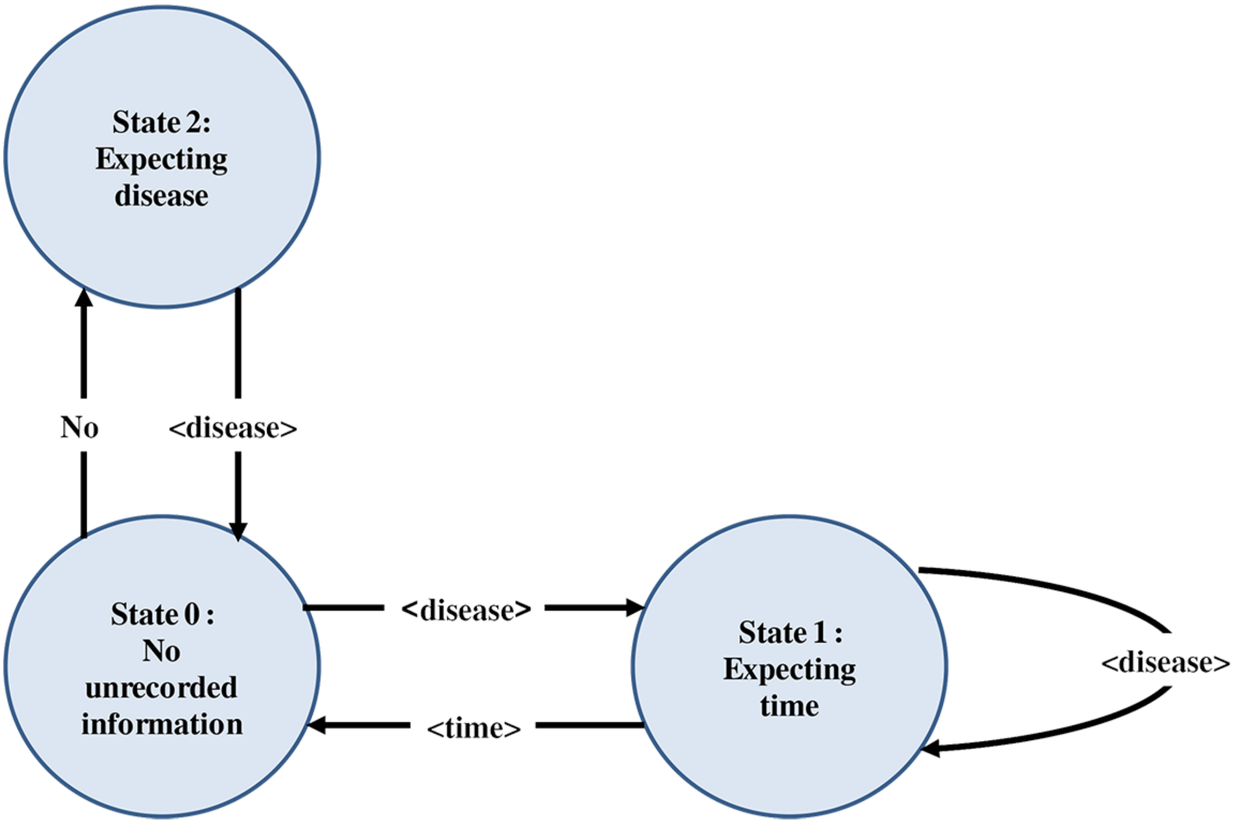
This step involved correctly clustering the information, that is, finding and establishing the relations between different tags (in this case, duration and disease name).

Specifically, once the duration tags were identified, it was important to associate the correct duration with the corresponding disease, which was a challenging part.

Therefore, we propose a novel, finite-state machine (FSM) to sequentially detect and cluster disease name(s) from the patient’s medical history records.

We defined 3 states in our FSM and the transition matrix that depends on the identified tag. In addition, we also defined a state-change action matrix, which is essentially an action associated with each transition. These are explained in detail below and illustrated in Figure 1.

Figure 1. Flowchart depicting the 3 states of the finite-state machine modeling.



State Definition

State 0

This is the starting state. It has no unrecorded information. In this state, previous disease is “null” (prev_disease = NULL).

State 1

This is the state where we have found a disease previously that has not yet been recorded or stored in prev_disease. Here, we expect to find a time-matching regular expression to complete the record for that disease.

State 2

This is the state that is reached after encountering a “NO” string from State 0. A “NO” string indicates absence of the particular disease that follows the word “NO.” This means that the next sentence is about a disease that is not present and should not be included in the list. The state-change matrix and state-change action matrix are presented in Tables 1 and 2, respectively.

Table 1. State-change matrix.

Disease state	Input string		
	“NO”	Disease	Time
0	2	1	Error
1	Error	1	0
2	Error	0	Error

Table 2. State-change action matrix.

State	Input string		
	“NO”	Disease	Time
0	N/A[a]	1. Print(Found a disease. Hoping for it to be followed by a time since when disease existed.) 2. Update prev_disease = <disease>.	Print(Warning: Found a time unit before finding a disease)
1	Print(Disease Name followed by NO. Not according to how it should be)	1. Record <prev_disease, No time>. 2. Print(Warning: prev_disease was not followed by a time.) 3. Update prev_disease = <disease>	1. Record<prev_disease, <time>>. 2. Print(Successfully detected disease and time since when). 3. Update prev_disease = NULL.
2	Print(Error: Found two consecutive NOs)	Print(Detected a “NO Disease statement. Ignoring and not recording.)	Print(Found a time unit after NO. Something wrong)

[a] N/A: not applicable.

FSM Output Measurements

We measured sensitivity, positive predictive value (PPV), and accuracy of the FSM to identify the disease name and associated keywords as well as the associated duration.

Results

We evaluated the information extraction accuracy of the FSM by comparing the results to those of an expert human gold standard. The human had good knowledge about the medical terms used and their diagnoses. The report was further crosschecked by another person to minimize human error. In all, 100 records were randomly sampled and manually annotated and compared to the output of the algorithm. The record was annotated as true positive only when all the disease names and any associated keywords were accurately extracted along with the accurate disease duration.

To compare the predictions of the FSM to a gold standard (ie, manually annotated data, in our case), a confusion matrix was used. Table 3 represents a generic 2×2 confusion matrix used to identify the diagnosis. Table 4 represents a confusion matrix used to identify the duration associated with that particular diagnosis.

For the extraction of disease names and associated keywords, we reported an accuracy of 95%, sensitivity of 94.9%, and PPV of 100%. For the extraction of the disease duration, we reported an accuracy of 93%, sensitivity of 92.9%, and PPV of 100%.

Table 3. Confusion matrix to determine the diagnosis of disease (n=100).

Gold-standard method	FSM[a] result	
	Predicted “Yes”	Predicted “No”
Actual “Yes”	94	5
Actual “No”	0	1

[a] FSM: finite-state machine.

Table 4. Confusion matrix to determine duration of the associated diagnosis (n=100).

Gold-standard method	FSM[a] result	
	Predicted "Yes"	Predicted "No"
Actual "Yes"	92	7
Actual "No"	0	1

[a] FSM: finite-state machine.

Discussion

Principal Findings

In this study, we demonstrated that FSM can be used to accurately identify the disease name, associated keywords, and disease duration from a large cohort of patient records obtained using an EMR system that has been implemented across a large, multitier ophthalmology network in India. Many previous studies have used regular expressions and natural language processing (NLP) to extract disease names or keywords [5-17]. Hobbs et al [18] used cascading finite-state automatas for extracting information from natural language text [18]. Leroy et al [19] used finite-state automata to structure the relation between extract entities, but attempts to extract the duration of the disease along with the disease name itself have not been made previously.

A variety of valuable medical information is stored in texts that are unstructured, but there are many challenges in dealing with such data as the text may contain many errors, incorrect usage of grammar, and improper structural framework, which would increase the challenges in analyzing and processing of data.

Unstructured data gives a wider picture about patient data and aids clinicians in connecting the dots and presenting a more accurate picture of the health of the patient. Extracting useful information from these records help doctors to identify a patient's medical history and also make important predictions.

There are inherent challenges in how information in an unstructured format is inputted into the EMR. This is governed by the training, literacy, and typing skill of the user in question. These challenges include errors in spelling and nonconformity of the structure of the data inputted. Automating the analysis of data in such formats helps reduce the time taken for manual mining of information.

Comparison With Prior Work

In a study on asthma by Zeng et al [5], which involved using NLP for extracting principal diagnosis, comorbidity, and smoking status, the accuracy of the algorithm was 82%, 87%, and 90%, respectively. Rosier et al [6] used clinical records to extract data on pacemaker implantation procedures by using regular expressions. The system extracted information with a very high PPV (>95%) and sensitivity (>90%). In a study by Murtaugh et al [7], which involved extraction of body weight values from clinical notes, the accuracy was 98.3% and precision was 98.8%. These values are similar to our findings in this study involving identification of systemic diseases and their durations.

Systemic diseases are frequently considered to be the underlying cause of many medical conditions. Systemic disease history is a particularly important component in the examination of patients with eye disease. Various systemic diseases affect the eye, notably diabetic retinopathy, dry eye disease, cataract, and thyroid eye disease [20]. An understanding of the duration of the systemic diseases is vital to prognosticate the severity of the ocular condition and the treatment outcomes.

In a systemic review based on extracting information from the text of EMRs to improve case detection, Ford et al [8] compared the accuracy of case-detection algorithms by comparing codes and text. For codes-only algorithms, the median sensitivity was 61.7% and PPV was 72%. For text-only algorithms, the median sensitivity was 78.1% and PPV was 73%. Moreover, for a combination of text and codes, the median sensitivity was 78.1% and PPV was 86%. The medical conditions included in this review were respiratory infections, bowel disease, cancer, and diabetes. The algorithm sensitivity ranged from 48.4% to 99.2%, specificity ranged from 90% to 99.4%, and PPV ranged from 54% to 97.9% [8].

Zheng et al [13] used both structured and unstructured EMRs for developing and testing a web-based diabetes case detection algorithm. The NLP-based algorithm had a PPV of 90%. Petch et al [14] extracted 15 clinical features from dictated ambulatory consult notes by using a commercially available NLP-based tool. NLP performed best for features that were classified as simple, yielding an overall accuracy of 96%. However, the performance was lower for other features that were of moderate and complex linguistic complexity.

The scope of this study, as the first experiment of this nature, was to successfully categorize systemic diseases and their durations from a cohort of patient records. The next models will focus on categorizing clinical findings based on slit lamp examination of various parts of the eye and the plan of management written by the health care provider. The tasks we undertook in this study were relatively challenging. The major challenges were that a patient's medical history may contain information about multiple diseases. The presence of the name of a disease does not always imply the patient was diagnosed with that disease. We can have instances where a doctor may write that the patient had no history of a particular disease. Moreover, not all disease names identified in the data had a duration associated with them. For example, there can be 3 disease names and only 2 duration related tags. All these challenges were addressed by the current methodology of FSM, as described in this study.

Study Limitations

One of the limitations of this study is that if the duration of disease were preceded by the disease name, it could not be identified and that disease names could not be identified if there were any spelling mistakes.

Since the current dataset had negligible spelling mistakes, and the disease names were always followed by the duration, the state space of the current FSM was small. However, an advantage of modeling this as an FSM is that it can be easily extended to run on datasets where these assumptions do not hold true. Thus, we propose that FSM is a very robust framework to address challenges of automated systemic disease and duration categorization. Our findings also suggest that this method can be used more generally for both clinical and research purposes, to identify the disease and duration.

Future Directions

Future work involves using the FSM on more datasets, understanding the complexities of the unstructured datasets that are used as inputs, and incorporating more changes to make the FSM more robust. This is an ongoing process of periodically analyzing the input data to modify the state changes to enable a more accurate categorization of the required variables.

The adoption of EMR in a large country like India is rather low. There are various associated challenges, but the potential long-term benefits for research and education are promising. Structured datasets from the EMR are crucial for any meaningful research to be conducted. Unstructured datasets also need to be analyzed in an automated fashion to minimize the time required for analyses.

Conclusions

In conclusion, we present a novel technique that was developed to analyze unstructured data of systemic diseases and their durations in a large cohort of patient records in a multitier ophthalmology network in India.

Acknowledgments

We wish to acknowledge the support of our Department of eyeSmart EMR & AEye team and are especially grateful to Mr. Mohammad Pasha for his assistance.

Conflicts of Interest

None declared.

References

1. Das A, Kammari P, Vadapalli R, Basu S. Big data and the eyeSmart electronic medical record system - An 8-year experience from a three-tier eye care network in India. Indian J Ophthalmol 2020 Mar;68(3):427-432 [FREE Full text] [doi: 10.4103/ijo.IJO_710_19] [Medline: 32056994]
2. Sun W, Cai Z, Li Y, Liu F, Fang S, Wang G. Data processing and text mining technologies on electronic medical records: a review. J Healthc Eng 2018;2018:4302425 [FREE Full text] [doi: 10.1155/2018/4302425] [Medline: 29849998]
3. Kong H. Managing unstructured big data in healthcare system. Healthc Inform Res 2019 Jan;25(1):1-2 [FREE Full text] [doi: 10.4258/hir.2019.25.1.1] [Medline: 30788175]
4. Rusu O, Halcu I, Grigoriu O, Neculoiu G, Sandulescu V, Marinescu M. Converting unstructured and semi-structured data into knowledge. : IEEE; 2013 Presented at: 2013 11th RoEduNet International Conference; Jan 7-13, 2019; Sinaia, Romania p. 1-4. [doi: 10.1109/roedunet.2013.6511736]
5. Zeng QT, Goryachev S, Weiss S, Sordo M, Murphy SN, Lazarus R. Extracting principal diagnosis, co-morbidity and smoking status for asthma research: evaluation of a natural language processing system. BMC Med Inform Decis Mak 2006 Jul 26;6:30 [FREE Full text] [doi: 10.1186/1472-6947-6-30] [Medline: 16872495]
6. Rosier A, Burgun A, Mabo P. Using regular expressions to extract information on pacemaker implantation procedures from clinical reports. AMIA Annu Symp Proc 2008 Nov 06:81-85. [Medline: 18998970]
7. Murtaugh M, Gibson B, Redd D, Zeng-Treitler Q. Regular expression-based learning to extract bodyweight values from clinical notes. J Biomed Inform 2015 Apr;54:186-190 [FREE Full text] [doi: 10.1016/j.jbi.2015.02.009] [Medline: 25746391]
8. Ford E, Carroll J, Smith H, Scott D, Cassell J. Extracting information from the text of electronic medical records to improve case detection: a systematic review. J Am Med Inform Assoc 2016 Sep;23(5):1007-1015 [FREE Full text] [doi: 10.1093/jamia/ocv180] [Medline: 26911811]
9. Bui DDA, Zeng-Treitler Q. Learning regular expressions for clinical text classification. J Am Med Inform Assoc 2014;21(5):850-857. [doi: 10.1136/amiajnl-2013-002411] [Medline: 24578357]
10. Liang Z, Chen J, Xu Z, Chen Y, Hao T. A pattern-based method for medical entity recognition from Chinese diagnostic imaging text. Front Artif Intell 2019 May 14;2. [doi: 10.3389/frai.2019.00001]
11. Guetterman TC, Chang T, DeJonckheere M, Basu T, Scruggs E, Vydiswaran VV. 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]
12. Kaufman DR, Sheehan B, Stetson P, Bhatt AR, Field AI, Patel C, et al. Natural language processing-enabled and conventional data capture methods for input to electronic health records: a comparative usability study. JMIR Med Inform 2016 Oct 28;4(4):e35 [FREE Full text] [doi: 10.2196/medinform.5544] [Medline: 27793791]

13. Zheng L, Wang Y, Hao S, Shin AY, Jin B, Ngo AD, et al. Web-based real-time case finding for the population health management of patients with diabetes mellitus: a prospective validation of the natural language processing-based algorithm with statewide electronic medical records. JMIR Med Inform 2016 Nov 11;4(4):e37 [FREE Full text] [doi: 10.2196/medinform.6328] [Medline: 27836816]
14. Petch J, Batt J, Murray J, Mamdani M. Extracting clinical features from dictated ambulatory consult notes using a commercially available natural language processing tool: pilot, retrospective, cross-sectional validation study. JMIR Med Inform 2019 Nov 01;7(4):e12575 [FREE Full text] [doi: 10.2196/12575] [Medline: 31682579]
15. Chen T, Dredze M, Weiner JP, Hernandez L, Kimura J, Kharrazi H. Extraction of geriatric syndromes from electronic health record clinical notes: assessment of statistical natural language processing methods. JMIR Med Inform 2019 Mar 26;7(1):e13039 [FREE Full text] [doi: 10.2196/13039] [Medline: 30862607]
16. Agaronnik ND, Lindvall C, El-Jawahri A, He W, Iezzoni LI. Challenges of developing a natural language processing method with electronic health records to identify persons with chronic mobility disability. Arch Phys Med Rehabil 2020 Oct;101(10):1739-1746. [doi: 10.1016/j.apmr.2020.04.024] [Medline: 32446905]
17. Koleck T, Dreisbach C, Bourne P, Bakken S. Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. J Am Med Inform Assoc 2019 Apr 01;26(4):364-379 [FREE Full text] [doi: 10.1093/jamia/ocy173] [Medline: 30726935]
18. Hobbs J, Appelt D, Bear J, Israel D, Kameyama M, Stickel M. FASTUS: A Cascaded Finite-State Transducer for Extracting Information from Natural-Language Text. In: Roche E, Schabes Y, editors. Finite-state Language Processing. Cambridge, Massachusetts: MIT Press; May 1997:383-406.
19. Leroy G, Chen H, Martinez JD. A shallow parser based on closed-class words to capture relations in biomedical text. J Biomed Inform 2003 Jun;36(3):145-158 [FREE Full text] [doi: 10.1016/s1532-0464(03)00039-x] [Medline: 14615225]
20. Pinazo-Durán MD, Zanón-Moreno V, García-Medina JJ, Arévalo JF, Gallego-Pinazo R, Nucci C. Eclectic ocular comorbidities and systemic diseases with eye involvement: a review. Biomed Res Int 2016;2016:6215745 [FREE Full text] [doi: 10.1155/2016/6215745] [Medline: 27051666]

Abbreviations

EMR: electronic medical record
FSM: finite-state machine
NLP: natural language processing
PPV: positive predictive value

Edited by C Lovis; submitted 22.09.20; peer-reviewed by P Loher, L Rusu; comments to author 07.11.20; revised version received 12.11.20; accepted 17.11.20; published 17.12.20.

Please cite as:
Sai Prashanthi G, Deva A, Vadapalli R, Das AV
Automated Categorization of Systemic Disease and Duration From Electronic Medical Record System Data Using Finite-State Machine Modeling: Prospective Validation Study
JMIR Form Res 2020;4(12):e24490
URL: http://formative.jmir.org/2020/12/e24490/
doi:10.2196/24490
PMID:33331823

Original Paper

Starter Kit for Geotagging and Geovisualization in Health Care: Resource Paper

Quan Do[1], DPhil; David Marc[2*], DPhil; Marat Plotkin[3*], MD; Brian Pickering[1*], MD; Vitaly Herasevich[1*], DPhil, MD

[1]Mayo Clinic, Rochester, MN, United States

[2]College of St Scholastica, Duluth, MN, United States

[3]Emory University, Atlanta, GA, United States

[*]these authors contributed equally

Corresponding Author:
Vitaly Herasevich, DPhil, MD
Mayo Clinic
200 First St SW
Rochester, MN
United States
Phone: 1 507 255 4055
Email: vitaly@mayo.edu

Abstract

Background: Geotagging is the process of attaching geospatial tags to various media data types. In health care, the goal of geotagging is to gain a better understanding of health-related questions applied to populations. Although there has been a prevalence of geographic information in public health, in order to effectively use and expand geotagging across health care there is a requirement to understand other factors such as the disposition, standardization, data sources, technologies, and limitations.

Objective: The objective of this document is to serve as a resource for new researchers in the field. This report aims to be comprehensive but easy for beginners to understand and adopt in practice. The optimal geocodes, their sources, and a rationale for use are suggested. Geotagging's issues and limitations are also discussed.

Methods: A comprehensive review of technical instructions and articles was conducted to evaluate guidelines for geotagging, and online resources were curated to support the implementation of geotagging practices. Summary tables were developed to describe the available geotagging resources (free and for fee) that can be leveraged by researchers and quality improvement personnel to effectively perform geospatial analyses primarily targeting US health care.

Results: This paper demonstrated steps to develop an initial geotagging and geovisualization project with clear structure and instructions. The geotagging resources were summarized. These resources are essential for geotagging health care projects. The discussion section provides better understanding of geotagging's limitations and suggests suitable way to approach it.

Conclusions: We explain how geotagging can be leveraged in health care and offer the necessary initial resources to obtain geocodes, adjustment data, and health-related measures. The resources outlined in this paper can support an individual and/or organization in initiating a geotagging health care project.

(JMIR Form Res 2020;4(12):e23379) doi:10.2196/23379

KEYWORDS

geographic mapping; medicalGIS guidelines; information storage and retrieval; mapping; geotagging; data visualization; population; public health

Introduction

Background

Medical geographic information systems (MedicalGIS), in general, and geotagging, in particular, have been employed in epidemiological research and population health practice to understand the correlation between locational information and health or health care. The instruments supporting MedicalGIS include GIS, big data, artificial intelligence such as agent-based modeling, cellular automata, disease surveillance, and analytical models. The uses of this technology in public health include but are not limited to estimating possible access to care, exploring the population distributions of health (eg, neighborhood socioeconomic conditions and population density), determining spatial allocations of diseases, and identifying the proximity to environmental health hazards. While GIS is a type of system that combines multilayers of geographic data to smoothly facilitate information and knowledge about data that is related to the location [1], geotagging is the process of attaching geographical identification tags to various media such as videos, photos, and websites [2]. These geospatial tags carry information such as longitude, latitude, name of a location, and distance. Geotagging carries both pros (eg, improving safety on the trail, targeting marketing, providing more context) and cons (eg, privacy and security, overtourism) [2].

The application of geotagging in health care traces back before the modern medicine era and the development of computers. During the 1854 cholera outbreak in London, Dr John Snow traced the cause of this pandemic by identifying that clusters of the disease plotted on a map near a city water pump [3]. This early example of using location with other data sources exemplifies the potential opportunities that geotagging can serve in health care. Today, the field that focuses on the use of geographic information and methods to study disease and health is known as medical geography or MedicalGIS. MedicalGIS began with early maps to understand the spread of disease and now includes the application of GIS to visualize and analyze trends related to the treatment, prevention, progression, and impact of disease [4].

With the evolution of medical geography, there has been greater national attention on creating a health care system to support the analysis of geographic location. An example of a national effort to adopt geotagging is from the US government's report Healthy People 2010, [5] which set a goal of geotagging 100% of all national, state, and local health datasets. Unfortunately, in the final report for this initiative, there was no measurable improvement in geotagging major health systems [5]. Despite the limitations in a recent national endeavor to adopt geotagging in health care data, there are clear examples where geotagging has shown promise.

Today, geotagging is used in many industries including retail, government, insurance, technology, and health care. In health care, the goal of geotagging is to gain better understanding of health-related questions applied to populations. Traditionally, there has been a prevalence of geographic information in public health. However, in order to effectively use and expand geotagging in health care, there is a requirement to understand other factors such as disposition, standardization, data sources, and limitations. It is important to have knowledge of the accompanying software to make use of the data for application. This paper will consolidate basic knowledge to assist in the initiation of a geotagging analysis in health care. The paper is designed as a resource document by suggesting the optimal geocodes, their sources, and a rationale for their use.

Geotagging in Public Health

Geotagging has been used for multiple purposes in health care. Some of the main benefits of geotagging are (1) being a better tool for health care data's visualization [6,7], (2) supporting time domain data analysis [8], (3) facilitating cross-analysis of multiple data types [9-11], and (4) enabling automatic data analysis [12-14].

Geotagging plays a crucial role in the visualization of geographically related health care data. Geotagging facilitates not only the analysis of this type of data, but also policy decision making [6], planning of new research, and data acquisition activities [7]. In a recent research study, geographical data visualization was used to explore the distribution and potential level of fluoride and arsenic in drinking water in Mexico in order to approximate the corresponding health burden for proper public health policies [6]. The study confirmed the association between fluoride contamination and aridity, together with the variation of fluoride concentrations in arid and humid states [6]. In China, GPS and GIS were used to locate and identify the distribution of malaria. Mobile terminal data acquisition was developed to collect such data for the purpose of malaria prevention and control [7].

Geotagging also helped to analyze geographically related time-varying data. It can be employed to evaluate the effectiveness of interventions (eg, evaluating the effectiveness of social distancing for individual neighborhoods in order to provide a more reasonable distribution of resources to assist high-risk populations) [15], predict trends in population health and health care (eg, using Twitter data to advance Zika virus surveillance) [16], determine the urgency level of interventions (eg, generating acute ischemic stroke alert or detecting COVID-19 symptoms, testing access, and recovery) [17,18], and identify areas that require further investigation (eg, worsening of the population health due to environmental pollution) [19]. Researchers of a study conducted in Japan in 2018 obtained geographical movement of physicians to determine the most favorable location and the median distance of geographical movement of female and male groups of health care providers [8]. The authors concluded that physicians preferred to work in urban areas but there was a higher rate of female physicians who had their first preference of working location as urban areas. Also, the overall distance of geographical movement among female providers was lower compared with male physicians. Finally, the greatest moving distance occurred between their second and fourth years after obtaining their medical license [8].

Data integration is essential in this era of big data. Geotagging contributes an important role in supporting cross-analysis of multiple data types such as environmental data, industrial data, demographic data, and sociological data. GIS has been used to determine potential patients and target locations for marketing purposes at hospitals [9]. Remote sensing data was integrated with geotagging to track the water quality of the Ismailia Canal in Egypt [10]. Transportation data generated and collected from different types of sensors was also used in an artificial information (AI)-based deep learning system to track patterns in the mobility of residents for developing better transportation models of smart cities [11].

When all the data are connected through a common geospatial variable, it is possible to develop algorithms for a software system to automatically analyze the data for a possible correlation as soon as new data become available and mark the possible correlation for further analysis. The development of a geotagging system for real-time early warning for flooding in Vietnam was recently reported [12]. Also, a geotagging project in Thailand was completed to identify the distribution of dengue hemorrhagic fever cases with spatial information. The project was conducted with an aim to provide appropriate interventions focusing on controlling and preventing the disease and developing strategic planning [13].

In the United States, a group of public health researchers obtained about 80 million geotagged tweets from more than 600,000 nonduplicate Twitter users over a year period in order to determine the associations between Twitter indicators such as happiness, diet, substance use (alcohol, drug, smoking), physical activity, and the county-level health outcomes. The study drew conclusions for the probable associations between Twitter indicators of happiness, food, and physical activity and the decrease in premature mortality, obesity, and physical inactivity rate [14]. Tweets related to alcohol use also indicated a greater alcohol-use correlated mortality rate. The study indicated that real-time data collected from social media platforms can make it possible for public health experts to detect movement of norms, sentiment, and behaviors that may signify health emerging issues or outbreaks [14].

The recent outbreak of the coronavirus emphasizes the crucial role of MedicalGIS. An AI-based app was developed to predict whether a patient had been infected with the COVID-19 coronavirus by the sound of their voice, breathing, and coughing [20]. Even though this technology was expected to support the detection of emerging hotspots, its use of a camera to collect location information, voice, and sound carried the potential for data compromise. The app was reported to have a not very high accuracy rate of 70% [20]. Given the risk of data leaking and low accuracy rate, whether the benefits outweigh the risks is still under debate. Although experts claimed that thermal screening does not yet have the ability to detect the coronavirus, it was an effective method to screen Ebola patients at airports [21]. Recently, there was also an effort to employ deep learning and thermal-based cameras to detect coronavirus patients based on body temperature in public places [22]. GPS data can be mapped with these systems to recognize the patient's location and determine the distribution of patients in a region. In the United States, Apple and Google developed a contact tracing app named COVID-19 Exposure Notifications to support health care agencies in tracking the places that an infected patient has been and the people with whom they have been in close proximity [23]. The US government also has been discussing methods of using smartphone location data to fight the coronavirus with technology companies and public health experts, including tracking whether people are maintaining appropriate social distancing [23]. However, preserving privacy policies presented barriers to the implementation of such tools. For other types of data, removing individual identifiers can be sufficient to protect people's privacy. But that may not be the case with geotagging data.

Despite of the advantages and disadvantages, geotagging and geovisualization in health care are essential due to their abilities to support decision making, preventive care, and epidemiological efforts. The challenge lies in identifying situations where these solutions would be most effective and beneficial.

Geotagging Best Practices: How to Get Started

To effectively use geotagging in health care, it is necessary to use the correct types of data (ie, base layers). The North American Association of Central Cancer Registries Inc created a geotagging best practice guide for the cancer registry community to ensure a high level of confidence, reliability, standardization, and accuracy in geotagging endeavors [24]. This guide outlines the components of geotagging with information to work effectively with geocodes at national, state, county, and city-block levels.

Another example of a best practice is the Digital National Spatial Data framework [25], which is meant to "assemble geographic information that describes the arrangement and attributes of features and phenomena on the Earth. The infrastructure includes the materials, technology, and people necessary to acquire, process, store, and distribute such information to meet a wide variety of needs" [26]. This framework is unique in that spatial data are not the only component; of equal importance are the people, organizations, and technology that make spatial analysis a national effort. The Digital National Spatial Data framework also proposes methods for disseminating geospatial data and illustrates the importance of funding geotagging projects.

Despite the motivation to create greater standardization and develop frameworks to assemble the information, there is an immediate need to understand the basic elements of geotagging and the data sources where the geocodes can be referenced. Therefore, the remainder of this paper will examine the necessary base layers of geotagging that can be adopted in health care and examine the sources of information that can assist in the implementation of MedicalGIS.

Objectives

The objective of this paper is to consolidate basic knowledge to help medical professionals begin using geotagging in health care. This paper is designed as a resource document by suggesting the optimal geocodes, their sources, and a rationale for use. Geotagging limitations are also discussed.

Methods

This paper consolidates the available geotagging resources (free and for fee) that can be leveraged by researchers and quality improvement personnel to effectively perform geospatial analyses primarily targeting the US health care and discusses the following layers and workflow of the geotagging framework:

1. Geotagging (first layer): translating of text-based descriptions to map coordinates and placing the data on the map
2. Geomapping (middle layer): finding the appropriate or available granularity level (city, county, zip code, and country) of the data you are working on and mapping/tagging the data to that geographic level
3. Adjustment and analysis: looking at the other data available for this geographic location and comparing this location to

other locations with similar data and calculating necessary statistics
4. Conclusions and action: evaluating whether the outcome is good or bad, drawing conclusions, and recommending further research

The paper concludes with a summarization of tools that could be used for geotagging, including software applications that are either licensed or open-source platforms.

Results

Information on the base geotagging layers that can be used to place a data point on a map is shown in Multimedia Appendix 1. The base layers are presented at various levels of granularity. At the most practical granular level is the ZIP+4 code, which can be used to present data at a geographic layer within a 5-digit zip code mapping segment. At the least granular level is the country, which presents data for each country in the world. The information is presented in order from the most granular to the least granular and includes a proposed preferred name for specific variables in future datasets.

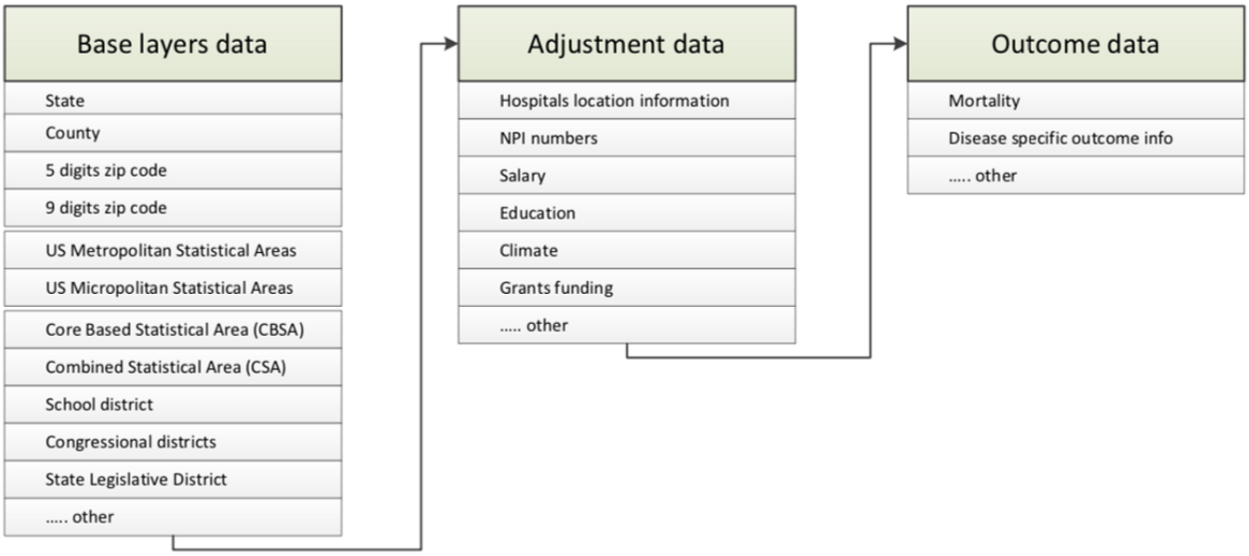
Information on resources that can be used to adjust geocodes is shown in Multimedia Appendix 2. For instance, the intent of an analysis may be to assess the mortality rates for heart failure across all hospitals in the United States. In order to effectively plot this data a zip code alone will not suffice. The distinct hospitals with accompanying geocodes need to be known. Multimedia Appendix 2 offers resources that can be used for adjustment to effectively conduct a spatial analysis.

Information on data that can be used to evaluate outcomes based on a geographic location and adjustment data is shown in Multimedia Appendix 3. The data includes information such as health outcomes, financial information, prevalence of chronic diseases, and patient satisfaction. The intent of Multimedia Appendix 3 is to present a measured outcome at a specific geographic level (from Multimedia Appendix 1) with or without accompaniment of the adjustment data (from Multimedia Appendix 2).

Although the information in these tables is not an exhaustive list of resources that can be used to effectively adopt geospatial analysis in a health care setting, the resources are still comprehensive and applicable to commonly conducted analyses. These resources can be adopted to initiate an evaluation of hospital performance, patient demographics, health disparities, and the incidence of disease.

Figure 1 offers a graphical depiction of how the data from Multimedia Appendices 1-3 can be used in concert to conduct a geospatial analysis. Beginning with the base layer data presented in Multimedia Appendix 1, the data can be used to construct the boundaries for the geospatial analysis. The intent may be to represent data at the county, state, combined statistical area, or other applicable levels. Adjustment data shown in Multimedia Appendix 2 may be added to the analysis to segment the data by a specific category, such as hospital or provider, or based on socioeconomic variables. The outcome data in Multimedia Appendix 3 compose the final layer to a geospatial analysis and are used to generate the measure. These measures are often compared across the base layers and adjustment layers.

Figure 1. Proposed 3-layer model (green blocks) for using geocoded health care data with example data elements (gray blocks).



Discussion

Principal Findings

Despite the significant impact that geotagging can have in health care, there has been limited adoption. This is largely attributed to challenges associated with adopting software and acquiring data sources and a skills gap with implementation and interpretation. This paper summarizes the necessary data layers of geotagging that can be adopted in health care and examines the sources of information that can assist new users in the implementation and use of MedicalGIS.

Spatial analysis methods have expanded over the past half century resulting in an increased number of GIS software tools or updates to existing tools. In many instances software vendors are able to accommodate updates to the original software, while others have opted to provide access to third party applications for extensible use. Additionally, many frameworks offer programming languages, software development kits, and/or extraordinary interfaces for developing novel analytical tools or approaches.

There are numerous software applications that can be adopted for geospatial analysis. Many are commercial products and require a license for use while others are noncommercial products that are either open source or free. These tools are made available either as a web-based or locally installed application. These applications typically include functionalities to support rapid map generation, dynamic interaction and summaries, drill-down capabilities, cross-tabulations, and a widespread suite of graphical options including 2D and 3D analyses. The advantage of commercial products is the degree of support and training that is offered. There are typically ample user guides, white papers, and training offered by licensed software vendors. Noncommercial products typically have limited training opportunities, more verbose and technical user guides, and little technical support but may offer more extensibility and flexibility when working with data.

Some of the free tools, such as the R statistical programming language (R Foundation for Statistical Computing), require significant training on how to use a scripting language to generate a geographic analysis. Despite the steep learning curve, R is free to implement and offers incredible flexibility with analyses. Licensed software such as ArcView GIS (Esri) offers training and support that may assist with an easier transition to implement geotagging analysis. However, the licensed software may be cost prohibitive for some parties. There is also significant growth in the development of business intelligence software such as Tableau (Tableau Software LLC) and QlikView (QlikTech International AB) that offer dashboard views of data and can automate the geotagging process. These tools are particularly useful for organizational monitoring and decision making and require very little training. The choice between a commercial or noncommercial product is usually related to user experience, cost, and time demands of performing an analysis. To learn more about GIS software applications, the Centers for Disease Control and Prevention developed a list of common tools used for geospatial analysis [27].

Limitations

Although positive impacts have been realized with geotagging, significant challenges still remain, data privacy being one. The Health Insurance Portability and Accountability Act of 1996 regulations restrict the use of patient data for research if there is a risk of potentially identifying a person. Therefore, there are limits to the level at which an analysis can be conducted, which is largely restricted to zip code rather than a specific address. However, there are existing limitations in obtaining accurate location data in a small area. As the recent coronavirus outbreak has put lives of millions of individuals worldwide at stake, it is understandable that many governments accepted the risk of data compromise by tracking people's location data in efforts to contain the virus. In China, visitors to public buildings must scan QR codes using their smartphones so the app can monitor people's movements/locations and later notify the users if they have been in close contact with COVID-19 patients. Telecom providers (through call histories) and technology companies (via smartphone apps) in England, Italy, Austria, Israel, Germany, and the United States are collecting or sharing either (1) personal location and use data or (2) anonymous and aggregated data to facilitate mapping and contact tracing [28]. The anonymous aggregated location data were used to map hot spots and determine movement patterns. However, there were arguments that location data does not become anonymous by simply deleting the identifiers; therefore, releasing it may lead to data compromise. Also, the collected data may not be representative as only about 67% of people worldwide have cell phones [28]. The individual location data were used instead to identify people who came into close contact with COVID-19 patients. Nonetheless, GPS/global navigation satellite system technology is practically accurate to within 16 ft. radius under clear sky [29] while the contact tracing would be more accurate if it can correctly obtain positioning data within 6 ft. The accuracy is also reduced near obstacles such as trees, bridges, or buildings [29], and COVID-19 has a higher infection rate inside buildings where human interactions mostly occur and the air is more compact. There is no doubt that geotagging is powerful, but its benefits should overcome the risks, and determining scenarios where it can outweigh the risks is always crucial.

Additionally, many of the external data sources are licensed and require a fee to obtain access. Information integration is often crucial for decision making. Multiple studies have been performed to detect areas or communities most vulnerable to COVID-19. As senior populations were considered at higher risk, Massachusetts Institute of Technology employed geotagged data analysis to determine if senior care facilities were lacking resources by integrating the bed availability data and safety data at assisted living, long-term care facilities, and nursing homes [30]. In an effort to prove that communities experiencing health care disparities are more susceptible to COVID-19, the *Los Angeles Times* created a map reflecting quality of care such as health care provider supply, insurance coverage, and rate of poverty [30]. The *New York Times* also generated a map presenting the number of people and the daily in and out frequency of captives in prisons all over the United States. The map suggested that nearly 20,000 people enter and leave prisons weekly, making the high risk of developing outbreaks from these centralized facilities clear [30].

There is also a skill gap in how to use geographic data and software effectively. Health care organizations typically do not have a GIS expert on staff, requiring them to purchase costly software or hire consultants to perform geographic analyses. While many GIS software applications can be costly and require training to use, there are free and open source software options that can be explored to minimize the cost. The need for training, however, is not often overcome.

Research also indicated that public health studies incorporating geocoded databases should also develop methods to evaluate and verify accuracy [31], as the error rate is often very high. Other issues are the cost of using this technology and hardware and software, training, and maintenance costs. GIS software is traditionally expensive. Although the cost of GIS hardware has decreased over time, it is still considered costly. GIS training is time-consuming and, therefore, expensive. As a result, software maintenance cost is still high.

Another challenge is to identify the required level of analysis for a geotagging project and select the appropriate base layer to support this requirement. For instance, a project that requires a block view of data may be difficult to implement given the challenges with acquiring data at this level. Commonly, the lowest level data is supplied at the 5-digit zip code. At this level, a regional analysis can be performed, but a more granular analysis (ie, block or address) is not possible. Also, frequent updates to 5-digit zip codes require organizations to maintain current sources. When the data intended to be geocoded is not current, this poses a challenge in establishing accurate geocodes.

Data quality is often a concern with geocodes. The quality of geographic data in health care is a continual challenge due to frequent updates to geocodes, lack of uniformity in their adoption, and inconsistencies in data collection parameters [32]. Not only are the geocodes themselves often limited and outdated, but the health care data sources, such as the public sources shown in Multimedia Appendix 3, should be subject to scrutiny. The data are often aggregated from other data sources or suppressed in areas with a low sample size. These limitations to the public data can result in a lack of complete coverage in geographic regions, particularly in rural areas with a lower population. The real value of public health care data comes from combining several sources. Unique identifiers for providers, hospitals, counties, states, and regions can be leveraged to combine data sources. For example, census data for each county can be combined with county health ranking data. A challenge with combining data sources is to resolve many-to-many relationships. For example, a zip code can be related to zero, one, or many core-based statistical areas (CBSA). That is, a single zip code may be mapped to a micropolitan and metropolitan area. Therefore, an analyst can only obtain an approximation of the rural/urban classification for any given zip code. Although such limitations exist, careful data handling can resolve many of these limitations by combining base layer geocodes. For instance, Topologically Integrated Geographic Encoding and Referencing Shapefiles published by the US Census can be used to obtain latitudinal and longitudinal boundaries for CBSAs. These boundaries can be compared with the latitudinal and longitudinal coordinates for a single hospital. If that hospital is found to fall within the boundaries of a CBSA, an analyst can classify the hospital as falling within a rural, micropolitan, or metropolitan area with greater confidence. Finally, there is inconsistent application of geotagging standards and best practices to identify the best data sources and software.

Conclusion

This paper provides clear instructions and explanations on how geotagging can be leveraged in health care. It also offers the necessary initial resources to obtain geocodes, adjustment data, and health-related measures. Despite the motivation to move forward with geotagging in health care, there remain significant challenges around training and acquiring the appropriate software. Data integration, data security, and privacy are crucial and need to be addressed carefully. This technology can be very helpful; however, benefits and potential risks should be carefully identified before implementation. The resources outlined in this paper can support an individual and/or organization in initiating a geotagging health care project.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Base layers for geotagging health care data.

[DOCX File , 17 KB - formative_v4i12e23379_app1.docx]

Multimedia Appendix 2

Adjustment data to limit the geocodes to specific categories (eg, hospitals, providers).

[DOCX File , 15 KB - formative_v4i12e23379_app2.docx]

Multimedia Appendix 3

Public data sources that offer outcome measures for geotagging.

[DOCX File , 18 KB - formative_v4i12e23379_app3.docx]

References

1. GIS and Public Health at CDC. Atlanta: Centers for Disease Control and Prevention URL: https://www.cdc.gov/gis/index.htm [accessed 2020-11-28]
2. Toth J. The good, the bad, and the geotag. 2020 Sep 03. URL: https://www.theoutbound.com/jill-toth/the-good-the-bad-and-the-geotag [accessed 2020-11-28]
3. Glenn EH. The ghost map: the story of London's most terrifying epidemic—and how it changed science, cities, and the modern world. J Am Planning Assoc 2008 Jul 30. [doi: 10.1080/01944360802146329]
4. Grady SC, Wadhwa V. Today and tomorrow: reflections on health and medical geography. Soc Sci Med 2015 May;133:212-215. [doi: 10.1016/j.socscimed.2015.02.030] [Medline: 25754540]

5. Healthy People 2010 final review. Hyattsville: National Center for Health Statistics; 2012. URL: https://www.cdc.gov/nchs/data/hpdata2010/hp2010_final_review.pdf [accessed 2020-11-28]
6. Alarcón-Herrera MT, Martin-Alarcon DA, Gutiérrez M, Reynoso-Cuevas L, Martín-Domínguez A, Olmos-Márquez MA, et al. Co-occurrence, possible origin, and health-risk assessment of arsenic and fluoride in drinking water sources in Mexico: geographical data visualization. Sci Total Environ 2020 Jan 01;698:134168. [doi: 10.1016/j.scitotenv.2019.134168] [Medline: 31505353]
7. Ying H, Wei-Ming W. [Preliminary exploration of mobile terminal data acquisition in malaria control]. Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi 2016 Jun 06;28(3):336-337. [doi: 10.16250/j.32.1374.2016107] [Medline: 29469435]
8. Ide H, Doi S, Atarashi H, Fujita S, Koike S. The distance and chance of lifetime geographical movement of physicians in Japan: an analysis using the age-period-cohort model. Hum Resour Health 2018 Jun 13;16(1):26 [FREE Full text] [Medline: 29895306]
9. Gandolf S. GIS: the power of visualizing hospital and healthcare data. Healthcare Success. URL: https://healthcaresuccess.com/blog/hospital-marketing/gis-power-visualizing-hospital-healthcare-data.html [accessed 2020-11-28]
10. El-Rawy M, Fathi H, Abdalla F. Integration of remote sensing data and in situ measurements to monitor the water quality of the Ismailia Canal, Nile Delta, Egypt. Environ Geochem Health 2020 Jul;42(7):2101-2120. [doi: 10.1007/s10653-019-00466-5] [Medline: 31823180]
11. Leung CK, Braun P, Cuzzocrea A. AI-based sensor information fusion for supporting deep supervised learning. Sensors (Basel) 2019 Mar 18;19(6) [FREE Full text] [doi: 10.3390/s19061345] [Medline: 30889840]
12. Nguyen HT, Duong TQ, Nguyen LD, Vo TQ, Tran NT, Dang PD, et al. Development of a spatial decision support system for real-time flood early warning in the Vu Gia-Thu Bon River Basin, Quang Nam Province, Vietnam. Sensors 2020 Mar 17;20(6):1667. [doi: 10.3390/s20061667]
13. Kajornkasirat S, Muangprathub J, Boonnam N. Online advanced analytical service: profiles for dengue hemorrhagic fever transmission in Southern Thailand. Iran J Public Health 2019 Nov;48(11):1979-1987 [FREE Full text] [Medline: 31970096]
14. Nguyen QC, McCullough M, Meng H, Paul D, Li D, Kath S, et al. Geotagged US Tweets as predictors of county-level health outcomes, 2015-2016. Am J Public Health 2017 Nov;107(11):1776-1782. [doi: 10.2105/AJPH.2017.303993] [Medline: 28933925]
15. Hong B. Exposure density and neighborhood disparities in COVID-19 infection risk: using large-scale geolocation data to understand burdens on vulnerable communities. 2020. URL: https://arxiv.org/abs/2008.01650 [accessed 2020-11-28]
16. Masri S, Jia J, Li C, Zhou G, Lee M, Yan G, et al. Use of Twitter data to improve Zika virus surveillance in the United States during the 2016 epidemic. BMC Public Health 2019 Jun 14;19(1):761 [FREE Full text] [doi: 10.1186/s12889-019-7103-8] [Medline: 31200692]
17. Seah HM, Burney M, Phan M, Shell D, Wu J, Zhou K, et al. Code stroke alert—concept and development of a novel open-source platform to streamline acute stroke management. Front Neurol 2019;10:725 [FREE Full text] [doi: 10.3389/fneur.2019.00725] [Medline: 31417478]
18. Mackey T, Purushothaman V, Li J, Shah N, Nali M, Bardier C, et al. Twitter. JMIR Public Health Surveill 2020 Jun 08;6(2):e19509 [FREE Full text] [doi: 10.2196/19509] [Medline: 32490846]
19. Persico C. The effects of increased pollution on COVID-19 cases and deaths. 2020 Jul 17. URL: https://www.eenews.net/assets/2020/07/17/document_gw_02.pdf [accessed 2020-11-28]
20. Collins S. New app collects the sounds of COVID-19. 2020 Apr 06. URL: https://www.cam.ac.uk/research/news/new-app-collects-the-sounds-of-covid-19 [accessed 2020-11-28]
21. Millar A. Coronavirus screening at airports: the problem with thermal detection. 2020 Apr 06. URL: https://airport.nridigital.com/air_apr20/coronavirus_screening_at_airports_the_problem_with_thermal_detection [accessed 2020-11-28]
22. Sullivan M. Updated: this AI camera detects people who may have COVID-19. Fast Company. URL: https://www.fastcompany.com/90479220/this-ai-camera-detects-people-who-may-have-covid-19 [accessed 2020-11-28]
23. Edwards N. Tech giants like Apple, Experian, And Google fight pandemic with digital tools. Forbes. 2020 May 21. URL: https://www.forbes.com/sites/neiledwards/2020/05/21/apple-experian-and-google-fight-pandemic-with-digital-apps/#5987e3851e6f [accessed 2020-06-16]
24. Goldberg D. A geocoding best practice guide. North American Association of Central Cancer Registries, Inc. 2008 Nov 10. URL: https://www.naaccr.org/wp-content/uploads/2016/11/Geocoding_Best_Practices.pdf [accessed 2020-11-28]
25. Rushton G, Gittler J, Greene B. Geocoding Health Data: The Use of Geographic Codes in Cancer Prevention and Control, Research, and Practice. Boca Raton: CRC Press; 2008.
26. National Research Council. National spatial data infrastructure evolution. In: Toward a Coordinated Spatial Data Infrastructure for the Nation. Washington: The National Academies Press; 1993:12-18.
27. GIS and Public Health at CDC: GIS software & tools. Atlanta: Centers for Disease Control and Prevention URL: https://www.cdc.gov/gis/software-tools.htm [accessed 2020-11-28]
28. Kumar M. Using location data to tackle the coronavirus pandemic. Geoawesomeness. 2020 Mar 30. URL: https://www.geoawesomeness.com/using-location-data-to-tackle-the-coronavirus-pandemic/ [accessed 2020-11-28]

29. Diggelen F, Enge P. The world's first GPS MOOC and worldwide lab using smartphones. 2015 Presented at: Proceedings of the 28th International Technical Meeting of the Satellite Division of The Institute of Navigation; September 14-18, 2015; Tampa. [doi: 10.33012/2017.15124]
30. Choudhary M. COVID-19: maps show vulnerable communities of the US. GeospatialWorld. 2020 May 19. URL: https://www.geospatialworld.net/blogs/covid-19-maps-show-vulnerable-communities-of-the-us/ [accessed 2020-11-28]
31. Krieger N, Waterman P, Lemieux K, Zierler S, Hogan JW. On the wrong side of the tracts? Evaluating the accuracy of geocoding in public health research. Am J Public Health 2001 Jul;91(7):1114-1116. [doi: 10.2105/ajph.91.7.1114] [Medline: 11441740]
32. Fletcher-Lartey SM, Caprarelli G. Application of GIS technology in public health: successes and challenges. Parasitology 2016 Apr;143(4):401-415. [doi: 10.1017/S0031182015001869] [Medline: 26831619]

Abbreviations

AI: artificial information
CDSA: core-based statistical areas
MedicalGIS: medical geographic information systems

Edited by C Lovis; submitted 10.08.20; peer-reviewed by T Heponiemi, YL Cheong; comments to author 28.09.20; revised version received 08.10.20; accepted 07.11.20; published 24.12.20.

Please cite as:
Do Q, Marc D, Plotkin M, Pickering B, Herasevich V
Starter Kit for Geotagging and Geovisualization in Health Care: Resource Paper
JMIR Form Res 2020;4(12):e23379
URL: http://formative.jmir.org/2020/12/e23379/
doi:10.2196/23379
PMID:33361054

Original Paper

Development and Feasibility of an App to Decrease Risk Factors for Type 2 Diabetes in Hispanic Women With Recent Gestational Diabetes (Hola Bebé, Adiós Diabetes): Pilot Pre-Post Study

Ellen W Seely[1], MD; Patricia Flynn Weitzman[2], PhD; Dharma Cortes[2], PhD; Sara Romero Vicente[2], BSc; Sue E Levkoff[2,3], ScD, MSW, SM

[1]Endocrinology, Diabetes and Hypertension Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States
[2]Environment and Health Group, Cambridge, MA, United States
[3]University of South Carolina, Columbia, SC, United States

Corresponding Author:
Ellen W Seely, MD
Endocrinology, Diabetes and Hypertension Division
Brigham and Women's Hospital
Harvard Medical School
221 Longwood Ave
Boston, MA, 02115
United States
Phone: 1 6177325012
Email: eseely@bwh.harvard.edu

Abstract

Background: Hispanic women have increased risk of gestational diabetes mellitus (GDM), which carries an increased risk for future type 2 diabetes, compared to non-Hispanic women. In addition, Hispanic women are less likely to engage in healthy eating and physical activity, which are both risk factors for type 2 diabetes. Supporting patients to engage in healthy lifestyle behaviors through mobile health (mHealth) interventions is increasingly recognized as a viable, underused tool for disease prevention, as they reduce barriers to access frequently experienced in face-to-face interventions. Despite the high percentage of smartphone ownership among Hispanics, mHealth programs to reduce risk factors for type 2 diabetes in Hispanic women with prior GDM are lacking.

Objective: This study aimed to (1) develop a mobile app (¡Hola Bebé, Adiós Diabetes!) to pilot test a culturally tailored, bilingual (Spanish/English) lifestyle program to reduce risk factors for type 2 diabetes in Hispanic women with GDM in the prior 5 years; (2) examine the acceptability and usability of the app; and (3) assess the short-term effectiveness of the app in increasing self-efficacy for both healthy eating and physical activity, and in decreasing weight.

Methods: Social cognitive theory provided the framework for the study. A prototype app was developed based on prior research and cultural tailoring of content. Features included educational audiovisual modules on healthy eating and physical activity; personal action plans; motivational text messages; weight tracking; user-friendly, easy-to-follow recipes; directions on building a balanced plate; and tiered badges to reward achievements. Perceptions of the app's acceptability and usability were explored through four focus groups. Short-term effectiveness of the app was tested in an 8-week single group pilot study.

Results: In total, 11 Hispanic women, receiving care at a federally qualified community health center, aged 18-45 years, and with GDM in the last 5 years, participated in four focus groups to evaluate the app's acceptability and usability. Participants found the following sections most useful: audiovisual modules, badges for completion of activities, weight-tracking graphics, and recipes. Suggested modifications included adjustments in phrasing, graphics, and a tiering system of badges. After app modifications, we conducted usability testing with 4 Hispanic women, with the key result being the suggestion for a "how-to tutorial." To assess short-term effectiveness, 21 Hispanic women with prior GDM participated in the pilot. There was a statistically significant improvement in both self-efficacy for physical activity (P=.003) and self-efficacy for healthy eating (P=.007). Weight decreased but not significantly. Backend process data revealed a high level of user engagement.

Conclusions: These data support the app's acceptability, usability, and short-term effectiveness, suggesting that this mHealth program has the potential to fill the gap in care experienced by Hispanic women with prior GDM following pregnancy. Future studies are needed to determine the effectiveness of an enhanced app in a randomized controlled trial.

Trial Registration: ClinicalTrials.gov NCT04149054; https://clinicaltrials.gov/ct2/show/NCT04149054

(JMIR Form Res 2020;4(12):e19677) doi:10.2196/19677

KEYWORDS

gestational diabetes mellitus; diabetes prevention; Hispanic; Latina; intervention; mobile apps; mHealth

Introduction

Gestational diabetes mellitus (GDM), defined as glucose intolerance diagnosed after the first trimester of pregnancy [1], occurs in 3%-7% of pregnancies in the United States. Due, in part, to the fact that 40% of Hispanic women in the United States of child-bearing age are obese, and 51% experience excessive weight gain during pregnancy [2-5], this group has 1.5 times the risk of GDM compared to non-Hispanic White women [6]. GDM carries an overall increased risk as high as 60% for the development of type 2 diabetes mellitus (T2DM) [7], placing Hispanic women with prior GDM at high risk for future T2DM. Furthermore, obesity, a major risk factor for GDM and the strongest modifiable risk factor for T2DM, is more prevalent among Hispanic than non-Hispanic White women [8,9].

It is widely acknowledged that Hispanic women in the United States experience disparities in health care access and utilization compared to non-Hispanic women [10]. Cultural, social, and economic barriers also lead to disparities in healthy lifestyle behaviors [10]. Hispanic women face sociocultural barriers to healthy eating (eg, cost of healthy food, knowledge about nutritional values of some foods, and family food preferences) [11], as well as structural barriers (eg, food deserts) to obtaining healthy foods [12]. Hispanic women are also less likely to be physically active compared to non-Hispanic White women [13-15]. While some barriers to physical activity are comparable to those for non-Hispanic women (eg, lack of time, lack of childcare, being tired, and having limited self-discipline) [13]. Some additional barriers may be culturally influenced such as being discouraged by family members and friends, as well as environmental factors, such as not having a safe place to exercise [16].

The Diabetes Prevention Program (DPP), which was delivered as a face-to-face intervention, demonstrated that T2DM can be prevented by lifestyle changes focused on healthy eating and physical activity in women with a remote history of self-reported GDM [17]. The scalability of face-to-face DPP-based programs has been a challenge, due to the costs involved in implementing an in-person intervention and the difficulties encountered in attending face-to-face programs [18,19], particularly in postpartum women with recent GDM [20]. A potential approach to overcome barriers to face-to-face implementation is through mobile health (mHealth) technologies that can enable greater patient access.

According to the Pew Research Center, approximately 80% of the Hispanic population owns a smartphone, which is comparable to White and Black populations [21], with Hispanics more likely to use their smartphone to seek health information than their White counterparts [22]. Hispanic people in the United States spend more time using apps than the general population [23]. These data suggest that mobile apps are a viable, underused tool for T2DM prevention in minority populations including Hispanic women with recent GDM.

These findings led us to develop and pilot test a culturally tailored, bilingual (Spanish/English), mobile app–based lifestyle program, *¡Hola Bebé, Adiós Diabetes!* (hereafter referred to as Hola Bebé), to reduce risk factors for T2DM in Hispanic women who have had GDM in the prior 5 years. The years after childbirth are well recognized as representing a “window of opportunity” to improve the future health of women who have had GDM, as demonstrated in our previous work [24,25] and by other studies [26,27]. The goal of the Hola Bebé pilot was to determine the feasibility, acceptability, and short-term effectiveness of an mHealth approach to increasing self-efficacy for healthy eating and increased physical activity, and promoting weight loss, in a population of Hispanic women with recent GDM.

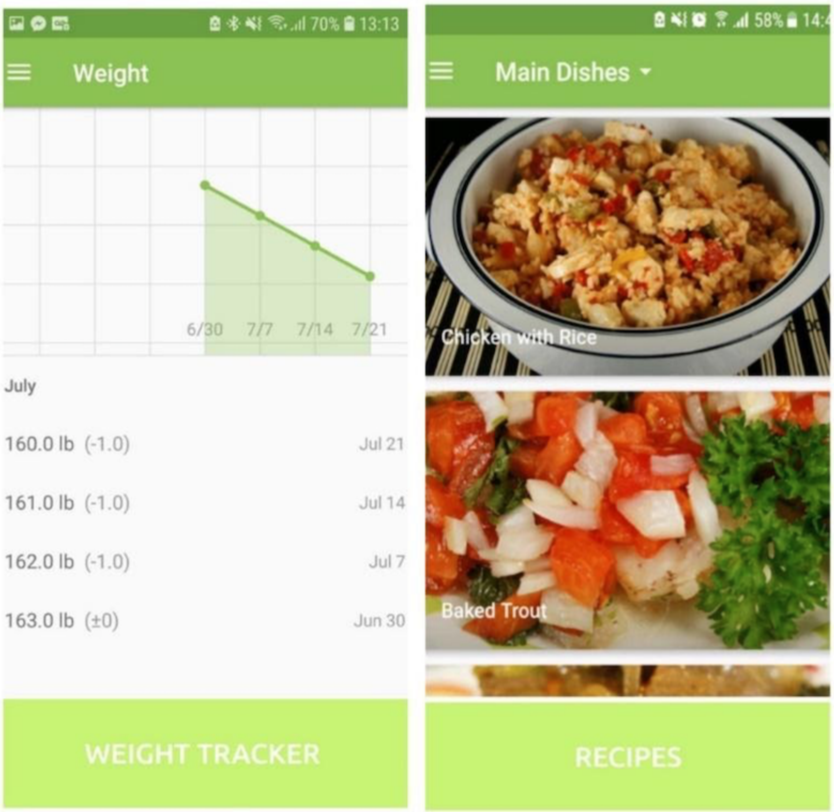
Methods

Overview

Social cognitive theory (SCT) provided the framework for the Hola Bebé intervention. Self-efficacy, the belief in one’s own capabilities to adopt and maintain behavior change [28,29], is a core component of SCT. For the intervention, we developed educational and motivational messages delivered through texts and videos to increase self-efficacy for healthy diet and physical activity. The focus on healthy eating and physical activity for the app was based on the DPP, which demonstrated that lifestyle change targeting healthy eating and increased physical activity led to a decrease in the development of T2DM in individuals at high risk for this condition including women with prior history of GDM [17]. Motivational messages were developed to target self-efficacy, which is associated with initiation and adherence to physical activity and other health-promoting activities [30,31]. Participants chose the times of day and frequency of the text messages. Cultural tailoring involved the development of the app first in Spanish, followed by translation into English with input from Hispanic women with a history of GDM who participated in every stage of app development.

The app included six educational audiovisual modules on healthy eating and physical activity; personal action plans for healthy eating and staying active; motivational and educational text messages; weight tracking; user-friendly, easy-to-follow recipes (Figure 1); directions on how to build a balanced plate; and tiered badges to reward achievements. For the action plans, participants were taught how to identify barriers to individualized goals and ways to overcome the barriers. Healthy eating advice was based on MyPlate [32]. Tiered badges could be earned by the participants with completion of a module, action plan, and/or inputting of weight. The app was developed to meet the 8th-grade literacy level. All content was in plain-language Spanish and English, with Spanish and English audio voiceover.

Figure 1. Screenshots from the ¡Hola Bebé, Adiós Diabetes! app.



July

160.0 lb (-1.0) Jul 21

161.0 lb (-1.0) Jul 14

162.0 lb (-1.0) Jul 7

163.0 lb (±0) Jun 30

Formative Phase

In our formative work, we solicited feedback on the acceptability and usability of the mHealth program through four focus groups. Inclusion criteria included Hispanic women, aged 18-45 years, with prior GDM in the past 5 years, and who received medical care at a federally qualified community health center (CHC), a Level 3 Patient-Centered Medical Home in the Greater Boston area. In total, 11 women participated in the acceptability focus groups. Participants were asked for feedback on the prototype of the app, including what feature(s) they found most useful, preferences for phrasing, wording, graphics, colors, and type of badge tier system. The sections they found most useful were (1) the audiovisual modules, especially those about how to make healthy choices when eating out; (2) the badges for completion of activities; (3) the weight tracking graphic; and (4) the recipes. They also recommended a color scheme from a menu of options and offered suggestions for certain adjustments in phrasing and graphics. Finally, they suggested that the tiering system of badges be based on a system of student achievement, such as “outstanding student” or “honor student.”

After modifications to the prototype app were made based on participants’ input, we conducted usability testing with 4 participants. Participants were given access to the app and asked to perform a number of tasks (eg, click on a tab, complete a module, and open an action plan tab), as well as to explore as they wished. The key result from the usability testing was that participants requested a “how-to tutorial” to make the app easier to use. Some, in fact, offered specific suggestions, such as “click here to add your weight.” Participants also asked for more tabs to better label and access specific sections of the app. In addition, for the action plan completion section, participants suggested instead of only providing a free-text box, there be an additional drop down menu of action plan options with prepopulated action plans to choose from (eg, “I will balance my plate at dinner,” “I will eat fruit and/or vegetable with every meal”). Feedback on the acceptability and usability was incorporated in the app prior to the pilot (Table 1).

Table 1. Features of the ¡Hola Bebé, Adiós Diabetes! app.

Domains	Features
Healthy eating education and tools	• Healthy eating educational audiovisual modules • Healthy recipes • MyPlate education and demonstration • Action plans for healthy eating
Physical activity education and tools	• Physical activity educational audiovisual modules • Exercise videos • Action plans to stay active
Text messaging	• Automated motivational text messages for healthy eating • Automated motivational text messages for physical activity
Self-monitoring	• Weight tracking • Gain badges by completing educational modules and action plans
Social network	• Sharing of experiences and recipes among participants through a community forum • Asking questions of other participants through a community forum
Operating system	• Android

Pilot Trial

A nurse and a medical assistant from the health center identified potential participants who were Hispanic and who had had GDM in the past 5 years from a list generated from the CHC database using the same inclusion criteria as in the formative phase. In addition, participants had to have or be willing to use an Android mobile phone for the study. For women with other smartphones, Android phones were offered on loan for the duration of the study. A member of the study team contacted those women who expressed interest in participating to provide additional information about the study. For those women who agreed to participate, the research assistant scheduled the first study visit. At the first study visit, informed written consent was obtained, and the participant’s weight and height were determined by the research assistant. Weight was measured with the DR400C/Detecto Portable Home Health Care Scale, which was zeroed prior to each weight determination, with the participant wearing light clothes. Participants were asked to complete self-efficacy questionnaires for healthy eating (20 items) and physical activity (12 items) developed by Sallis et al [33], scored on a scale of 1-5 with 5 being the most self-efficacious. These questionnaires have been widely used in research in both Spanish [34] and English. A research assistant helped the participants download (via the Google Play store) and open the app on their phone, and review the “how to use the app” tutorial. Participants were asked to watch one module weekly for the 8 weeks of study duration, complete the corresponding action plan, weigh themselves, and enter their weight into the app. At the end of the 8 weeks, baseline measures were repeated, and structured exit interviews, which focused on what participants liked best and areas for improvement, were performed. Primary outcomes included self-efficacy for healthy eating and self-efficacy for physical activity, with weight as a secondary outcome.

The study was approved by the Pearl Institutional Review Board and the board of the CHC. All participants signed written informed consent.

Statistical Analysis

Descriptive statistics were presented as mean (SD) and frequency (%). For pre-post comparisons from the pilot study, paired *t* tests were conducted with a 5% significance level.

Results

In total, 30 eligible women were identified from the CHC database; 4 women could not be contacted. Of this, 26 women were successfully notified about the study and 21 (88%) consented to participate. Reasons given for not participating included not being interested in participating (n=3), moving out of state (n=1), and not wanting to use a study-provided Android phone (n=1) (Table 2). At baseline, 21 participants were assessed and 18 completed the 8-week study.

Table 2. Baseline characteristics of the study population (N=21).

Characteristic	Value
Age (years), mean (SD)	33 (6.9)
Number of pregnancies, mean (SD)	3 (1.83)
Number of pregnancies with GDM[a], mean (SD)	1 (0.57)
Years post-GDM pregnancy, mean (SD)	2.9 (0.74)
Smartphone operating system, n (%)	
Android	14 (66.6)
iOS	7 (33.3)
Family history of diabetes in first-degree relative, n (%)	14 (66.6)
Ethnicity: Hispanic, n (%)	21 (100)
Primary language: Spanish, n (%)	21 (100)

[a] GDM: gestational diabetes mellitus.

Short-Term Effectiveness

Self-efficacy for healthy eating increased from 4.2 (SD 0.8) to 4.4 (0.7) (*P*=.007). Self-efficacy for physical activity increased from 3.0 (SD 0.6) to 3.4 (SD 0.6) (*P*=.003). The secondary outcome, weight, fell from 163 (SD 36) to 162 (SD 37) lbs (*P*=.16).

Engagement

Backend process data revealed a high level of user engagement. In total, 91% (19/21) of participants viewed audiovisual modules and created action plans. There was also a high level of engagement in earning badges, with 95% of participants (20/21) earning badges by completing a learning module and/or an action plan or weighing. Participants posted tips on the community forum for other participants, such as a family recipe, and asked questions that other participants answered. One woman did not participate in any of the trackable app features.

Exit Interviews

We conducted exit interviews at the conclusion of the pilot study. The following quotes are representative of the participants' experiences of using the app:

> *No cambiaría nada de la aplicación, me gusta todo (I would not change anything from the app, I like everything).*
>
> *Los videos de los módulos me han ayudado a entender la clase de alimentos que son buenos para mi. He comenzado a cambiar los granos por granos integrales y ahora me siento más saludable (The module videos helped me understand what kind of foods were good for me. I started changing my grains for whole wheat grains and now I feel healthier).*
>
> *Cuando voy a comer con mi familia a un restaurante, ya sé que clase de comida puedo ordenar y no sentirme culpable después (Whenever I go to eat with my family to a restaurant, I know which kind of food I can order and not feel guilty afterwards).*

Participants especially liked the personalized action plans, the motivational text messages, the at-home exercise videos, and the recipes. Women commented that they found the "how to use the app" tutorial to be helpful.

Participants had suggestions for incorporation in a future version of the app. They requested more exercise videos including Zumba and with the addition of music, expansion of the recipe section to include more Latin American dishes and vegetarian options, and explanation of portion sizes for each recipe that align with MyPlate. Women also requested videos for recipe preparation. Participants asked for an "ask the expert" option to submit specific exercise and diet questions on the community forum. The 7 iPhone users asked that an app be developed for use on an iPhone.

Discussion

Principal Findings

The *¡Hola Bebé, Adiós Diabetes!* mHealth program was designed to overcome access barriers to T2DM prevention support among Hispanic women with prior GDM. Pilot testing indicated that it was well accepted, usable, and showed preliminary effectiveness at increasing self-efficacy for both physical activity and healthy eating. Weight decreased over the 8-week period but not significantly.

Interventions delivered through apps have great potential to fill the gap experienced by individuals seeking care across a range of conditions. A classification scheme for analyzing apps for preventing and managing disease proposes three dimensions for analysis: health condition (physical versus mental); prevention versus management; and, according to Green and Kreuter's [35] Precede-Proceed Model, predisposing, enabling, and motivating factors [36]. Using this classification scheme, Hola Bebé, addresses a physical condition, that of GDM, for the prevention of type 2 diabetes, and includes factors related to all aspects of the model, such as predisposing (eg, educational audiovisual modules, healthy recipes), enabling (eg, MyPlate demonstrations, action plans, weight tracking, and badges), and motivational (motivational text messages, sharing experiences/recipes, and asking questions through a community forum).

Hola Bebé has the potential to fill the gap in care experienced by women with GDM following pregnancy. Over 86% of women with GDM have no contact with primary care in the first year post delivery, and close to 60% have no contact at 3 years post delivery [37]. This is despite recommendations from the American College of Obstetricians and Gynecologists [38] for referral to primary care and counseling for lifestyle modification in nutrition and exercise for women with a prior pregnancy complicated by GDM. Some have characterized women with prior GDM as falling into a “healthcare chasm” [39]; alternatively, others have referred to this more positively as “a fixable gap in women’s preventative healthcare” [37], which app technology can potentially address. A major advantage of an app-delivered program for Hispanic women is the widespread use of apps by this population [40], which experiences significant disparities in health care [10]. Additional strengths of using an app for behavior change include easy access potential for integration with other apps that commonly come with smartphones (eg, pedometer and music apps); faster speed, as data are stored on the smartphone; and the ease of receiving notifications.

Importantly, Hola Bebé takes advantage of the “window of opportunity” following a complicated pregnancy by bridging the gap in care through lifestyle counseling without dependency on visits to the health center and clinician [37]. This app also overcomes many barriers experienced by women who have young children at home and competing priorities for time; it can be used at home or work or while traveling, day or night, and in small doses whenever users have a few minutes. In addition, this app was culturally and linguistically tailored for Hispanic women and was developed first in Spanish. Finally, the app was designed through an iterative approach incorporating feedback from Hispanic women with recent gestational diabetes at several stages of development.

Limitations

Given the nature of the pilot study, we were limited by a small sample size, lack of a control group, and short study duration. A further limitation was the unavailability of the app for iOS users.

Conclusions

The widespread use of apps among Hispanic women of childbearing age holds promise for this particularly high-risk and underserved population to reduce risk factors for diabetes. This app-delivered program should be tested in a randomized controlled trial and be developed for iOS users.

Acknowledgments

This study was supported by a funding (2R44MD009454-03) from the National Institute on Minority Health and Health Disparities (NIMHD), the National Institutes of Health (NIH).

Conflicts of Interest

None declared.

References

1. American Diabetes Association. Management of Diabetes in Pregnancy: Standards of Medical Care in Diabetes-2020. Diabetes Care 2020 Jan;43(Suppl 1):S183-S192. [doi: 10.2337/dc20-S014] [Medline: 31862757]
2. Headen IE, Davis EM, Mujahid MS, Abrams B. Racial-ethnic differences in pregnancy-related weight. Adv Nutr 2012 Jan;3(1):83-94 [FREE Full text] [doi: 10.3945/an.111.000984] [Medline: 22332106]
3. Chasan-Taber L, Schmidt MD, Pekow P, Sternfeld B, Solomon CG, Markenson G. Predictors of excessive and inadequate gestational weight gain in Hispanic women. Obesity (Silver Spring) 2008 Jul;16(7):1657-1666 [FREE Full text] [doi: 10.1038/oby.2008.256] [Medline: 18451771]
4. Brawarsky P, Stotland NE, Jackson RA, Fuentes-Afflick E, Escobar GJ, Rubashkin N, et al. Pre-pregnancy and pregnancy-related factors and the risk of excessive or inadequate gestational weight gain. Int J Gynaecol Obstet 2005 Nov;91(2):125-131. [doi: 10.1016/j.ijgo.2005.08.008] [Medline: 16202415]
5. Walker LO, Hoke MM, Brown A. Risk factors for excessive or inadequate gestational weight gain among Hispanic women in a U.S.-Mexico border state. J Obstet Gynecol Neonatal Nurs 2009;38(4):418-429. [doi: 10.1111/j.1552-6909.2009.01036.x] [Medline: 19614877]
6. Casagrande SS, Linder B, Cowie CC. Prevalence of gestational diabetes and subsequent Type 2 diabetes among U.S. women. Diabetes Res Clin Pract 2018 Jul;141:200-208. [doi: 10.1016/j.diabres.2018.05.010] [Medline: 29772286]
7. Kim C, Newton KM, Knopp RH. Gestational diabetes and the incidence of type 2 diabetes: a systematic review. Diabetes Care 2002 Oct;25(10):1862-1868. [doi: 10.2337/diacare.25.10.1862] [Medline: 12351492]
8. Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. JAMA 2012 Feb 1;307(5):491-497. [doi: 10.1001/jama.2012.39] [Medline: 22253363]
9. Obesity and Hispanic Americans. HHS.gov. 2020 Mar 26. URL: https://minorityhealth.hhs.gov/omh/browse.aspx?lvl=4 [accessed 2020-10-29]
10. Paz K, Massey KP. Health Disparity among Latina Women: Comparison with Non-Latina Women. Clin Med Insights Womens Health 2016;9(Suppl 1):71-74 [FREE Full text] [doi: 10.4137/CMWH.S38488] [Medline: 27478393]

11. Cortés DE, Millán-Ferro A, Schneider K, Vega RR, Caballero AE. Food purchasing selection among low-income, Spanish-speaking Latinos. Am J Prev Med 2013 Mar;44(3 Suppl 3):S267-S273 [FREE Full text] [doi: 10.1016/j.amepre.2012.11.012] [Medline: 23415192]
12. Morrill KE, Lopez-Pentecost M, Ballesteros G, Pfander JL, Hingle MD, Klimentidis YC, et al. Weight loss interventions for Hispanic women in the USA: a protocol for a systematic review. Syst Rev 2019 Dec 01;8(1):301 [FREE Full text] [doi: 10.1186/s13643-019-1213-3] [Medline: 31787110]
13. Bautista L, Reininger B, Gay JL, Barroso CS, McCormick JB. Perceived barriers to exercise in Hispanic adults by level of activity. J Phys Act Health 2011 Sep;8(7):916-925 [FREE Full text] [doi: 10.1123/jpah.8.7.916] [Medline: 21885882]
14. Neighbors CJ, Marquez DX, Marcus BH. Leisure-time physical activity disparities among Hispanic subgroups in the United States. Am J Public Health 2008 Aug;98(8):1460-1464. [doi: 10.2105/AJPH.2006.096982] [Medline: 18048795]
15. Centers for Disease ControlPrevention (CDC). Prevalence of regular physical activity among adults--United States, 2001 and 2005. MMWR Morb Mortal Wkly Rep 2007 Nov 23;56(46):1209-1212 [FREE Full text] [Medline: 18030281]
16. Larsen BA, Pekmezi D, Marquez B, Benitez TJ, Marcus BH. Physical activity in Latinas: social and environmental influences. Womens Health (Lond Engl) 2013 Mar;9(2):201-210 [FREE Full text] [doi: 10.2217/whe.13.9] [Medline: 23477325]
17. Ratner RE, Christophi CA, Metzger BE, Dabelea D, Bennett PH, Pi-Sunyer X, Diabetes Prevention Program Research Group. Prevention of diabetes in women with a history of gestational diabetes: effects of metformin and lifestyle interventions. J Clin Endocrinol Metab 2008 Dec;93(12):4774-4779 [FREE Full text] [doi: 10.1210/jc.2008-0772] [Medline: 18826999]
18. Hernan WH, Brandle M, Zhang P, Williamson DF, Matulik MJ, Ratner RE, Diabetes Prevention Program Research Group. Costs associated with the primary prevention of type 2 diabetes mellitus in the diabetes prevention program. Diabetes Care 2003 Jan;26(1):36-47 [FREE Full text] [doi: 10.2337/diacare.26.1.36] [Medline: 12502656]
19. Ackermann RT, Finch EA, Brizendine E, Zhou H, Marrero DG. Translating the Diabetes Prevention Program into the community. The DEPLOY Pilot Study. Am J Prev Med 2008 Oct;35(4):357-363 [FREE Full text] [doi: 10.1016/j.amepre.2008.06.035] [Medline: 18779029]
20. Nicklas JM, Zera CA, Seely EW, Abdul-Rahim ZS, Rudloff ND, Levkoff SE. Identifying postpartum intervention approaches to prevent type 2 diabetes in women with a history of gestational diabetes. BMC Pregnancy Childbirth 2011;11:23 [FREE Full text] [doi: 10.1186/1471-2393-11-23] [Medline: 21435246]
21. Perrin A, Turner E. Smartphones help blacks, Hispanics bridge some-but not all- digital gaps with whites. Pew Research Center. 2019. URL: https://www.pewresearch.org/fact-tank/2019/08/20/smartphones-help-blacks-hispanics-bridge-some-but-not-all-digital-gaps-with-whites/ [accessed 2020-04-22]
22. Anderson M. Racial and ethnic differences in how people use mobile technology. Pew Research Center. 2015. URL: https://www.pewresearch.org/fact-tank/2015/04/30/racial-and-ethnic-differences-in-how-people-use-mobile-technology [accessed 2020-04-22]
23. Descubrimiento Digital: The Online Lives of Latinx Consumers (Rep). The Nielsen Company. 2018. URL: https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/the-online-lives-latinx-consumers.pdf [accessed 2020-10-28]
24. Bentley-Lewis R, Levkoff S, Stuebe A, Seely EW. Gestational diabetes mellitus: postpartum opportunities for the diagnosis and prevention of type 2 diabetes mellitus. Nat Clin Pract Endocrinol Metab 2008 Oct;4(10):552-558 [FREE Full text] [doi: 10.1038/ncpendmet0965] [Medline: 18779843]
25. Rich-Edwards JW, McElrath TF, Karumanchi SA, Seely EW. Breathing life into the lifecourse approach: pregnancy history and cardiovascular disease in women. Hypertension 2010 Sep;56(3):331-334 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.110.156810] [Medline: 20679178]
26. Beckles G, Thompson-Reid P. Diabetes & Women's Health across the Life Stages: A Public Health Perspective. Atlanta, GA: US Dept of Health and Human Services, Centers for Disease Control and Prevention; 2001.
27. Phelan S. Pregnancy: a "teachable moment" for weight control and obesity prevention. Am J Obstet Gynecol 2010 Feb;202(2):135.e1-135.e8 [FREE Full text] [doi: 10.1016/j.ajog.2009.06.008] [Medline: 19683692]
28. Bandura A. Self-efficacy: The Exercise of Control. New York, NY: WH Freeman/Times Books/Henry Holt & Co; 1997.
29. Bandura A. The Explanatory and Predictive Scope of Self-Efficacy Theory. Journal of Social and Clinical Psychology 1986 Sep;4(3):359-373. [doi: 10.1521/jscp.1986.4.3.359]
30. Desharnais R, Bouillon J, Godin G. Self-efficacy and outcome expectations as determinants of exercise adherence. Psychol Rep 2016 Sep 01;59(3):1155-1159. [doi: 10.2466/pr0.1986.59.3.1155]
31. Nansel TR, Haynie DL, Lipsky LM, Wang J, Mehta SN, Laffel LM. Relationships among parent and youth healthful eating attitudes and youth dietary intake in a cross-sectional study of youth with type 1 diabetes. Int J Behav Nutr Phys Act 2013;10(1):125. [doi: 10.1186/1479-5868-10-125]
32. What is MyPlate? Chose MyPlate, US Department of Agriculture. 2020. URL: https://www.choosemyplate.gov/eathealthy/WhatIsMyPlate [accessed 2020-12-18]
33. Sallis JF, Pinski RB, Grossman RM, Patterson TL, Nader PR. The development of self-efficacy scales for healthrelated diet and exercise behaviors. Health Educ Res 1988;3(3):283-292. [doi: 10.1093/her/3.3.283]
34. Sallis JF. Self-Efficacy for Diet and Exercise. DrJimSallis.org. URL: https://drjimsallis.org/measure_selfefficacy.html [accessed 2020-10-28]

35. Green LW, Kreuter MW. Health promotion planning: An educational and environmental approach. New York: McGraw-Hill Humanities; 2004.
36. Wang A, An N, Lu X, Chen H, Li C, Levkoff S. A classification scheme for analyzing mobile apps used to prevent and manage disease in late life. JMIR Mhealth Uhealth 2014 Feb 17;2(1):e6 [FREE Full text] [doi: 10.2196/mhealth.2877] [Medline: 25098687]
37. Bernstein JA, Quinn E, Ameli O, Craig M, Heeren T, Lee-Parritz A, et al. Follow-up after gestational diabetes: a fixable gap in women's preventive healthcare. BMJ Open Diabetes Res Care 2017;5(1):e000445 [FREE Full text] [doi: 10.1136/bmjdrc-2017-000445] [Medline: 28948028]
38. Committee on Practice Bulletins—Obstetrics. Practice Bulletin No. 180: Gestational Diabetes Mellitus. Obstet Gynecol 2017 Jul;130(1):e17-e37. [doi: 10.1097/AOG.0000000000002159] [Medline: 28644336]
39. Wilkinson SA, Lim SS, Upham S, Pennington A, O'Reilly SL, Asproloupos D, et al. Who's responsible for the care of women during and after a pregnancy affected by gestational diabetes? Med J Aust 2014 Aug 04;201(3 Suppl):S78-S81. [doi: 10.5694/mja14.00251] [Medline: 25047889]
40. La Oportunidad Latinx: Cultural currency and the consumer journey. The Nielson Company. 2019. URL: https://www.nielsen.com/us/en/insights/report/2019/la-oportunidad-latinx/ [accessed 2020-04-22]

Abbreviations

CHC: community health center
DPP: Diabetes Prevention Program
GDM: gestational diabetes mellitus
mHealth: mobile health
SCT: social cognitive theory
T2DM: type 2 diabetes mellitus

Edited by G Eysenbach; submitted 30.04.20; peer-reviewed by E Jones, A Wang, K Sincalir; comments to author 11.07.20; revised version received 30.10.20; accepted 03.11.20; published 31.12.20.

Please cite as:
Seely EW, Weitzman PF, Cortes D, Romero Vicente S, Levkoff SE
Development and Feasibility of an App to Decrease Risk Factors for Type 2 Diabetes in Hispanic Women With Recent Gestational Diabetes (Hola Bebé, Adiós Diabetes): Pilot Pre-Post Study
JMIR Form Res 2020;4(12):e19677
URL: http://formative.jmir.org/2020/12/e19677/
doi:10.2196/19677
PMID:33382039

Original Paper

Citizen Responses to Government Restrictions in Switzerland During the COVID-19 Pandemic: Cross-Sectional Survey

Kevin Selby[1], MD, MAS; Marie-Anne Durand[2], PhD; Alexandre Gouveia[1], MD; Francesca Bosisio[3], PhD; Gaia Barazzetti[3], PhD; Maxime Hostettler[1], BSc, MSc; Valérie D'Acremont[1], MD, PhD; Alain Kaufmann[3], MSc; Christian von Plessen[1,4], MD, PhD

[1]Center for Primary Care and Public Health (Unisanté), University of Lausanne, Lausanne, Switzerland
[2]University Toulouse III Paul Sabatier, Toulouse, France
[3]The ColLaboratory - Participatory, Collaboratory and Action-Research Unit, University of Lausanne, Lausanne, Switzerland
[4]Institute for Clinical Research, University of Southern Denmark, Odense, Denmark

Corresponding Author:
Kevin Selby, MD, MAS
Center for Primary Care and Public Health (Unisanté)
University of Lausanne
Rue de Bugnon 44
Lausanne, 1011
Switzerland
Phone: 41 0798940957
Email: kevin.selby@unisante.ch

Abstract

Background: The success of government-recommended mitigation measures during the COVID-19 pandemic depends largely on information uptake and implementation by individual citizens.

Objective: Our aim was to assess citizens’ knowledge and perceptions about COVID-19 recommendations in the Canton of Vaud, Switzerland.

Methods: A cross-sectional electronic survey with open and closed questions was disseminated by community-based partners prior to the relaxation of government restrictions. Outcomes included citizen knowledge (9-question measure) and worry about the virus, perception of government measures, and recommendations for improvements. Comparisons used linear regression, controlling for age, sex, education, and health literacy. Free-text answers were analyzed thematically.

Results: Of 807 people who accessed the survey, 684 (85%) completed all questions and 479 (60%) gave free-text recommendations. Overall, 75% were female, the mean age was 48 years, and 93% had high health literacy. Knowledge scores were high, with a median score of 8 out of 9. Mean levels of worry about the COVID-19 pandemic were higher in women than men (55/100 versus 44/100, $P<.001$), and in respondents with lower health literacy (57/100 versus 52/100, $P=.03$). Self-reported adherence to recommendations was high (85%) and increased with age and worry (both $P<.001$). Respondents rated their own adherence higher than others (85% versus 61%, $P<.001$). Moreover, 34% of respondents reported having self-quarantined; this rose to 52% for those aged ≥75 years. Those who had self-quarantined reported higher levels of fear. Nearly half (49%) of respondents felt the government response had been adequate, though younger age and higher levels of worry were associated with considering the response to be insufficient (both $P<.001$). Analysis of open-text answers revealed 4 major themes: access to and use of masks, gloves, and hand sanitizer; government messaging; lockdown and lockdown exit plan communication; and testing for COVID-19.

Conclusions: Knowledge, adherence, and satisfaction regarding government recommendations and response were high in this sample, but many desired greater access to personal protective equipment. Those with lower health literacy and those who have been in self-isolation reported greater concerns about the pandemic.

(JMIR Form Res 2020;4(12):e20871) doi:10.2196/20871

KEYWORDS

COVID-19; coronavirus; Switzerland; mitigation strategies; citizen knowledge

Introduction

The global threat posed by the COVID-19 pandemic has led governments to impose unprecedented restrictions on personal movement and conduct [1]. On February 25, 2020, Switzerland had its first confirmed case; just 4 weeks later, there were 10,772 cases and 165 deaths [2]. On March 13 and 16, federal public health officials announced restrictions on gatherings and recommended multiple behavior changes intended to slow the spread of the virus and protect those most vulnerable to severe outcomes of infection [3,4]. Further, testing for COVID-19 was limited to high-risk populations. These restrictions depended largely on individuals self-isolating in case of symptoms and self-quarantining in case of contact with someone with symptoms. The government relied on cooperation from the public to adapt policies to their own situation. Little is known about how individuals understood, perceived, and implemented the recommendations in their personal life. Adherence to recommendations is typically influenced by the individual's understanding of the information, perception of the virus, related threat, and perceived impact on their lives. Rapid online surveys of people's perceptions during the COVID-19 outbreak suggested that participants generally had a good knowledge of the primary disease transmission modes and common symptoms [5]. However, about one-third of the sample had misconceptions about ways to prevent an infection and recommended care-seeking behaviors. In India, a cross-sectional online survey of 2459 participants also revealed good knowledge, which was higher among respondents aged >40 years, people with higher educational attainment, those living in urban areas, and those in health care professions [6]. As far as could be determined in April 2020, no other survey of adult knowledge and perceptions of the COVID-19 pandemic had been conducted in Switzerland.

In this context, we aimed to assess citizens' knowledge and perceptions about COVID-19 recommendations in the Canton of Vaud, Switzerland. Further, we collected information about individual behaviors and perceived adherence to population-level recommendations and suggestions for improvement.

Methods

We conducted a cross-sectional electronic survey using REDCap (Vanderbilt University) between April 17 and 28, 2020, prior to the easing of restrictions in Switzerland. The survey was distributed on social media by multiple community partners, including a local consumer organization, health promotion organizations, and university community members. Ethics approval was not required as all data collected were anonymous [7]. The survey was supported by the health department of the Canton of Vaud, a French-speaking canton in Switzerland.

The electronic survey consisted of 20 questions about participants' knowledge and perception, as well as individual behaviors and perceived adherence, regarding population-level recommendations about COVID-19 (Multimedia Appendix 1). The survey also assessed information sources and worry. We used an open-text response to collect suggestions for improvement regarding additional measures that the government could implement to limit the spread of the virus (see Table 1 for an overview of the survey structure). The survey was developed in English, translated to French, and user-tested by 6 non–medical professionals. Demographic data collected included age, sex, number of persons in household, Canton of residence, highest level of education, a validated health literacy item [8], and whether they had been tested for COVID-19. A knowledge score was developed with 9 true-false items about current government recommendations. We used lists of common changes due to the restrictions, enabling factors and barriers to implementation, and protection measures. Visual analogue scores were used to measure worry about the pandemic, self-reported adherence, and perception of government measures.

Table 1. Overview of survey structure.

Question number	Section
1-6	Sociodemographic questions
7-8	COVID-19 testing and symptoms
9	Information sources about COVID-19
10	Worry about the new coronavirus
11	Knowledge about government restrictions
12-18	Individual behaviors and perceived adherence to recommendations
19	Perception of the government measures to limit the spread of the new coronavirus
20	Additional suggestions to limit the spread of the new coronavirus (open-text answer)

We used descriptive statistics for participant characteristics, including means for continuous variables, except for perceptions of government measures, which were dichotomized because of a nonnormal distribution. Scores below 45 were classified as considering the government response inadequate. We used multivariable regression to explore associations between participant characteristics and level of worry and self-reported adherence; both models included participant age (continuous), gender (male or female), education (university level or not), health literacy (high or low), and knowledge of current government recommendations (number of answers correct). We used logistic regression to explore associations between considering government measures inadequate and participant age, gender, education, health literacy, knowledge of current government recommendations, and worry. The estimation command "predict" was used in STATA (StataCorp) to calculate absolute differences between groups, using the models controlling for other factors. Analyses were limited to completed surveys. A *P* value <.05 was considered statistically significant. All quantitative analyses were performed with STATA and free-text answers were analyzed thematically by one author (MD) using MAXQDA (VERBI Software).

Results

Overview

Of 807 people who accessed the survey, 684 (85%) completed it and 479 (60%) gave free-text recommendations for improvement. Respondents were 75% female, with a mean age of 48 years and 93% self-reporting high health literacy (Table 2). The primary sources of information about COVID-19 were television and radio (73%), government sources (72%), and newspapers (63%). Knowledge scores were high, with a median score of 8/9. The most common incorrect answer was agreeing that people must stop using all public transport, which was not a federal recommendation (32% incorrect). Overall, 34% of respondents reported having self-quarantined, a proportion that increased with age (27% of those aged >30 years versus 52% for those aged ≥75 years, *P*<.001).

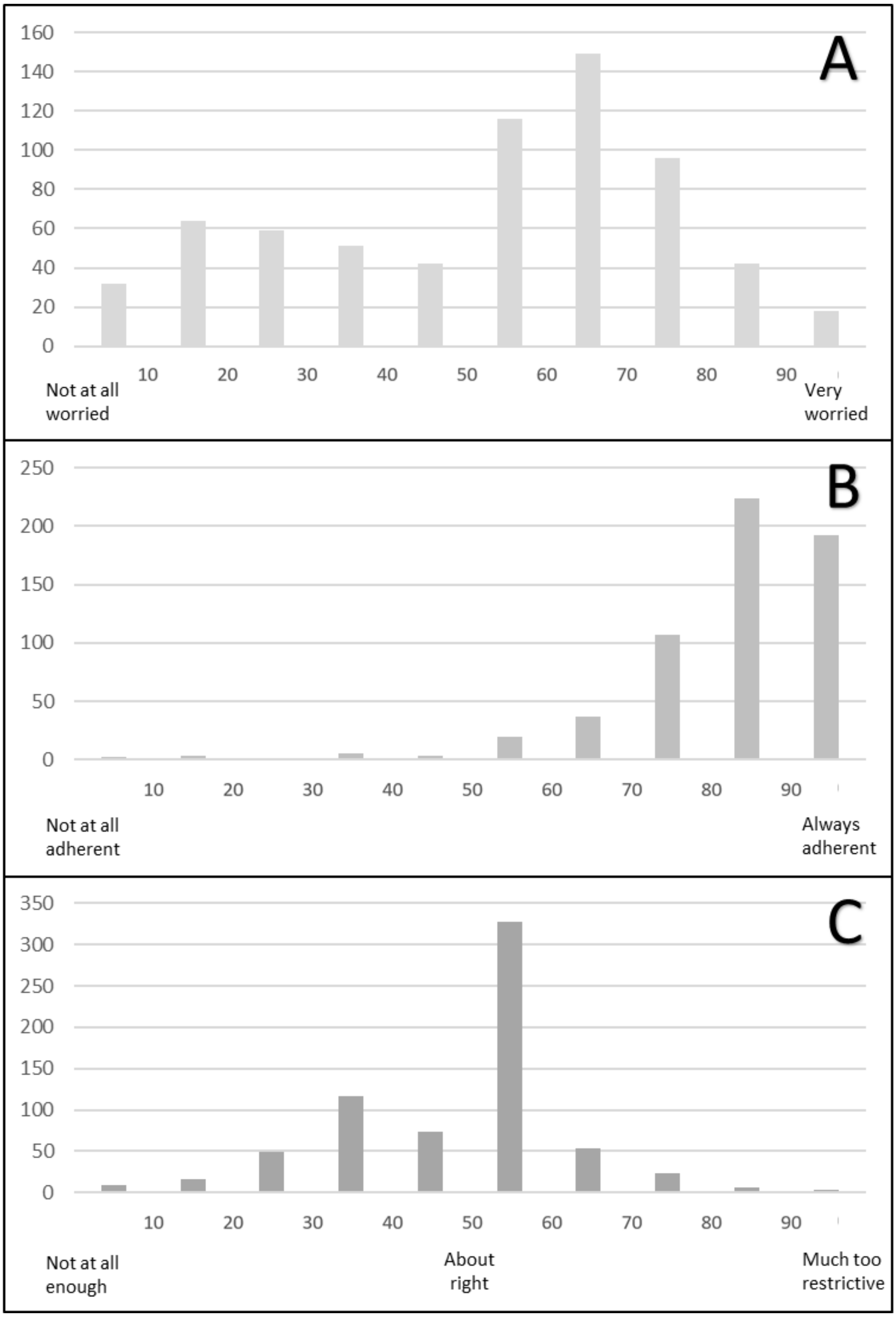
Table 2. Demographic characteristics of respondents who completed the questionnaire (N=684).

Characteristic	Values, n (%)
Age (years)	
18-39	223 (33)
40-64	351 (51)
≥65	110 (16)
Sex	
Female	165 (24)
Male	518 (76)
Canton of residence (31 missing)	
Vaud	581 (88)
Other	77 (12)
Highest education level (5 missing)	
High school or apprenticeship	239 (35)
University or professional school	440 (65)
Health literacy[a] (1 missing)	
Low health literacy	48 (7)
High health literacy	635 (93)
Tested for COVID-19 by reverse transcription polymerase chain reaction test (1 missing)	
Yes	44 (6)
No	639 (94)
Members in household (4 missing)	
≤2	365 (53)
≥3	315 (47)
Knowledge score on 9 true-false questions	
≤7	222 (33)
8 or 9	461 (67)

[a] Dichotomized, with "Often" and "Always" as high health literacy, and "Never," "Rarely," and "Sometimes" as low health literacy.

After controlling for participant age, level of education, and knowledge of current government recommendations, mean levels of worry were higher in women than in men (55/100 versus 44/100, *P*<.001), in respondents with lower health literacy (57/100 versus 52/100, *P*=.03), and among those who had been in self-isolation or self-quarantine (59/100 versus 49/100, *P*<.001). The distribution of responses is shown in Figure 1. Nearly half (49%) of respondents considered the government response to have been adequate, responding in the middle of the 0-100 scale (Figure 1). Younger age and higher worry were associated with considering the response to be insufficient (both *P*<.001). Self-reported adherence to government recommendations was high (85%) and increased with age and level of worry (both *P*<.001; Figure 1). Self-reported adherence was higher than the perceived adherence of other people (85% versus 61%, *P*<.001). Government (82%) and other online sources of information (38%), help from family and friends (32%), and employers (25%) were considered to be useful for following the recommendations. Basic needs (57%), family obligations like caring for the elderly (20%), and difficulties with changing habits (10%) were barriers. The majority of respondents had intervened to encourage others to follow recommendations (77%).

Figure 1. Distribution of participant responses to three questions using a visual analogue scale (0-100). (A) Overall worry about COVID-19. (B) Self-reported adherence to federal recommendations. (C) Adequacy of government recommendations to limit the spread of COVID-19.



Thematic Analysis

Thematic analysis of open-text recommendations for improvements to the government response revealed 4 major themes, presented below alongside participant remarks.

Access and Use of Masks, Gloves, and Hand Sanitizer

Many felt masks should become compulsory outside the home, particularly in crowded areas (eg, public transport, at work), where social distancing is difficult.

> *We should implement the use of masks (so with distribution to the population) for each time you leave the house in order to have a visual reminder of the need to follow recommendations, while also decreasing contamination.*

Many advocated for free distribution to the entire population (including gloves and sanitizer if possible), or access at affordable prices.

> *Ensure that masks are available for everyone, even if their usefulness remains controversial.*

Government Messaging

Although many participants considered the government messages and response adequate and expressed gratitude, some asked for more transparency and precise messages, centralized information, and specific information for groups at risk.

> *Give clear and comprehensible recommendations. Avoid jargon.*

Lockdown and Lockdown Exit Plans

A considerable number of people favored slower easing of restrictions, or even an extension of lockdown with stricter measures, particularly for vulnerable people.

> *Do not rush deconfinement, it’s too early and I am at risk, it scares me…*

> *Enforce complete confinement during at least two weeks.*

Further, some called for tighter policing.

> *Be even more firm (fines? More insistent?) about hand washing and especially social distancing.*

> *1. More surveillance and fines that are a stronger deterrent, 2. Systematic surveillance of rules in businesses.*

A number of people recommended postponing the opening of schools and day cares, and making the return to school optional.

> *Wait longer before reopening schools, because with the recent reopening of stores and other places, there will be a resurgence in cases and if children return to school, it’s grandparents who will care for them after school because parents will be back at work.*

Testing

Large-scale deployment of COVID-19 testing was a recurring theme.

> *Begin testing everyone with symptoms and not just those who are at risk or are medical personnel.*

Further, the use of serological tests, the tracing of chains of transmission, and the isolation of positive cases were also mentioned repeatedly.

> *Impose electronic surveillance with tracking (smartphones) for all citizens during the pandemic.*

Discussion

In this cross-sectional, electronic survey distributed by social media, knowledge and self-reported adherence were high in this sample from Switzerland. Despite high overall satisfaction, free-text answers revealed a desire for greater availability of protective equipment, clearer and more coherent communication by the government, and large-scale testing for the virus. People in isolation and those with lower health literacy reported greater concerns about the pandemic, associations that persisted after controlling for demographic features.

While several larger survey studies on the same subject matter are underway in Europe [9], few studies with quantitative data about individual responses are available for comparison. A Chinese survey also showed high knowledge scores (10.8/12) and high confidence that COVID-19 would be controlled (91%) [10]. In an online survey by the Swiss public broadcasting corporation, 48% of respondents reported major concerns about COVID-19 and 49% felt the government response was too slow [11]. Satisfaction with the government response was higher in a second survey on April 8, 2020 [12]. Support for government restrictions was higher in Italian- and French-speaking areas, which had a higher density of cases, than in German-speaking areas of Switzerland. We did not find data regarding barriers and enablers of adherence to COVID-19 recommendations or data to be stratified by health literacy. Several other surveys conducted in the United States [5], United Kingdom [5], India [6], Egypt [13], and Malaysia [14] have shown adequate overall knowledge about COVID-19 among respondents. However, specific misconceptions and variations in knowledge according to age, income level, educational attainment, race, and ethnicity have been identified. A US national survey of 5198 people suggested that the largest differences in COVID-19–related knowledge and behaviors were associated with race/ethnicity, sex, and age [15]. African American participants, men, and people younger than 55 years were less likely to know how the disease spread, were less likely to know the symptoms, and left the home more often [15].

In our sample, self-reported adherence was high. A study of citizens’ adherence to COVID-19 mitigation recommendations in 3 countries (United States, Kuwait, and South Korea) suggested that in all 3 countries, government response efforts and business reopening agreements, as well as the intensity of information source use, social media use, and knowledge about COVID-19, all influenced either self-adherence or perception of the adherence of others [16].

The finding that those who had been in self-isolation and had lower health literacy were more worried about COVID-19 has important implications. Social isolation among older adults contributes to not only greater risk of depression and anxiety, but also cardiovascular, autoimmune, and neurological disease

[17]. Thus, great care must be taken to help this group. Further, there is well-founded concern that marginalized groups may be disproportionately affected by government measures and the disease itself [18]. Information about COVID-19 in the media is often sensationalized and contradictory, making it challenging for those with lower health literacy to effectively assess risk. As far as could be determined, few studies have so far explored health literacy in the context of COVID-19–related knowledge and perceptions. One study of 630 US adults with chronic conditions found that, contrary to our sample, people who had lower health literacy were less likely to be worried about COVID-19 and to believe they would become infected [19]. Those with lower health literacy also reported greater confidence in the government response. A cross-sectional study of COVID-19–related health literacy was conducted among 1037 German people aged 16 and older [20]. The results highlighted that while the overall health literacy level was high, nearly half of respondents (47.8%) reported difficulties judging whether or not to trust media information on COVID-19. As expected, confusion about COVID-19–related information was significantly higher among respondents with lower health literacy. Accessible information, that adheres to widely accepted plain language principles, and is adapted to those who have lower health literacy and might feel most vulnerable, seems warranted.

Strengths of this survey included rapid data collection prior to easing restrictions, giving a first assessment of citizen responses to restrictions in Switzerland. The primary limitations were those inherent in cross-sectional, open, online survey samples. Moreover, self-reporting may be subject to desirability biases. We could not determine temporality between factors such as fear of the virus and self-isolation. Our sample had a higher proportion of women (76% versus 51%) and university graduates (65% versus 45%) than the Vaud population. Although rates of survey completion were high, we do not know our exact response rate. Finally, because of the novel subject matter, we could not use previously validated survey questions.

In conclusion, knowledge and self-reported adherence were high in an online sample from the Canton of Vaud. Levels of worry about COVID-19 were generally high, particularly among those in isolation and with lower health literacy. Future research should aim to better understand the concerns and needs of these groups and envision target support measures. Special effort will be needed to limit not only the direct health effects of COVID-19, but also psychological distress created by government restrictions as the pandemic continues to evolve.

Acknowledgments

We would like to thank all of our community partners for their responsiveness and support, especially Yannis Papadaniel from the Fédération Romande des Consommateurs.

Conflicts of Interest

None declared.

Multimedia Appendix 1
English-language version of questionnaire.
[PDF File (Adobe PDF File), 178 KB - formative_v4i12e20871_app1.pdf]

References

1. Fauci AS, Lane HC, Redfield RR. Covid-19 — Navigating the Uncharted. N Engl J Med 2020 Mar 26;382(13):1268-1269. [doi: 10.1056/nejme2002387]
2. Probst D. COVID-19 Info Switzerland. URL: https://www.corona-data.ch/ [accessed 2020-05-01]
3. Le Conseil fédéral renforce les mesures contre le coronavirus pour protéger la santé de la population et soutient les secteurs touchés. Confédération suisse. URL: https://www.admin.ch/gov/fr/accueil/documentation/communiques.msg-id-78437.html#:~:text=Berne%2C%2013.03.2020%20%2D%20Afin,de%20plus%20de%20100%20personnes [accessed 2020-11-25]
4. Coronavirus: Federal Council declares "extraordinary situation" and introduces more stringent measures. The Federal Council. URL: https://www.admin.ch/gov/en/start/documentation/media-releases.msg-id-78454.html#:~:text=Bern%2C%2016.03.,terms%20of%20the%20Epidemics%20Act [accessed 2020-11-25]
5. Geldsetzer P. Use of Rapid Online Surveys to Assess People's Perceptions During Infectious Disease Outbreaks: A Cross-sectional Survey on COVID-19. J Med Internet Res 2020 Apr 02;22(4):e18790 [FREE Full text] [doi: 10.2196/18790] [Medline: 32240094]
6. Narayana G, Pradeepkumar B, Ramaiah JD, Jayasree T, Yadav DL, Kumar BK. Knowledge, perception, and practices towards COVID-19 pandemic among general public of India: A cross-sectional online survey. Curr Med Res Pract 2020 Jul;10(4):153-159 [FREE Full text] [doi: 10.1016/j.cmrp.2020.07.013] [Medline: 32839725]
7. Commission cantonal d'éthique de la recherche sur l'être humain. Canton de Vaud. URL: http://cer-vd.ch/soumission/premiers-pas.html [accessed 2020-05-01]

8. Chew LD, Griffin JM, Partin MR, Noorbaloochi S, Grill JP, Snyder A, et al. Validation of screening questions for limited health literacy in a large VA outpatient population. J Gen Intern Med 2008 May 12;23(5):561-566 [FREE Full text] [doi: 10.1007/s11606-008-0520-5] [Medline: 18335281]
9. COVID-19 and Confinement: A Large-Scale French Survey of Social Challenges and Health. INSERM Press Office. 2020 Apr 10. URL: https://presse.inserm.fr/en/covid-19-and-confinement-a-large-scale-french-survey-of-social-challenges-and-health/39099/ [accessed 2020-11-25]
10. Zhong B, Luo W, Li H, Zhang Q, Liu X, Li W, et al. Knowledge, attitudes, and practices towards COVID-19 among Chinese residents during the rapid rise period of the COVID-19 outbreak: a quick online cross-sectional survey. Int J Biol Sci 2020;16(10):1745-1752 [FREE Full text] [doi: 10.7150/ijbs.45221] [Medline: 32226294]
11. Kissling V. Santé, quotidien, autorités, ce que pensent les Suisses du coronavirus. Radio Télévision Suisse. 2020 Mar 25. URL: https://www.rts.ch/info/suisse/11191960-sante-quotidien-autorites-ce-que-pensent-les-suisses-du-coronavirus.html [accessed 2020-11-25]
12. Didier Kottelat MH. Pas de majorité pour un relâchement du semi-confinement en Suisse, selon un sondage. Radio télévision Suisse. URL: https://www.rts.ch/info/suisse/11229396-pas-de-majorite-pour-un-relachement-du-semiconfinement-en-suisse-selon-un-sondage.html [accessed 2020-11-25]
13. Abdelhafiz AS, Mohammed Z, Ibrahim ME, Ziady HH, Alorabi M, Ayyad M, et al. Knowledge, Perceptions, and Attitude of Egyptians Towards the Novel Coronavirus Disease (COVID-19). J Community Health 2020 Oct 21;45(5):881-890 [FREE Full text] [doi: 10.1007/s10900-020-00827-7] [Medline: 32318986]
14. Azlan AA, Hamzah MR, Sern TJ, Ayub SH, Mohamad E. Public knowledge, attitudes and practices towards COVID-19: A cross-sectional study in Malaysia. PLoS One 2020 May 21;15(5):e0233668 [FREE Full text] [doi: 10.1371/journal.pone.0233668] [Medline: 32437434]
15. Alsan M, Stantcheva S, Yang D, Cutler D. Disparities in Coronavirus 2019 Reported Incidence, Knowledge, and Behavior Among US Adults. JAMA Netw Open 2020 Jun 01;3(6):e2012403 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.12403] [Medline: 32556260]
16. Al-Hasan A, Yim D, Khuntia J. Citizens' Adherence to COVID-19 Mitigation Recommendations by the Government: A 3-Country Comparative Evaluation Using Web-Based Cross-Sectional Survey Data. J Med Internet Res 2020 Aug 11;22(8):e20634 [FREE Full text] [doi: 10.2196/20634] [Medline: 32716896]
17. Armitage R, Nellums LB. COVID-19 and the consequences of isolating the elderly. Lancet Public Health 2020 May;5(5):e256 [FREE Full text] [doi: 10.1016/S2468-2667(20)30061-X] [Medline: 32199471]
18. Bodenmann P, Pahud-Vermeulen B, Bouche L, Sanchis Zozaya J, Bauermeister M, Berzig A. [Left behind populations, COVID-19 and risks of health inequities : a guide of the local social-health network (Vaud, Switzerland)]. Rev Med Suisse 2020 Apr 29;16(N° 691-2):859-862. [Medline: 32348054]
19. Wolf MS, Serper M, Opsasnick L, O'Conor RM, Curtis L, Benavente JY, et al. Awareness, Attitudes, and Actions Related to COVID-19 Among Adults With Chronic Conditions at the Onset of the U.S. Outbreak. Annals of Internal Medicine 2020 Jul 21;173(2):100-109. [doi: 10.7326/m20-1239]
20. Okan O, Bollweg TM, Berens E, Hurrelmann K, Bauer U, Schaeffer D. Coronavirus-Related Health Literacy: A Cross-Sectional Study in Adults during the COVID-19 Infodemic in Germany. Int J Environ Res Public Health 2020 Jul 30;17(15):1 [FREE Full text] [doi: 10.3390/ijerph17155503] [Medline: 32751484]

Edited by G Eysenbach; submitted 31.05.20; peer-reviewed by M El Tantawi, N Aly; comments to author 21.08.20; revised version received 15.10.20; accepted 03.11.20; published 03.12.20.

Please cite as:
Selby K, Durand MA, Gouveia A, Bosisio F, Barazzetti G, Hostettler M, D'Acremont V, Kaufmann A, von Plessen C
Citizen Responses to Government Restrictions in Switzerland During the COVID-19 Pandemic: Cross-Sectional Survey
JMIR Form Res 2020;4(12):e20871
URL: https://formative.jmir.org/2020/12/e20871
doi:10.2196/20871
PMID:33156809

Original Paper

Iterative Adaptation of a Tuberculosis Digital Medication Adherence Technology to Meet User Needs: Qualitative Study of Patients and Health Care Providers Using Human-Centered Design Methods

Devika Patel[1*], MS; Christopher Allen Berger[2*], MS, MD; Alex Kityamuwesi[3], MD; Joseph Ggita[3], MA; Lynn Kunihira Tinka[3], BA; Patricia Turimumahoro[3], MBChB; Joshua Feler[4], MS, MD; Lara Chehab[1], MPH; Amy Z Chen[5], BA; Nakull Gupta[5], BE; Stavia Turyahabwe[6], MPH, MD; Achilles Katamba[3,7], MBChB, PhD; Adithya Cattamanchi[2,3], MSc, MD; Amanda Sammann[1], MPH, MD

[1]Department of Surgery, University of California, San Francisco, San Francisco, CA, United States

[2]Division of Pulmonary and Critical Care Medicine and Center for Tuberculosis, University of California, San Francisco, San Francisco, CA, United States

[3]Uganda Tuberculosis Implementation Research Consortium, Kampala, Uganda

[4]School of Medicine, Yale University, New Haven, CA, United States

[5]Everwell Health Solutions, Bangalore, India

[6]Uganda National Tuberculosis and Leprosy Programme, Kampala, Uganda

[7]Department of Medicine, Makerere University College of Health Sciences, Kampala, Uganda

[*]these authors contributed equally

Corresponding Author:
Christopher Allen Berger, MS, MD
Division of Pulmonary and Critical Care Medicine and Center for Tuberculosis
University of California, San Francisco
1001 Potrero Avenue
San Francisco, CA, 94110
United States
Phone: 1 4152064679
Email: Christopher.berger@ucsf.edu

Abstract

Background: Digital adherence technologies have been widely promoted as a means to improve tuberculosis medication adherence. However, uptake of these technologies has been suboptimal by both patients and health workers. Not surprisingly, studies have not demonstrated significant improvement in treatment outcomes.

Objective: This study aimed to optimize a well-known digital adherence technology, 99DOTS, for end user needs in Uganda. We describe the findings of the ideation phase of the human-centered design methodology to adapt 99DOTS according to a set of design principles identified in the previous inspiration phase.

Methods: 99DOTS is a low-cost digital adherence technology wherein tuberculosis medication blister packs are encased within an envelope that reveals toll-free numbers that patients can call to report dosing. We identified 2 key areas for design and testing: (1) the envelope, including the form factor, content, and depiction of the order of pill taking; and (2) the patient call-in experience. We conducted 5 brainstorming sessions with all relevant stakeholders to generate a suite of potential prototype concepts. Senior investigators identified concepts to further develop based on feasibility and consistency with the predetermined design principles. Prototypes were revised with feedback from the entire team. The envelope and call-in experience prototypes were tested and iteratively revised through focus groups with health workers (n=52) and interviews with patients (n=7). We collected and analyzed qualitative feedback to inform each subsequent iteration.

Results: The 5 brainstorming sessions produced 127 unique ideas that we clustered into 6 themes: rewards, customization, education, logistics, wording and imagery, and treatment countdown. We developed 16 envelope prototypes, 12 icons, and 28 audio messages for prototype testing. In the final design, we altered the pill packaging envelope by adding a front flap to conceal the pills and reduce potential stigma associated with tuberculosis. The flap was adorned with either a blank calendar or map of Uganda. The inside cover contained a personalized message from a local health worker including contact information, pictorial pill-taking instructions, and a choice of stickers to tailor education to the patient and phase of treatment. Pill-taking order was indicated with colors, chevron arrows, and small mobile phone icons. Last, the call-in experience when patients report dosing was changed to a rotating series of audio messages centered on the themes of prevention, encouragement, and reassurance that tuberculosis is curable.

Conclusions: We demonstrated the use of human-centered design as a promising tool to drive the adaptation of digital adherence technologies to better address the needs and motivations of end users. The next phase of research, known as the implementation phase in the human-centered design methodology, will investigate whether the adapted 99DOTS platform results in higher levels of engagement from patients and health workers, and ultimately improves tuberculosis treatment outcomes.

(JMIR Form Res 2020;4(12):e19270) doi:10.2196/19270

KEYWORDS

human-centered design; tuberculosis; digital adherence technology; digital health; adherence; medication; treatment; outcome; lung; respiratory; infectious disease

Introduction

Background

Tuberculosis (TB) is now the leading infectious killer globally, with an estimated 10 million people falling sick annually and nearly 1.5 million dying of the disease [1]. Patient adherence to and completion of TB treatment is a key barrier to reducing mortality and amplification of drug resistance [2]. Recent data suggested that adherence rates of less than 90%, or missing more than 3 doses in a month, can decrease treatment success and cure rates [3]. Directly observed therapy (DOT), wherein a health worker observes a patient taking each daily dose of TB medicines, has been a core component of the World Health Organization (WHO)-recommended strategy for supervising TB treatment [4]. However, DOT is associated with numerous challenges, including questionable effectiveness in improving treatment outcomes [5,6], as well as time and cost burdens for patients and health workers [7,8].

Digital adherence technologies (DATs) have emerged as an alternative to facilitate and promote TB treatment adherence and completion [9-11]. The range and breadth of DATs have expanded rapidly as companies continue to develop unique technological platforms and devices, including (1) short message service (SMS) text message reminders; (2) “dose-in-hand” event monitors, including electronic pill boxes [12,13] and sleeves that fit over medication blister packs [14]; (3) video-observed treatment to remotely record pill ingestion using smartphones; and (4) wireless-observed treatment, in which silicon sensors embedded in TB pills document ingestion [11,15]. Each of these products connects with virtual platforms that allow providers to view real-time adherence data for their patients.

Despite the use of DATs in TB programs [16], supporting data are limited and show minimal to no improvement in treatment success and mortality [9,17]. Additionally, several large-scale implementation studies have shown suboptimal patient engagement with various DATs [18-20]. In response, the WHO and other large funding agencies have called for country-level assessments and stakeholder mapping to support rollout of and user engagement with DATs [15,21].

This study was the second phase in the adaptation of the 99DOTS envelope using the human-centered design (HCD) methodology. HCD is a well-established problem-solving and innovation methodology with the potential to improve adoption and implementation challenges that have impeded the impact of DATs. HCD has been employed in a variety of industries over the past 40 years, notably in business and technology [22-25]. This approach uses qualitative research methods to understand the values and motivations of users, and to systematically incorporate user input to guide the design or adaptation of products and interventions [24]. The primary focus on the user throughout the entire process—problem framing, solution development, implementation, and evaluation—may yield more efficacious and sustainable results [26].

99DOTS is a novel sleeve-based DAT wherein TB medication blister packs are housed within an envelope that reveals hidden, toll-free numbers when pills are pushed out. Patients call these toll-free numbers daily to confirm medication dosing. Clinic staff can access real-time dosing information from both a web and mobile phone app to guide patient follow-up. SMS text messaging reminders, both to patients for dosing and to providers to follow up on patients who have missed doses, are also a core feature of the platform [13,14].

Objective

The first phase of this study used HCD’s approach to human factors research to identify key insights and design opportunities for modifying the design of the original 99DOTS product [27]. The purpose of this second phase of the study was to address these design opportunities through modification of key features of the 99DOTS platform in order to improve TB medication adherence in Uganda.

Methods

Context and Setting

This was the second phase of a 3-phase study using HCD methodology to adapt the 99DOTS platform [13]. In HCD, these phases are termed inspiration, ideation, and implementation. Previously, in the inspiration phase, we conducted 67 semistructured interviews with TB patients, family members, health workers, and community leaders at 8 rural and periurban TB treatment centers in Uganda. Key themes and quotes were elicited from interviews and translated into actionable insights and design opportunities (Table 1) [27]. Here, we report on the ideation phase, which focused on the development and testing of solutions in order to address the insights and design opportunities identified in the inspiration phase. In reporting qualitative aspects of the study, we followed the Consolidated Criteria for Reporting Qualitative Research (COREQ-2) checklist (Multimedia Appendix 1) [28]. The study was approved by the University of California San Francisco, Committee on Human Research, San Francisco, CA, USA (protocol number 221678, June 11, 2018) and the Makerere University School of Public Health Research Ethics Committee, Kampala, Uganda (protocol number HS2436, July 19, 2018).

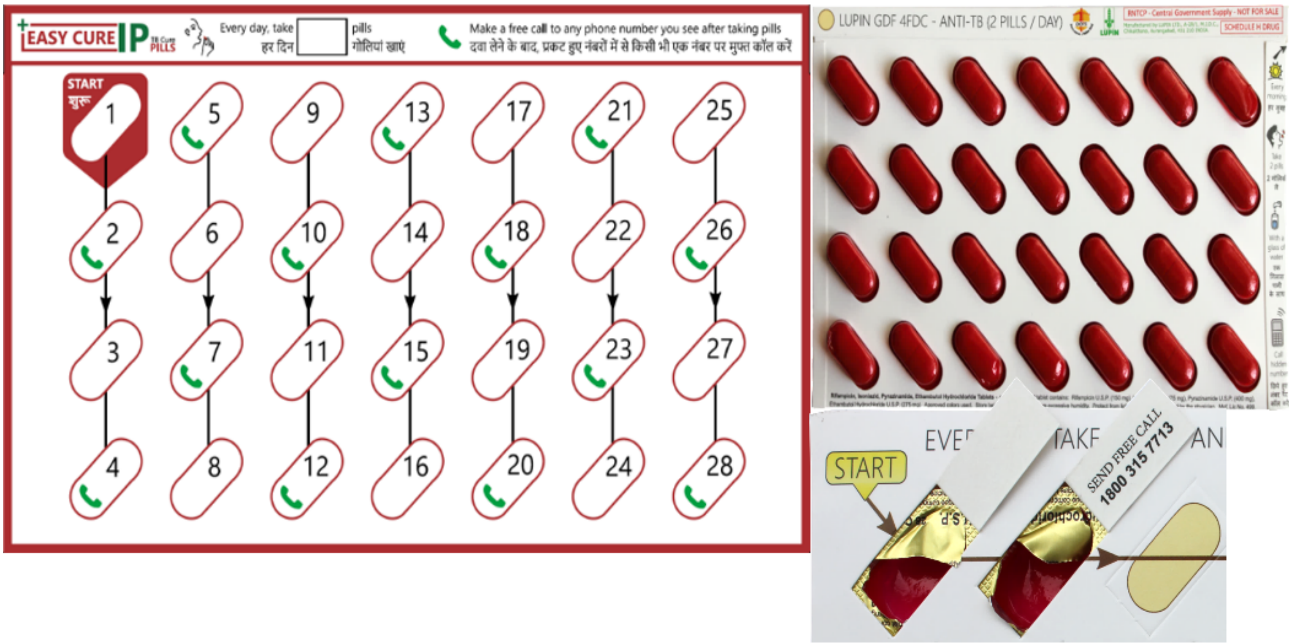
Table 1. Insights and opportunities from the inspiration phase.

Insight	Opportunity
1. Social stigma may be feared as much as the disease.	Make tuberculosis treatment discreet by avoiding nonconspicuous labels and visuals.
2. Packaging is used as an ad hoc and accidental reminder system.	Guide sequential pill-taking behavior and accurately track treatment milestones.
3. Health is defined by a return to normal strength and capability, which occurs before treatment completion.	Reinforce the importance of treatment completion to sustain health at every patient point of contact.
4. Feeling a personal connection with health care workers is as important as receiving medical care.	Leverage the relationship between patients and community health workers to personalize the treatment experience.
5. Local messages feel authentic; endorsements by famous figureheads are viewed with suspicion.	Incorporate authentic and personal adherence messages from the local community in the tuberculosis treatment program.
6. Patients are motivated by service to their community, but do not have enough knowledge education to be maximally effective.	Empower individuals to be community stewards of infection control practices through language and literacy appropriate education.
7. Words and visuals can have multiple interpretations and an unintended negative impact on adherence behavior.	Communicate clearly and accurately with simple graphics and culturally and linguistically appropriate text.

99DOTS System and Redesign Opportunities

This study focused on the redesign of 2 components of the original 99DOTS platform: (1) the pill packaging envelope that reveals hidden, toll-free numbers called by patients to report dosing and (2) the patient experience when calling toll-free numbers (Figure 1). In the original platform, pills are visible on one side of the pill packaging envelope and, as pills are pushed out of the blister pack, a toll-free number is revealed on the opposite side. Upon calling the toll-free number patients hear a beeping sound and the call is logged by the 99DOTS server to indicate a taken dose. The process is repeated on each day the patient is scheduled to take TB medications. A daily dosing history is thus created that providers can access by logging in to an online dashboard via a desktop or mobile phone app and can be used to send SMS text message alerts to providers to notify them of missed consecutive doses. Adaptable features associated with the original pill packaging envelope included its form factor (physical packaging design and external images), content (instructional education and messaging), and visuals (wayfinding instructions for pill taking and revealed phone numbers). The adaptable feature associated with the patient call-in experience was what patients hear when making calls to self-report dosing.

Figure 1. The original 99DOTS product consists of an envelope that fits around the tuberculosis medication blister pack distributed to patients. When patients remove pills from the blister pack each day, a hidden toll-free phone number is revealed. Patients call the number to record medication ingestion, and all calls are logged as a taken dose on the 99DOTS platform. Providers log in to an online dashboard to access patients' adherence history via a desktop or mobile phone app. Text messages are sent to patients daily to remind them to take their medications and to providers to alert them about patients who have missed consecutive doses.



Prototype Development

From the 5 brainstorming sessions informed by the interviews conducted in the previous phase and structured around the design opportunities identified in the inspiration phase, we built a diverse list of unique concepts that we could develop into prototypes. Each session started with a "How might we...?" question that was written to maximize brainstorming output without suggesting a solution [29]. Brainstorming sessions included 3 designers and at least 2 other members of the research team. Each participant was given 3 minutes to generate and capture ideas on sticky notes. The ideas were shared and displayed on foam boards. Participants collaboratively generated additional ideas and discussed whether any potential ideas had been missed. When all 5 brainstorming sessions were complete, we reviewed each idea and clustered redundant or related ideas to represent discrete concepts that we could develop into a prototype. We then organized the concepts into more general themes. Written notes and photographs of the thematic displays captured the output of each brainstorming session.

The general themes and associated concepts were presented to senior investigators (AS, AC, and AKa), who selected the concepts that had the most promise to be developed into prototypes based on the following criteria: (1) whether the concept was clearly linked to a design opportunity identified in the previous phase of research; (2) whether the concept could be applied to one of the adaptable 99DOTS features, and (3) whether the concept was feasible based on the constraints of the 99DOTS platform and available resources. We selected a minimum of 4 concepts addressing at least 2 different design opportunities for each feature of the pill packaging and for the patient call-in experience. Pill packaging prototypes were developed by visual designers using Adobe Creative Cloud (Adobe Inc), and audio messages for the patient call-in experience were drafted in Microsoft Word (Microsoft Corporation). These prototypes were then iteratively revised by team members in Uganda and the United States until consensus was reached. Each prototype selected for field testing was translated from English into Luganda, the primary language spoken at the study research sites.

Prototype Testing

Participant Recruitment

The prototypes were tested in September 2018 by 1 design researcher and 5 local research staff. We recruited health workers and patients (users) for testing from 18 Uganda National Tuberculosis and Leprosy Program TB treatment units participating in a stepped-wedge randomized trial of 99DOTS. To ensure a diverse representation of treatment units, we recruited at least one TB care provider from each of the 18 health centers. All health workers involved in providing TB treatment who attended the public randomization ceremony for the parent trial were invited to participate in a focus group. Prototypes were tested in 5 focus groups with 52 health workers, with each focus group consisting of 9 to 13 participants. During the health facility site visits, we also conducted semistructured interviews with all patients who presented for a TB treatment follow-up visit (n=7). Focus groups were conducted in English, and patient interviews were conducted in English or Luganda, depending on patient preference.

Pill Packaging Envelope

We used a standardized questionnaire (Multimedia Appendix 2) in focus groups and interviews to compare a series of up to 6 prototypes for each feature of the pill packaging envelope. Positive and negative feedback were obtained for each prototype, and users were asked to propose changes to improve each prototype. Feedback received during both focus groups and interviews was recorded in written notes by Ugandan research staff, then reviewed independently by 2 US researchers and clustered into themes.

Patient Call-In Experience

We first tested audio message prototypes in focus groups with health workers during the randomization ceremony. Audio messages were presented by thematic category and read aloud in English. Feedback was solicited after each message, which included real-time edits to the language, suggestions for improvement, and removal of messages that did not resonate with users. The revised messages were then clustered into themes and further evaluated through interviews with patients. Patients were asked what audio messages they would like to hear from their health care workers and then given examples of the different messaging themes generated from the health worker focus groups.

Prototype Iteration

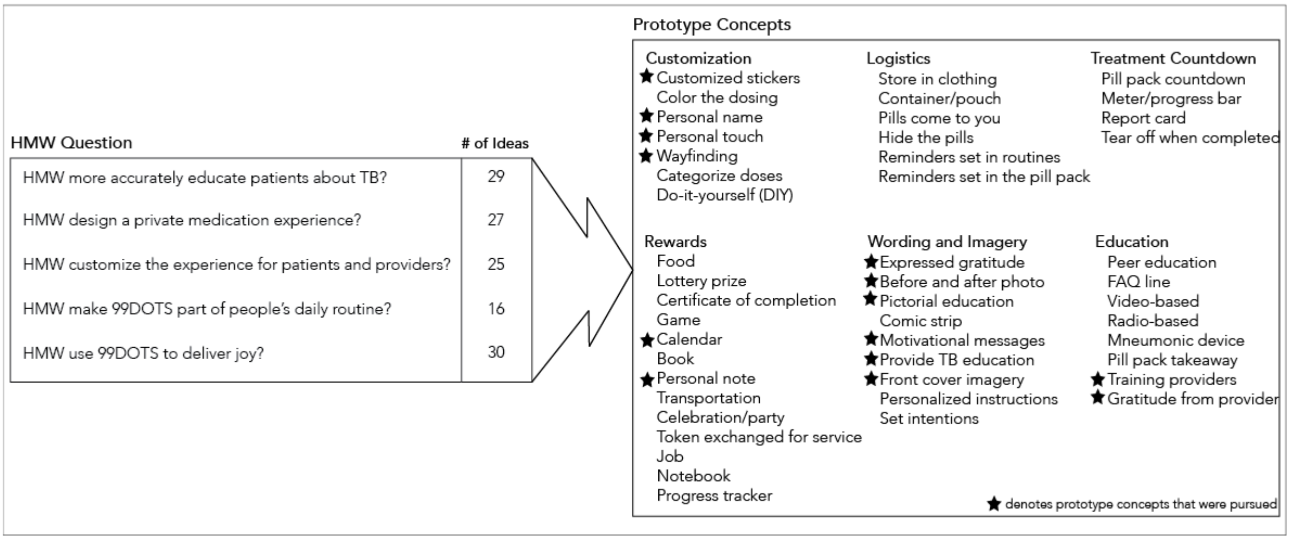
We analyzed qualitative feedback collected during prototype testing to inform the next set of changes to each component. We analyzed the resulting interview transcripts using an inductive approach to thematic analysis [30]. All transcripts were read independently by 2 to 3 members of the research team to identify key themes across the interviews. The reviewers then met to review all thematic categories and ensure that accurate descriptions and supporting quotes depicting positive and negative user feedback were assigned to each theme. Following each iteration, the prototype was refined based on the list of themes that emerged from user feedback, alignment with a particular design opportunity, and feasibility of implementation in the 99DOTS system. We used the final list of themes from each testing session to refine each prototype. The final designs for the pill packaging components and the patient call-in experience were revised over multiple cycles of iteration and feedback from the entire research team, including Uganda- and US-based personnel, until consensus was reached.

Results

Brainstorming

Brainstorming sessions addressed the following “How might we...?” questions: (1) “How might we more accurately educate patients about TB?”; (2) “How might we design a private medication experience?”; (3) “How might we customize the experience for patients and providers?”; (4) “How might we make 99DOTS part of people’s daily routine?”; and (5) “How might we use 99DOTS to deliver joy?” Overall, the brainstorms generated 127 unique ideas. The first session, on education, generated 29 ideas; the second session, on privacy, generated 27 ideas; the third session, on customization, generated 25 ideas; the fourth session, on daily routine, generated 16 ideas; and the fifth session, on delivering joy, generated 30 ideas. We clustered the ideas into 47 unique concepts grouped into 6 major themes: rewards (14 prototype concepts), customization (7 prototype concepts), education (8 prototype concepts), logistics (5 prototype concepts), wording and imagery (9 prototype concepts), and treatment countdown (4 prototype concepts) (Figure 2). These themes were linked back to the design opportunities identified in the previous phase of research (Table 1) to identify which themes had concordance with the design opportunities. This, in turn, informed the prototypes chosen for development and testing.

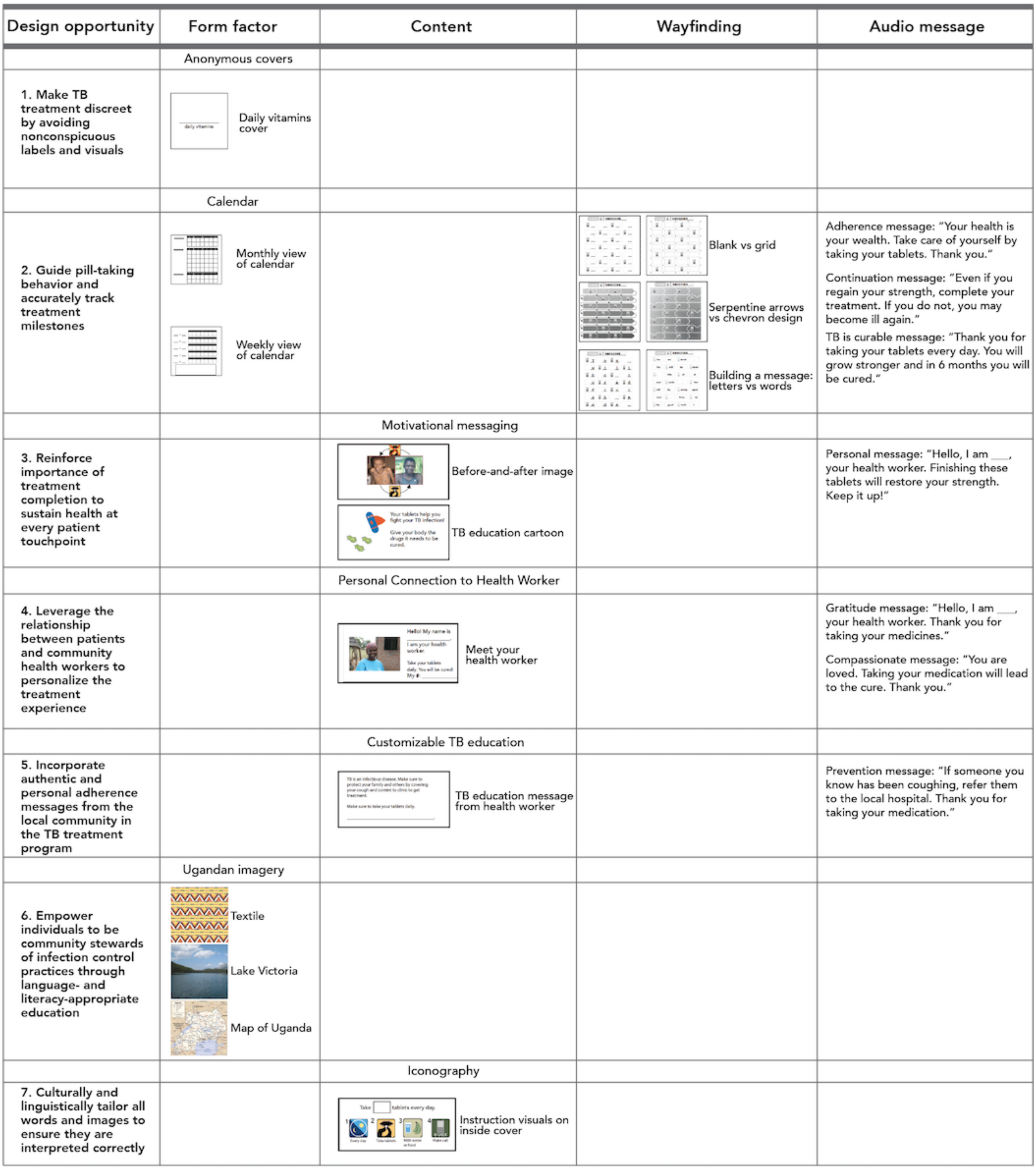
Figure 2. A total of 5 brainstorming sessions centered on one “How might we?” (HMW) question (left). Participants were each allowed 3 minutes to put 1 unique idea on a sticky note, which was later visibly displayed to the group and summated under each HMW question. After all brainstorm sessions were complete, ideas were clustered into 6 themes: customization, rewards, logistics, wording and imagery, treatment countdown, and education. Those ideas that addressed the design opportunities and were feasible to implement moved forward for testing (designated by a star). FAQ: frequently asked questions; TB: tuberculosis.



Prototype Development

We developed 16 packaging prototypes, 12 icons, and 28 messages for testing. For each prototype, Figure 3 indicates the design opportunity addressed and the adapted 99DOTS feature.

Figure 3. For each design opportunity, we prototyped different form factors, content, wayfinding, and motivational messaging. TB: tuberculosis.



Design opportunity	Form factor	Content	Wayfinding	Audio message
	Anonymous covers			
1. Make TB treatment discreet by avoiding nonconspicuous labels and visuals	Daily vitamins cover			
	Calendar			
2. Guide pill-taking behavior and accurately track treatment milestones	Monthly view of calendar Weekly view of calendar		Blank vs grid Serpentine arrows vs chevron design Building a message: letters vs words	Adherence message: "Your health is your wealth. Take care of yourself by taking your tablets. Thank you." Continuation message: "Even if you regain your strength, complete your treatment. If you do not, you may become ill again." TB is curable message: "Thank you for taking your tablets every day. You will grow stronger and in 6 months you will be cured."
		Motivational messaging		
3. Reinforce importance of treatment completion to sustain health at every patient touchpoint		Before-and-after image TB education cartoon		Personal message: "Hello, I am ___, your health worker. Finishing these tablets will restore your strength. Keep it up!"
		Personal Connection to Health Worker		
4. Leverage the relationship between patients and community health workers to personalize the treatment experience		Meet your health worker		Gratitude message: "Hello, I am ___, your health worker. Thank you for taking your medicines." Compassionate message: "You are loved. Taking your medication will lead to the cure. Thank you."
		Customizable TB education		
5. Incorporate authentic and personal adherence messages from the local community in the TB treatment program		TB education message from health worker		Prevention message: "If someone you know has been coughing, refer them to the local hospital. Thank you for taking your medication."
	Ugandan imagery			
6. Empower individuals to be community stewards of infection control practices through language- and literacy-appropriate education	Textile Lake Victoria Map of Uganda			
		Iconography		
7. Culturally and linguistically tailor all words and images to ensure they are interpreted correctly		Instruction visuals on inside cover		

Pill Packaging Envelope Form Factor

Given the importance of discretion (design opportunity 1; Figure 3), we determined that the pills had to be covered. In collaboration with our 99DOTS production partners, we decided that the most feasible design was an additional front flap to cover the pill pack. We subsequently developed 6 discrete prototypes for the front flap imagery: (1) 1 anonymous cover (a daily vitamins label) that addressed discretion (design opportunity 1); (2) 2 calendar covers (weekly and monthly) to track treatment milestones (design opportunity 2); and (3) 3 covers with Ugandan imagery (a map of Uganda, Lake Victoria, and a Ugandan textile design) that addressed national and community pride (design opportunity 6).

Pill Packaging Envelope Content

We developed 4 prototypes for the inside cover. These were (1) 2 educational messages about treatment completion: a before-and-after treatment photo of a patient with TB, and a cartoon reinforcing the importance of TB treatment completion (design opportunity 3) and (2) 2 personal messages: 1 introductory message from a health worker to reinforce their relationship (design opportunity 4) and 1 tailored, personal message from the health worker (design opportunity 5). We also developed 12 unique icons to illustrate the instructions for taking the medication: 2 icons for illustrating daily dosing, 6 icons to illustrate taking pills by mouth, and 4 icons to represent a toll-free phone call.

Pill Packaging Envelope Wayfinding

To assist patient understanding of the order in which to take pills, we developed 4 visual prototypes: a serpentine pattern, a chevron pattern, a customizable pattern, and an unannotated blister pack (design opportunity 2).

Patient Call-In Experience

Given the importance of encouraging adherence at every contact point with a patient (design opportunity 3), we determined that the call-in experience would benefit from adaptation from the original system. Originally, the call concluded with a short “thank you” before its automatic termination. We adapted this feature to include a changing daily audio message. Messages were developed to represent 7 themes: gratitude, adherence, reminders, prevention, encouragement, reassurance that TB is curable, and compassion from health care workers (design opportunities 3, 4, 5, and 6).

Prototype Testing

We tested prototypes with 59 users overall: 5 focus groups comprising 52 health workers representing all 18 health centers enrolled in the stepped-wedge trial, and 7 patients at 2 of the 18 health centers.

Pill Packaging Envelope Form Factor

When we tested the anonymous cover option, users did not view the daily vitamins label favorably. Although mislabeling the cover as daily vitamins was perceived to avoid stigma, health workers and patients were concerned it would lead to unintended use by children or other members of the family. One patient noted: “Those who are not sick would want to take the vitamins.”

Testing of the 2 calendar cover options (weekly vs monthly calendar) revealed that patients universally liked the idea of the calendar to track their daily dosing and treatment. One patient noted: “The calendar can show me the days I have taken the medicine.” However, there was no clear preference for the weekly or monthly option. Health workers expressed mixed sentiments. Some liked that it would help patients track their medications, whereas others were concerned that it would duplicate work they are already doing: “It would be extra work for patient and health care worker.”

Of the 3 Ugandan imagery prototypes tested, most patients (5 of 7) preferred the map of Uganda over the images of Lake Victoria and a Ugandan textile. There was no consensus among health workers. However, patients and health workers believed the map was useful because it indicated that TB treatment is a national effort, and patients stated that they would use the map for directions when traveling: “It is our country, if I am going anywhere I can refer to the map.”

Pill Packaging Envelope Content

Among the educational messages tested, the before-and-after treatment photo of a patient with TB was preferred to the TB treatment cartoon. Health workers believed the before-and-after photo was a powerful reminder of the consequence of nonadherence but were still concerned that it might propagate stigma: “It is a beautiful message but can lead to stigma if someone else sees it.” In the before-and-after photo of a patient with TB, we used circular arrows to show the link between the before image and the after image. Health workers did not like the circular arrows because they suggested a cycle and were concerned that the image represented an extreme case of TB that might not be relevant to the majority of patients: “The picture shows the worst stage you can die, not all patients get wasted like that.” However, all patients believed the images were motivating: “It encourages me to take my medicine; if I don’t take, I will look like this.” Conversely, the TB treatment cartoon was misunderstood by every patient.

There was enthusiasm among both health workers and patients for multiple personal messages. Connecting patients directly to their health worker through the health worker introduction was favored by all patients, especially the inclusion of the health worker’s phone number: “I can easily call or consult my health worker.” The tailored message was universally liked because it addressed patient concerns such as protecting their family and educating others: “If another person has the condition, I can show it.” Patients and health care workers liked the space for custom messages and stated that they wanted their health worker to “add that TB is curable to give hope.” Two patients and all providers requested that the education have more pictures and less text to support individuals with language and literacy challenges.

We tested the 12 unique icons against each other in each category. There was no consensus on the daily dosing icon. Patients liked the sun image because it implied that medications were to be taken during the day. Health care workers preferred the image with a sun and moon because it allowed customization of the dosing time but were concerned that it might be misinterpreted as twice daily dosing. There was universal consensus on the take-pills-by-mouth image, where the majority preferred the image with the side profile and hand holding the pill. For the phone call icon, patients preferred the image with the phone and “0 UGX” (ie, 0 Ugandan schillings) at the bottom because it was most clear. Providers either had no preference or agreed with the patient choice. The biggest concern remained the need to reinforce that this was a toll-free call.

Pill Packaging Wayfinding

Patients and health workers preferred the graphics that clearly defined the direction of pill taking, but there was no consensus on whether this should include the chevron or serpentine pattern. Instead of taking the pills in a left-to-right direction across rows, they preferred to take the pills from the top to the bottom of the pill pack. Patients and health workers preferred color to black-and-white prototypes.

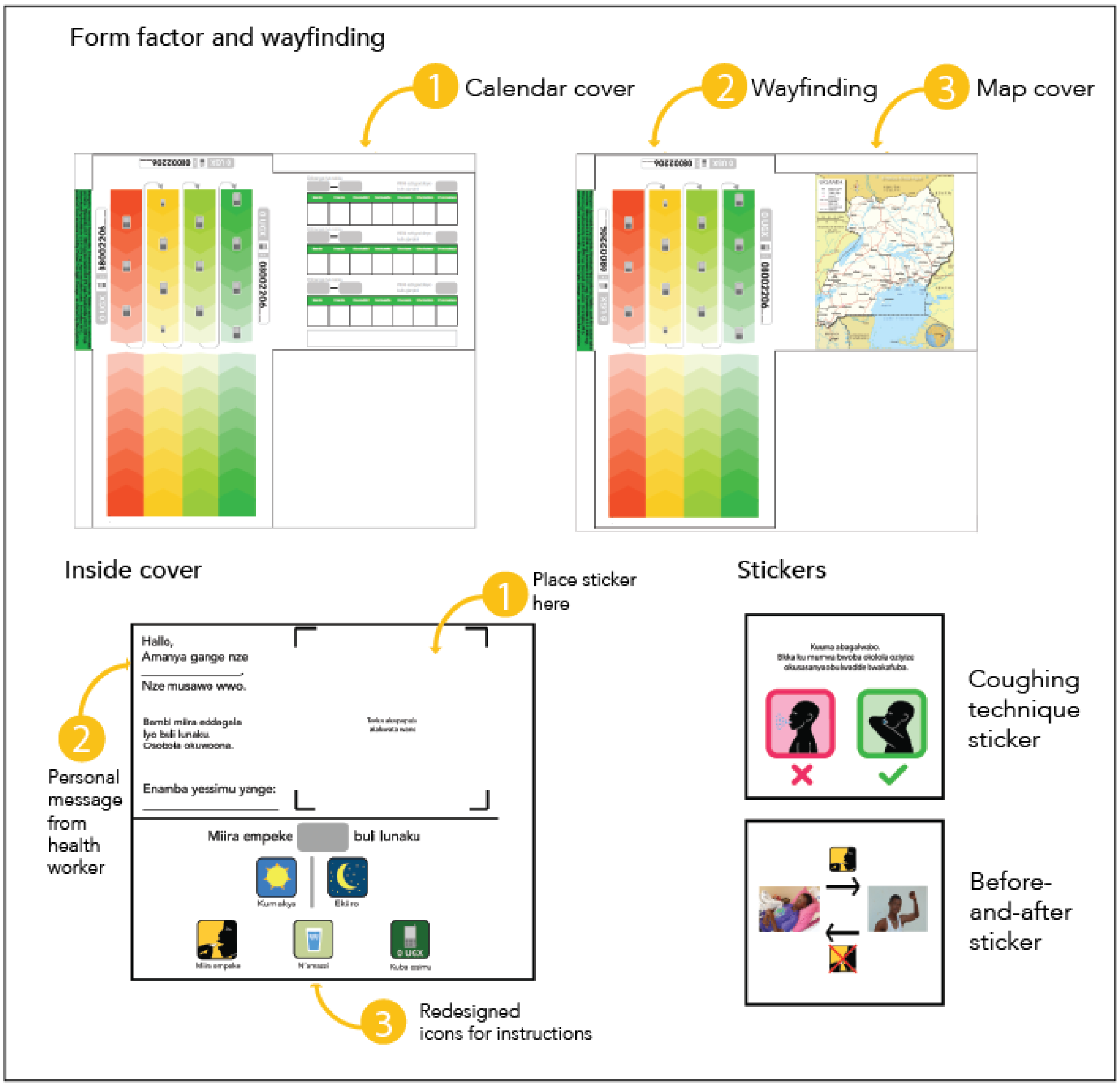
Patient Call-In Experience

We tested 7 themes of audio messaging. The themes that were universally preferred by health workers and patients were reassurance that TB is curable; prevention, especially messages about how to protect others; and encouragement. Patients also preferred compassionate messages from health workers: “Ask if I am ok, have I treated my body well?” Providers additionally preferred messages about adherence and appointment reminders.

Iteration to Final Prototype

We adapted the final prototype described here for the 2 phases of TB treatment (initiation and continuation), which have different pills, doses, and pill packaging shapes (Figures 4 and 5). The form factor included a cover flap with 2 different image options (a map of Uganda or a weekly calendar). This design maintained discretion (design opportunity 1) by providing an unlabeled cover. The covers had more than 1 image to avoid the risk that, with time, a single image would become associated with TB.

Figure 4. Final prototypes of the form factor, content, and wayfinding for the intensive phase of tuberculosis (TB) treatment (first 2 months). A front flap was added to the envelope to conceal the medications with an image of either a calendar (1) or a map of Uganda (3). The wayfinding (2) used color and a chevron pattern with mobile phone icons to indicate dosing days and pill-taking order. The inside covers were customized to include a personal message and contact information for local health workers (2), pictorial pill-taking instructions (3), and several sticker designs to be chosen by the patient (1). Intensive-phase stickers focused on coughing technique when a person may still be infectious and the importance of medication adherence to return to health. Translation of Luganda: personal messages from health worker: “Hello! My name is _____. I am your health worker. Please take your tablets daily. You can be cured! My telephone number: _____”; coughing technique sticker: “Protect your loved ones. Cover your mouth when you cough to prevent the spread of TB.”; blank sticker space: “Place sticker here”; icons: “Take ____ tablets every day”, “Morning”, “Night”, “Take tablets”, “With water”, “Make call”.



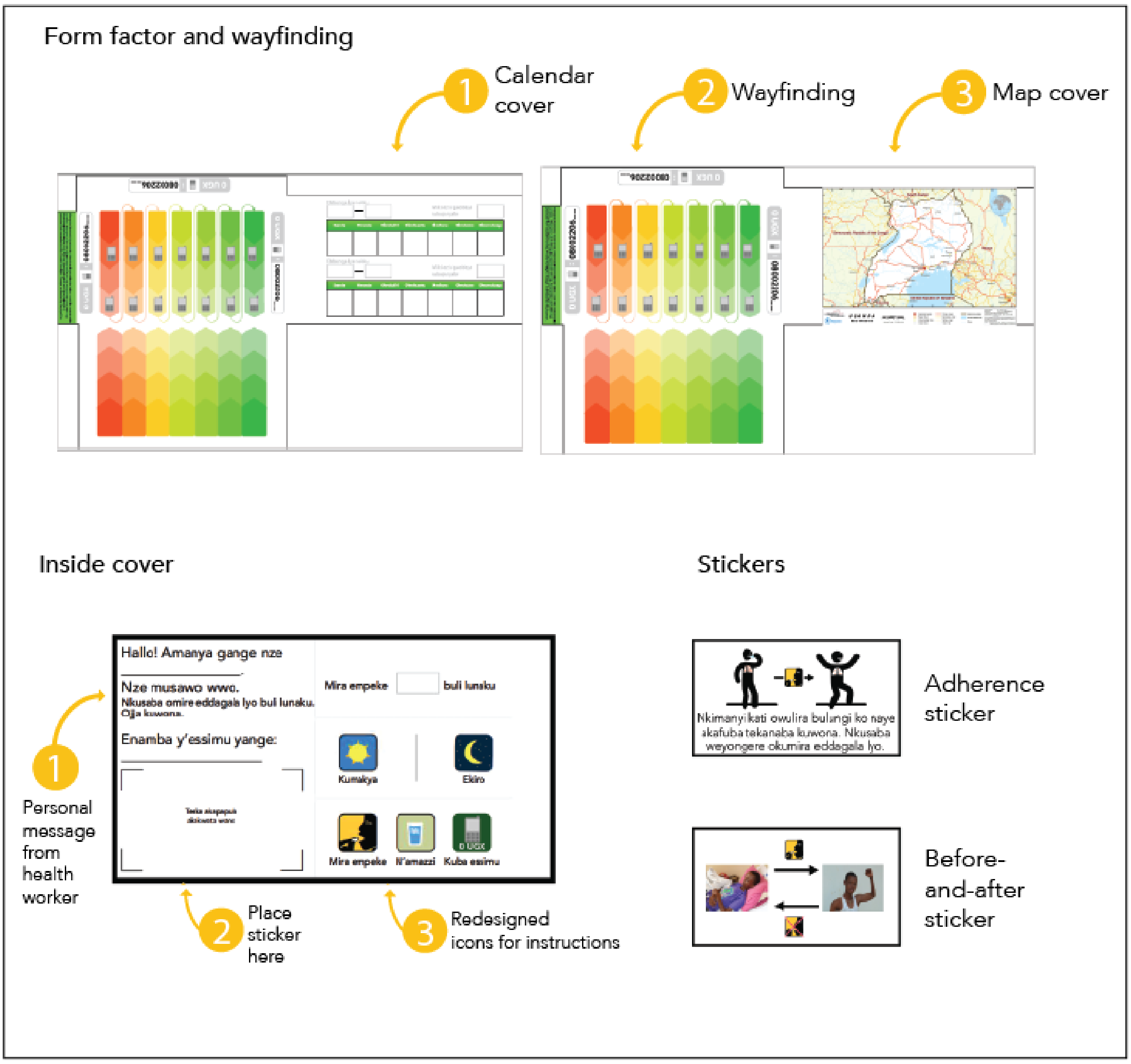
The final pill pack content included a number of unique features:

- A compassionate message from the health worker with space to write the health worker’s name and phone number in order to support this important relationship (design opportunity 4) and to personalize the pill package experience (design opportunity 5).
- Instructional visual icons, which went through multiple iterations of changes to ensure appropriate interpretation (design opportunity 7). Notable changes included separate

icons of a sun and moon to allow patients to select their dosing time, and an icon to indicate that medications should be taken with water so as not to risk nonadherence for people experiencing food insecurity.

- A customizable section on the inside flap with an option to choose 1 of 3 educational or motivational stickers: a before-and-after picture to reinforce treatment adherence (design opportunity 3), which we adapted from feedback to include a less cachectic person and bidirectional rather than circular arrows (design opportunity 7); a cough technique illustration tailored for patients in the initiation phase to support infection control (design opportunity 6); and an illustration tailored for patients in the continuation phase showing that treatment completion leads to cure (design opportunity 3) (Figures 4 and 5).

Figure 5. Final prototypes of the form factor, content, and wayfinding for the continuation phase of tuberculosis (TB) treatment (final 4 months). A front flap was added with an image of either a calendar (1) or a map of Uganda (3). The wayfinding (2) used color and a chevron pattern with mobile phone icons to indicate dosing days and pill-taking order. The inside covers were customized to include a personal message and contact information for local health workers (1), pictorial pill-taking instructions (3), and several sticker designs to be chosen by the patient (2). Both continuation phase stickers focused on the importance of medication adherence, even when a person is feeling better, until the end of treatment. Translation of Luganda: personal messages from health worker: "Hello! My name is ______. I am your health worker. Please take your tablets daily. You can be cured! My telephone number: ______"; adherence sticker: "I know you feel better now, but TB hasn't yet been cured. Please continue taking your medicines."; blank sticker space: "Place sticker here"; icons: "Take _____ tablets every day", "Morning", "Night", "Take tablets", "With water", "Make call".



All of these images prioritized visual graphics over text per our user feedback and design opportunity 7.

The wayfinding graphics were multicolored with a graduated chevron pattern and serpentine directional arrows to guide pill taking (design opportunity 2). We adjusted this pattern to honor the cultural norm by changing the direction of pill taking to flow from top to bottom in columns and included colors that had been tested and approved by patients (design opportunity 7).

A total of 20 audio messages were recorded by local health workers (Textbox 1). We recorded 10 general messages, each beginning with a personal introduction: “Hello, I am your health worker...” (design opportunities 4 and 5). The themes of the general messages included gratitude, that TB is curable, and the importance of adherence. We tailored 5 messages to patients in the initiation phase with the themes of prevention and that TB is curable, and 5 messages to the continuation phase with themes of gratitude and adherence.

Textbox 1. Final audio messages.

GENERAL (10 messages played through entire treatment):

Gratitude

- Hello, I am ______ your health worker. Thank you for taking you medicines.

Adherence

- Hello, I am ______ your health worker. If you finish taking all medicines, you will be cured.
- Hello, I am ______ your health worker. It is essential that you take your medicines. It will lead you to the cure.
- Hello, I am your health worker. Continue to take your tablets each and every day!

Encouragement

- Hello, I am your health worker. Do not lose hope!
- Hello, I am ______ your health worker. Finishing these tablets will restore your strength. Keep it up!

Compassion from health worker

- Hello, I am your health worker. We believe in you.

Prevention

- Hello, I am your health worker. Cover your cough and take your medicines to protect your family and friends.
- Hello, I am your health worker. Please contact me if you feel unwell.

Sputum checkup (time specific)

- This is your reminder to go to clinic and get your sputum checkup. This will help us know whether you are getting better or not.

INITIATION MESSAGES (5 messages only played in initiation phase):

Reassurance that tuberculosis (TB) is curable

- TB is curable if you take your medicine following the health workers’ instructions.
- Thank you for taking your tablets every day. You will grow stronger and in 6 months you can be cured.
- The infection is curable. Remember, you will not have it for life.

Prevention

- Protect your family and friends. Please take your tablets daily.
- Please bring coughing relatives with you for checkup.

CONTINUATION (5 messages only played in continuation phase)

Adherence

- Thank you for taking your medication. By taking your medication every day, you can be cured at the end of 6 months.
- If you don’t take all of your pills, you can get sick again. Thank you.
- You may not feel sick, but TB is still in your body. Please take your tablet for the remaining months to be cured.
- TB can come back if you do not complete your treatment. Take your tablets every day.

Prevention

- Prevention begins with you. Complete your treatment and ensure your health.

Discussion

Principal Findings

In this study representing the ideation phase of HCD research, we adapted existing features of 99DOTS—a low-cost DAT—to better meet user needs and impact TB medication adherence in Uganda. We leveraged the insights and design opportunities identified from the previous inspiration phase to guide tangible, user-centered design alterations to the 99DOTS platform [27,31]. We first identified key components of the 99DOTS platform for customization: the envelope and patient call-in experience. Through a series of brainstorming sessions, we developed a wide range of ideas that we then narrowed based on feasibility and grounding in the outlined design opportunities. We then field tested the resulting prototypes through iterative cycles of feedback and design to create the final prototype that we used in an ongoing stepped-wedge trial of the 99DOTS technology in Uganda.

Current data indicate that the potential impact of DATs has been hindered by a failure to meet the needs and desires of patients and health workers [18,32,33]. As widespread deployment of DATs continues in several high-burden TB countries [14,19], multinational agencies have called for more detailed stakeholder engagement in an effort to adapt these technologies [15,21,34]. The disconnect between currently available technologies and the needs and desires of users has been notable in multiple settings [18,20,33,35-38]. Specifically, evaluations of the 99DOTS platform have highlighted reduced patient contact [18,19], a lack of personalized engagement between providers and patients [20,39], limited patient literacy [33], and ongoing stigma resulting in health care avoidance as key barriers to the successful rollout of 99DOTS [20,38]. In addition, the inspiration phase of our research further highlighted the use of packaging as a reminder system (insight 2), the importance of defining health per local constructs (insight 3), and the need to educate patients about their disease (insight 6). To address these barriers to the adoption and implementation of 99DOTS, our adapted 99DOTS platform includes visuals to guide pill-taking behavior, discreet images depicting messaging about health as a return of strength, and culturally and linguistically tailored educational and motivational messages from health workers.

The WHO's *Handbook for the Use of Digital Technologies to Support Tuberculosis Medication Adherence* recommends that, when implementing digital health solutions, it is necessary for implementors to thoroughly assess the context of the health care system and position the patients at the center, with their needs and challenges incorporated into the implementation and decision-making process [11]. HCD, when done correctly, facilitates this process by providing a structured, user-centered approach to iteratively develop solutions that target barriers and facilitators identified in formative research. Determining how best to address barriers and facilitators is a key challenge for implementation science. Our previous experience [40], as well as that of others, shows that implementation failures are common, even when interventions are informed by implementation theories or frameworks [41-43]. To overcome this challenge, HCD allows for a broad range of solutions to be considered and either rejected or refined through user testing prior to large-scale implementation studies. Although HCD is relatively new in global health research, there are now several examples of HCD-informed approaches in addressing implementation problems. In Kenya, Catalani et al used HCD to create a TB clinical decision support system to improve uptake of TB prevention therapy among people living with HIV [26], and a group in Colombia recently demonstrated the use of HCD to develop 4 novel prototypes to improve cervical cancer screening [44]. These examples, in addition to our work, add to the growing literature on the potential benefits of using the HCD approach in addressing implementation barriers.

Limitations

Our study has several important limitations. The outputs of brainstorming sessions in part reflect the participants included. While we did not include patients and health workers for logistical reasons, the sessions included a diverse group of designers and researchers, including Ugandan social scientists and public health researchers. The relatively small number of patients available to be interviewed for prototype testing is a limitation, though patient perspectives obtained during the prior phase of research informed the initial prototypes. We plan for additional patient interviews to obtain further feedback during the subsequent implementation phase of the project.

Conclusions

We adapted a DAT to the expressed needs and desires of health workers and TB patients using HCD methodology. The results of our ongoing trial will confirm whether the adaptations made to strengthen connections between 99DOTS and its intended users result in improved TB medication adherence and treatment outcomes.

Acknowledgments

We would like to express our appreciation to the health workers and patients who participated in this study and provided open and honest feedback that was critical to adapting the 99DOTS platform. We would also like to acknowledge Talemwa Nalugwa and Sara Nabwire, who helped with facilitating focus groups.

Funding for this work was provided by the STOP TB Partnership (TB REACH Wave 6), the UCSF Nina Ireland Program for Lung Health, and the US National Institutes of Health (R01HL144406).

Authors' Contributions

AS, AC, AKa, and ST conceived of the study. DP, CAB, AKi, JG, LKT, and PT collected data. AS, DP, CAB, JF, AKi, JG, LKT, PT, AC, and AKa analyzed data. AS, AC, DP, CAB, and LC drafted the manuscript. All authors critically reviewed and revised the manuscript.

Conflicts of Interest

AS, LC, and DP are human-centered design consultants for The Empathy Studio.

Multimedia Appendix 1
COREQ-2 checklist.
[PDF File (Adobe PDF File), 67 KB - formative_v4i12e19270_app1.pdf]

Multimedia Appendix 2
Focus-group standardized questionnaire.
[DOCX File , 147 KB - formative_v4i12e19270_app2.docx]

References

1. World Health Organization. Global Tuberculosis Report 2019. Geneva, Switzerland: WHO; 2019. URL: https://www.who.int/tb/publications/global_report/en/ [accessed 2020-11-02]
2. World Health Organization. Guidelines for Treatment of Tuberculosis, 4th edition. Geneva, Switzerland: WHO; 2010. URL: https://wwwwhoint/tb/publications/2010/9789241547833/en/ [accessed 2020-11-02]
3. Imperial MZ, Nahid P, Phillips PPJ, Davies GR, Fielding K, Hanna D, et al. A patient-level pooled analysis of treatment-shortening regimens for drug-susceptible pulmonary tuberculosis. Nat Med 2018 Nov;24(11):1708-1715 [FREE Full text] [doi: 10.1038/s41591-018-0224-2] [Medline: 30397355]
4. World Health Organization. International Standards for Tuberculosis Care (ISTC) and the Patients' Charter for Tuberculosis Care. Geneva, Switzerland: WHO; 2006.
5. Karumbi J, Garner P. Directly observed therapy for treating tuberculosis. Cochrane Database Syst Rev 2015 May 29(5):CD003343 [FREE Full text] [doi: 10.1002/14651858.CD003343.pub4] [Medline: 26022367]
6. Volmink J, Garner P. Directly observed therapy for treating tuberculosis. Cochrane Database Syst Rev 2007 Oct 17(4):CD003343. [doi: 10.1002/14651858.CD003343.pub3] [Medline: 17943789]
7. McKay B, Castellanos M, Ebell M, Whalen CC, Handel A. An attempt to reproduce a previous meta-analysis and a new analysis regarding the impact of directly observed therapy on tuberculosis treatment outcomes. PLoS One 2019;14(5):e0217219 [FREE Full text] [doi: 10.1371/journal.pone.0217219] [Medline: 31120965]
8. McLaren ZM, Milliken AA, Meyer AJ, Sharp AR. Does directly observed therapy improve tuberculosis treatment? More evidence is needed to guide tuberculosis policy. BMC Infect Dis 2016 Oct 04;16(1):537 [FREE Full text] [doi: 10.1186/s12879-016-1862-y] [Medline: 27716104]
9. Subbaraman R, de Mondesert L, Musiimenta A, Pai M, Mayer KH, Thomas BE, et al. Digital adherence technologies for the management of tuberculosis therapy: mapping the landscape and research priorities. BMJ Glob Health 2018;3(5):e001018 [FREE Full text] [doi: 10.1136/bmjgh-2018-001018] [Medline: 30364330]
10. World Health Organization. Digital Health for the END TB Strategy: An Agenda for Action. Geneva, Switzerland: WHO; 2015. URL: https://www.who.int/tb/publications/digitalhealth-TB-agenda/en/ [accessed 2020-11-02]
11. World Health Organization. Handbook for the Use of Digital Technologies to Support Tuberculosis Medication Adherence. Geneva, Switzerland: WHO; 2018. URL: https://wwwwhoint/tb/publications/2018/TB_medication_adherence_handbook_2018/en/ [accessed 2020-11-02]
12. evriMED. Somerset West, South Africa: Wisepill URL: https://www.wisepill.com/evrimed/ [accessed 2020-11-02]
13. 99DOTS, low-cost monitoring and improving medication adherence.: Everwell Health Solutions URL: https://www.99dots.org [accessed 2020-11-02]
14. Cross A, Gupta N, Liu B, Nair V, Kumar A, Kuttan R. 99DOTS: a low-cost approach to monitoring and improving medication adherence. 2019 Presented at: Tenth International Conference on Information and Communication Technologies and Development; Jan 4-7, 2019; Ahmedabad, India p. Article 15. [doi: 10.1145/3287098.3287102]
15. van Rest J. Digital adherence technologies. Den Haag, Netherlands: KNCV; 2020 Jan 21. URL: https://kncvtbc.atlassian.net/wiki/spaces/ADHERENCE/pages/222494774/Digital+Adherence+Technologies [accessed 2020-11-02]
16. KNCV Tuberculosis Foundation. ASCENT: Empowering Patients Through Digital Technology. Den Haag, Netherlands: KNCV; 2019 Jul 23. URL: https://www.kncvtbc.org/en/2019/07/23/ascent-empowering-patients-through-digital-technology-kncv-tuberculosis-foundation-leads-unitaid-funded-project/ [accessed 2020-11-02]

17. Ngwatu BK, Nsengiyumva NP, Oxlade O, Mappin-Kasirer B, Nguyen NL, Jaramillo E, Collaborative group on the impact of digital technologies on TB. The impact of digital health technologies on tuberculosis treatment: a systematic review. Eur Respir J 2018 Jan;51(1) [FREE Full text] [doi: 10.1183/13993003.01596-2017] [Medline: 29326332]
18. Liu X, Blaschke T, Thomas B, De Geest S, Jiang S, Gao Y, et al. Usability of a medication event reminder monitor system (MERM) by providers and patients to improve adherence in the management of tuberculosis. Int J Environ Res Public Health 2017 Sep 25;14(10) [FREE Full text] [doi: 10.3390/ijerph14101115] [Medline: 28946683]
19. Liu X, Lewis JJ, Zhang H, Lu W, Zhang S, Zheng G, et al. Effectiveness of electronic reminders to improve medication adherence in tuberculosis patients: a cluster-randomised trial. PLoS Med 2015 Sep;12(9):e1001876 [FREE Full text] [doi: 10.1371/journal.pmed.1001876] [Medline: 26372470]
20. Thekkur P, Kumar AN, Chinnakali P, Selvaraju S, Bairy R, Singh AR, et al. Outcomes and implementation challenges of using daily treatment regimens with an innovative adherence support tool among HIV-infected tuberculosis patients in Karnataka, India: a mixed-methods study. Glob Health Action 2019;12(1):1568826 [FREE Full text] [doi: 10.1080/16549716.2019.1568826] [Medline: 30712507]
21. Falzon D, Timimi H, Kurosinski P, Migliori GB, Van GW, Denkinger C, et al. Digital health for the End TB Strategy: developing priority products and making them work. Eur Respir J 2016 Jul;48(1):29-45 [FREE Full text] [doi: 10.1183/13993003.00424-2016] [Medline: 27230443]
22. Brown T. Design thinking. Harv Bus Rev 2008 Jun;86(6):84-92, 141. [Medline: 18605031]
23. Harte R, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, Quinlan LR, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. JMIR Hum Factors 2017 Mar 16;4(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.5443] [Medline: 28302594]
24. Brown T, Katz B. Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation. New York, NY: HarperBusiness; 2019.
25. Bazzano AN, Martin J, Hicks E, Faughnan M, Murphy L. Human-centred design in global health: a scoping review of applications and contexts. PLoS One 2017;12(11):e0186744 [FREE Full text] [doi: 10.1371/journal.pone.0186744] [Medline: 29091935]
26. Catalani C, Green E, Owiti P, Keny A, Diero L, Yeung A, et al. A clinical decision support system for integrating tuberculosis and HIV care in Kenya: a human-centered design approach. PLoS One 2014;9(8):e103205 [FREE Full text] [doi: 10.1371/journal.pone.0103205] [Medline: 25170939]
27. Berger C, Patel D, Kityamuwesi A, Ggita J, Tinka LK, Turimumahoro P, et al. Opportunities to improve digital adherence technologies and TB care using human-centered design. Int J Tuberc Lung Dis 2020 Oct 01;24(10):1112-1115. [doi: 10.5588/ijtld.20.0184]
28. 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]
29. IDEO.org. The Field Guide to Human-Centered Design: Design Kit. San Francisco, CA: IDEO.org; 2012:37.
30. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate - service users' experiences of doing collaborative research on mental health. Disabil Rehabil 2012;34(19):1608-1616. [doi: 10.3109/09638288.2012.656792] [Medline: 22489612]
31. Berger C, Patel D. Optimising novel tuberculosis medication adherence technologies utilising human-centred design. 2019 Presented at: 50th World Conference on Lung Health of the International Union Against Tuberculosis and Lung Disease; Oct 30-Nov 2, 2019; Hyderabad, India.
32. Haberer JE, Subbaraman R. Digital technology for tuberculosis medication adherence: promise and peril. Ann Am Thorac Soc 2020 Apr;17(4):421-423. [doi: 10.1513/AnnalsATS.202001-027ED] [Medline: 32233860]
33. Thomas BE, Kumar JV, Onongaya C, Bhatt SN, Galivanche A, Periyasamy M, et al. Explaining differences in the acceptability of 99DOTS, a cell phone–based strategy for monitoring adherence to tuberculosis medications: qualitative study of patients and health care providers. JMIR Mhealth Uhealth 2020 Jul 31;8(7):e16634 [FREE Full text] [doi: 10.2196/16634] [Medline: 32735220]
34. Zelezny-Green R, Vosloo S, Conole G. Digital Inclusion for Low-Skilled and Low-Literate People: A Landscape Review. Geneva, Switzerland: United Nations Educational, Scientific and Cultural Organization; 2018. URL: https://unesdoc.unesco.org/ark:/48223/pf0000261791 [accessed 2020-11-02]
35. Thomas BE, Kumar JV, Chiranjeevi M, Shah D, Khandewale A, Thiruvengadam K, et al. Evaluation of the accuracy of 99DOTS, a novel cellphone-based strategy for monitoring adherence to tuberculosis medications: Comparison of digital adherence data with urine isoniazid testing. Clin Infect Dis 2020 Mar 28. [doi: 10.1093/cid/ciaa333] [Medline: 32221550]
36. Wang N, Zhang H, Zhou Y, Jiang H, Dai B, Sun M, et al. Using electronic medication monitoring to guide differential management of tuberculosis patients at the community level in China. BMC Infect Dis 2019 Oct 15;19(1):844 [FREE Full text] [doi: 10.1186/s12879-019-4521-2] [Medline: 31615433]
37. Musiimenta A, Tumuhimbise W, Mugaba AT, Muzoora C, Armstrong-Hough M, Bangsberg D, et al. Digital monitoring technologies could enhance tuberculosis medication adherence in Uganda: mixed methods study. J Clin Tuberc Other Mycobact Dis 2019 Dec;17:100119 [FREE Full text] [doi: 10.1016/j.jctube.2019.100119] [Medline: 31788561]

38. Drabarek D, Anh NT, Nhung NV, Hoa NB, Fox GJ, Bernays S. Implementation of medication event reminder monitors among patients diagnosed with drug susceptible tuberculosis in rural Viet Nam: a qualitative study. PLoS One 2019;14(7):e0219891 [FREE Full text] [doi: 10.1371/journal.pone.0219891] [Medline: 31329610]
39. Thomas B, Kumar JV, Chiranjeevi M, Ramachandran G, Periyasamy M, Khandelwale AS, et al. Updated findings from an evaluation of the accuracy of 99DOTS, a digital technology for monitoring tuberculosis medication adherence in HIV co-infected and uninfected patients. 2019 Presented at: 50th Union World Conference on Lung Health; Oct 30-Nov 2, 2019; Hyderabad, India.
40. Meyer AJ, Armstrong-Hough M, Babirye D, Mark D, Turimumahoro P, Ayakaka I, et al. Implementing mHealth interventions in a resource-constrained setting: case study from Uganda. JMIR Mhealth Uhealth 2020 Jul 13;8(7):e19552 [FREE Full text] [doi: 10.2196/19552] [Medline: 32673262]
41. Dharmayat KI, Tran T, Hardy V, Chirambo BG, Thompson MJ, Ide N, et al. Sustainability of 'mHealth' interventions in sub- Saharan Africa: a stakeholder analysis of an electronic community case management project in Malawi. Malawi Med J 2019 Sep;31(3):177-183 [FREE Full text] [doi: 10.4314/mmj.v31i3.3] [Medline: 31839886]
42. Brown W, Giguere R, Sheinfil A, Ibitoye M, Balan I, Ho T, et al. Challenges and solutions implementing an SMS text message-based survey CASI and adherence reminders in an international biomedical HIV PrEP study (MTN 017). J Biomed Inform 2018 Apr;80:78-86 [FREE Full text] [doi: 10.1016/j.jbi.2018.02.018] [Medline: 29501908]
43. Bardosh KL, Murray M, Khaemba AM, Smillie K, Lester R. Operationalizing mHealth to improve patient care: a qualitative implementation science evaluation of the WelTel texting intervention in Canada and Kenya. Global Health 2017 Dec 06;13(1):87 [FREE Full text] [doi: 10.1186/s12992-017-0311-z] [Medline: 29208026]
44. Arrivillaga M, Bermúdez PC, García-Cifuentes JP, Botero J. Innovative prototypes for cervical cancer prevention in low-income primary care settings: a human-centered design approach. PLoS One 2020;15(8):e0238099 [FREE Full text] [doi: 10.1371/journal.pone.0238099] [Medline: 32833974]

Abbreviations

DAT: digital adherence technology
DOT: directly observed therapy
HCD: human-centered design
SMS: short message service
TB: tuberculosis
WHO: World Health Organization

Edited by G Eysenbach; submitted 17.04.20; peer-reviewed by R Mpofu, B Smith, A Kibballi Madhukeshwar, A Meyer; comments to author 08.05.20; revised version received 29.06.20; accepted 24.09.20; published 08.12.20.

Please cite as:
Patel D, Berger CA, Kityamuwesi A, Ggita J, Kunihira Tinka L, Turimumahoro P, Feler J, Chehab L, Chen AZ, Gupta N, Turyahabwe S, Katamba A, Cattamanchi A, Sammann A
Iterative Adaptation of a Tuberculosis Digital Medication Adherence Technology to Meet User Needs: Qualitative Study of Patients and Health Care Providers Using Human-Centered Design Methods
JMIR Form Res 2020;4(12):e19270
URL: https://formative.jmir.org/2020/12/e19270
doi:10.2196/19270
PMID:33289494

Original Paper

Psychological Effects of the COVID-19 Imposed Lockdown on Adults with Attention Deficit/Hyperactivity Disorder: Cross-Sectional Survey Study

Marios Adamou[1*], MD, MA, MSc, LLM, MBA, PhD, FRCPsych, FFOM; Tim Fullen[2*], DClinPsy; Nazmeen Galab[2], BSc; Isobel Mackintosh[2], BSc; Karl Abbott[2*], BSc; Deborah Lowe[2], BSc; Claire Smith[2*], BA

[1]School of Human and Health Sciences, University of Huddersfield, Huddersfield, United Kingdom

[2]South West Yorkshire Partnership NHS Foundation Trust, Wakefield, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Marios Adamou, MD, MA, MSc, LLM, MBA, PhD, FRCPsych, FFOM
School of Human and Health Sciences
University of Huddersfield
Queensgate
Huddersfield, HD1 3DH
United Kingdom
Phone: 44 1924316492
Email: m.adamou@hud.ac.uk

Abstract

Background: The psychological effects of the COVID-19 government-imposed lockdown have been studied in several populations. These effects however have not been studied in adult populations with attention deficit/hyperactivity disorder (ADHD).

Objective: We wanted to investigate the psychological effects of the COVID-19 imposed lockdown on an adult population with ADHD.

Methods: We conducted a cross-sectional survey by administering the Patient Health Questionnaire-9, Generalized Anxiety Disorder Scale-7, Brief Adjustment Scale-6, Perceived Stress Scale, and Multidimensional Scale of Perceived Social Support to a pragmatic sample of adults with ADHD.

Results: In total, 24 individuals (male: n=18, 75%; female: n=6, 25%; age: mean 21.75 years, SD 1.85 years) were included in this study. The adults with ADHD we surveyed had significant levels of emotional distress during the COVID-19 pandemic period. However, there was no evidence of significant deterioration to the mental health of our sample during the COVID-19 pandemic.

Conclusions: When treatment for ADHD is maintained, the effects of the COVID-19 pandemic on the mental health of adults with ADHD are mild. Targeted psychological interventions may be useful in such circumstances.

(JMIR Form Res 2020;4(12):e24430) doi:10.2196/24430

KEYWORDS

adult ADHD; pandemic; lockdown; COVID-19; well-being; psychological; intervention; ADHD

Introduction

Attention-deficit/hyperactivity disorder (ADHD) is one of the most common neuropsychiatric conditions. The pooled worldwide prevalence of ADHD is estimated to be approximately 5% in school-aged children, and impairing symptoms have persisted in adulthood in up to 65% of cases. The pooled estimated prevalence of ADHD in adults is approximately 2.5% [1]. ADHD is characterized by a persistent and impairing pattern of inattention and hyperactivity/impulsivity that causes significant impairment across many domains of mental health [2]. Along with these 3 main symptoms, people with ADHD also present with deficits in executive functions, behavior and emotion regulation, and motivation [3].

In December 2019, an outbreak of a novel coronavirus pneumonia occurred in Wuhan (Hubei, China), and subsequently attracted worldwide attention [4]. As a result, countries around the world introduced public health policies that have been commonly called a lockdown. These policies, which include individuals isolating at home and only allowing necessary movement, were indented to reduce the spread of the virus. The United Kingdom was put into lockdown on March 23, 2020.

As the lockdown progressed, concerns were raised about the effect of this lockdown on the mental health of individuals [5]. A general population survey administered by Ipsos MORI [6] has revealed widespread concerns about the impact of social isolation and social distancing on a person's well-being, such as increased anxiety, depression, stress, and other negative feelings.

The mechanism behind the impact of the lockdown can be understood by considering the lockdown as an event that has created stressful life events in a person's life. Stressful life events are described as discrete, quantifiable circumstances that can have a severe negative impact. A wealth of published evidence has shown that life event stressors are associated with the onset of depressive episodes and that there is a dose-response relationship between these stressors and depressive episodes [7]. Stressors are associated with severe initial depressive symptoms in both adult patients [8] in nonpatient community samples [9].

Stress is generally conceptualized as a negative cognitive-emotional state caused by an individual's perception of difficulty in adjusting to or managing life events [10]. Studies have shown that adults with ADHD have elevated physiological stress responses and higher levels of self-reported subjective stress compared to nondiagnosed controls [11,12]. In fact, the levels of stress for adults diagnosed with ADHD appear elevated in the presence of a stressor and in anticipation of the stressor. In addition, a study has suggested that adults with ADHD have greater difficulty in recovering from elevated stress levels than adults without ADHD [12], indicating that persons with ADHD may also face challenges in recovering from stress.

Based on this context, we believe that adults with ADHD are disproportionately affected by the lockdown compared to other populations. We therefore administered a survey to a group of adults with ADHD to investigate the psychological effects of the COVID-19 imposed lockdown on this population, understand this population's needs, and provide appropriate support.

Methods

We conducted a cross-sectional survey study. The convenience sample was derived from the caseload of the ADHD and Autism Service at South West Yorkshire Partnership NHS Foundation Trust. The ADHD and Autism Service provides input to approximately 1.2 million people in West Yorkshire, England, for people with ADHD and autism. The ADHD and Autism Service can only offer pharmacological interventions to people with ADHD aged >25 years. Therefore, only people aged <25 years were eligible for this study. Our survey was considered a service improvement activity because the results of the survey were to be used to identify people who were suitable for COVID-19–specific psychological intervention.

Our cohort consisted of 94 individuals with a preestablished diagnosis of ADHD (male: n=75; female: n=19) who were first contacted on May 18, 2020, for a period of 2 weeks over the telephone. After 3 unsuccessful attempts of making contact with an individual, the person was considered not-responding, and no further attempts were made. All individuals were receiving pharmacological treatment.

For the people who did respond to the telephone call, after the purpose of the well-being survey was explained using a predetermined script, the Patient Health Questionnaire-9 (PHQ-9) [13], Generalized Anxiety Disorder Scale-7 (GAD-7) [14], Brief Adjustment Scale-6 (BASE-6) [15], Perceived Stress Scale (PSS) [16], and Multidimensional Scale of Perceived Social Support (MSPSS) [17] were administered.

The PHQ-9 and GAD-7 use a scale of 0 (not at all) to 3 (nearly every day) to measure the severity of self-reported depression and anxiety, respectively, within the last two weeks. They have 9 and 7 questions, respectively. PHQ-9 scores ranging from 0 to 5 represent mild depression, 6-10 represent moderate depression, 11-15 represent moderately severe depression, and 16-20 represent severe depression. GAD-7 scores of ≤4 predict no anxiety, 5-9 predict mild anxiety, 10-14 predict moderate anxiety, and 15-21 predict severe anxiety.

The BASE-6 measures a person's general psychological adjustment within the last week, with ratings between 1 (not at all) and 7 (extremely) given for each of the 6 questions.

The PSS consists of 10 items designed to assess the degree to which common life situations in the past month are deemed stressful, and responses are made using a 5-point scale (0=never, 4=very often). Individual scores on the PSS can range from 0 to 40, with higher scores indicating higher perceived stress. Scores ranging from 0 to 13 indicate low perceived stress, 14-26 indicate moderate perceived stress, and 27-40 indicate high perceived stress.

The MSPSS measures the perception of support from the following 3 sources: family, friends, and a significant other. The scale is comprised of 12 items on a 7-point Likert scale, with 4 items for each subscale. Higher scores reflect more perceived social support. Mean scores ranging from 1 to 2.9 indicate low perceived support, 3-5 indicate moderate perceived support, and 5.1-7 indicate high perceived support.

The results from these scales, along with demographic information, were recorded using Microsoft Excel. The results were then exported to IBM SPSS statistical software for analysis.

Results

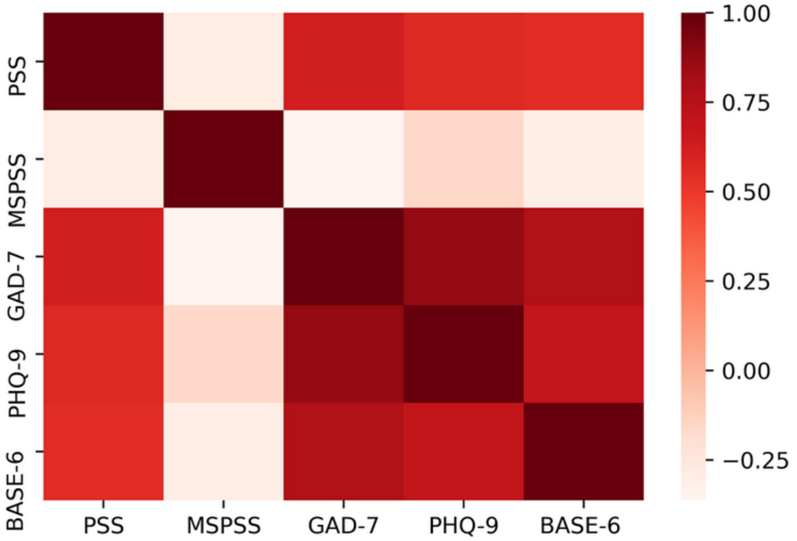
Of the 94 individuals eligible for the survey, 75 were male and 19 were female. The mean age of the total sample was 22.05 years (SD 1.837 years). Of these 94 individuals, 70 (74%; male: n=57; female: n=13) were excluded from this study, either because they could not be contacted or because they refused to take part in this study. Of these 70, 18 (26%; male: n=16; female: n=2) declined to take part in this study and 52 (74%; male: n=41; female: n=11) could not be contacted. In total, 24 patients (male: n=18, 75%; female: n=6, 25%) were included in this study.

There was no significant difference in age between males who were excluded from this study (mean 22.26 years, SD 1.86 years) and included males (mean 21.94 years, SD 1.74 years; t_{73}=0.64; *P*=.53). There was also no significant difference in age between females who were excluded from this study (mean 21.62 years, SD 2.18 years) and included females (mean 21.17 years, SD 1.83 years; t_{17}=0.47; *P*=.65).

For the 24 included participants, there were no missing data. The scores from the scales were checked for normality using the Kolmogorov-Smirnov test. All scores had a normal distribution. The mean age of all included patients was 21.75 years (SD 1.85 years). Based on our survey results, the mean PSS score was 21.5 (SD 7), the mean MSPSS score was 5.55 (SD 0.89), the mean GAD-7 score was 4.84 (SD 1.46), the mean PHQ-9 score was 10.88 (SD 6.83), and the mean BASE-6 score was 17.13 (SD 9.56).

A Pearson correlation coefficient was computed to assess the relationship between the variables. We found a positive correlation between PSS scores and GAD-7 (*r*=.623; *P*=.01), PHQ-9 (*r*=.572; *P*=.03), and BASE-6 (*r*=.553; *P*=.05) scores; GAD-7 scores and PSS (*r*=.623; *P*=.01), PHQ-9 (*r*=.872; *P*<.001) BASE-6 (*r*=.778; *P*<.001) scores; and PHQ-9 scores and PSS (*r*=.572; *P*=.03), GAD-7 (*r*=.872; *P*<.001), BASE-6 (*r*=.696; *P*<.001) scores. All correlations were significant at the .01 level and all tests were 2-tailed. There was no correlation between age or MSPSS scores with any other variables (Figure 1).

Figure 1. Heat map of correlations between the survey scale scores. BASE-6: Brief Adjustment Scale-6; GAD-7: Generalized Anxiety Disorder Scale-7; MSPSS: Multidimensional Scale of Perceived Social Support; PHQ-9: Patient Health Questionnaire; PSS: Perceived Stress Scale.



Discussion

Principle Results

Our study showed that the adults with ADHD we surveyed had significant levels of emotional distress during the COVID-19 pandemic. It was not clear if these levels were due to the psychological effects of the pandemic and the social changes it imposed or just what would be normally expected in this population. Furthermore, the pragmatic study design did not allow us to attribute a differential negative impact to the effect of the pandemic on this population and the general population. However, a discussion of the wider context of our findings generated some interesting ideas for future research.

In terms of anxiety, it is already known that adults with ADHD are associated with positive screening results [18]. In a sample of 141 adults with ADHD, Goniu and Moreno [19] reported that 48 (34.1%) adults met the criteria for mild anxiety, 31 (22%) met the criteria for moderate anxiety, and 18 (6.7%) met the criteria for severe anxiety. For our sample, the GAD-7 scores were in the mild to moderate range, even in the context of the global pandemic, indicating that our participants' anxiety was not much worse than what has been previously reported.

In terms of depression, surprisingly, there was not much information in the current literature to know what the expected PHQ-9 scores would be for adults with ADHD. We know that the expected prevalence rate of PHQ-9 depressive disorders in the general population is 9.2% [20] and that the PHQ-9 scores we reported (mean 10.88, SD 6.83) reliably differentiate our study population from the general population. However, we cannot claim that this difference was a result of the impact of the pandemic, especially when the comorbidity rate for adult ADHD with major depressive disorder has been reported to be as high as 16% [21], depending on the sample.

We used the PSS to measure participants' perception of stress. We believed that there was adequate reason for any person to be under some sort of personal stress due to the lockdown and its consequences for social life. For adults with ADHD specifically, it has been argued that their perception of stress would be even higher, as it has been suggested that adults with ADHD experience elevated perceived or subjective stress compared to adults without ADHD [22]. Our participants' perceived stress was only slightly higher than those reported in adults without ADHD [23]. However, the standard deviation in our sample's PSS scores suggest greater variability, which may be attributable to the social consequences of the lockdown. There is not enough research to allow us to make accurate comparisons with this scale for adults with ADHD. However, another study that used the PSS recruited a sample of people with posttraumatic stress disorder [24], and another study on people with ADHD [21] reported a lower mean PSS score than our study (mean 16.7, SD 1.6).

In terms of the level of psychological adjustment in 1 week, before we contacted our sample, their mean BASE-6 score was 17.13 (SD 9.56), which was similar to the clinical sample (mean 18.9, SD 7.0) and college student sample (mean 18.1 SD 6.8) scores reported during the evaluation of the BASE-6 [15]. However, it should be noted that the clinical and college student sample scores were reported outside of a pandemic context. Therefore, there is no evidence to suggest that the level of psychological adjustment of our sample was impaired when measured using the BASE-6. Furthermore, the reason why our participants' BASE-6 scores were not as high as expected was not because social support was available. This is because participants' MSPSS scores did not correlate with their BASE-6 scores.

Our participants' mean MSPSS score was 5.55 (SD 0.89), which indicated a high level of support. Previous studies that used this scale for people with ADHD [25,26] have reported that the perception on the importance of social support between people with ADHD and people without ADHD was similar, even though the scores differed between the groups.

In this study, the GAD-7, PHQ-9, PSS, and BASE-6 scores positively correlated with each other, suggesting that in circumstances where the impact of stress is poor, psychological adjustment, mood, and anxiety will deteriorate. Based on this, it might be useful to have an available form of psychological intervention specific to current circumstances to prevent further deterioration.

We are aware that this study has several limitations; the sample size was small, the dropout rate was high, and the sample was not representative of all adults due to only including people aged 18-25 years. Despite this, our findings provided a window into how adults with ADHD fared with the social changes caused by a global pandemic that restricted their movement and relationships. In our sample, participants seemed to manage the same as their expected baseline, and this may be somewhat explained by the fact that they were able to continue their medication. Our findings also suggest that it may be useful to have an available form of psychological intervention for increasing resilience in adults with ADHD, in case other life events or a prolonged pandemic cause more severe deterioration to their mental health.

Conclusions

We expected that adults with ADHD would find the restrictions imposed by the lockdown particularly difficult to adhere to, and as a result, their mental health would be at risk of deterioration. Our findings suggest that there was evidence of mental ill health, but this was not severe or disproportionate to what was expected for this population. This may have been due to all participants continuing their pharmacological treatment for ADHD. However, due to their reported distress, a preventative targeted psychological intervention may be useful.

Conflicts of Interest

None declared.

References

1. Thapar A, Cooper M. Attention deficit hyperactivity disorder. Lancet 2016 Mar 19;387(10024):1240-1250. [doi: 10.1016/S0140-6736(15)00238-X] [Medline: 26386541]
2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Arlington, VA: American Psychiatric Association; 2013:991.
3. Lopez PL, Torrente FM, Ciapponi A, Lischinsky AG, Cetkovich-Bakmas M, Rojas JI, et al. Cognitive-behavioural interventions for attention deficit hyperactivity disorder (ADHD) in adults. Cochrane Database Syst Rev 2018 Mar 23;3:CD010840 [FREE Full text] [doi: 10.1002/14651858.CD010840.pub2] [Medline: 29566425]
4. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. Lancet 2020 Feb 15;395(10223):470-473 [FREE Full text] [doi: 10.1016/S0140-6736(20)30185-9] [Medline: 31986257]
5. 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]
6. Covid-19 and mental wellbeing. Ipsos MORI. 2020 Apr 15. URL: https://www.ipsos.com/ipsos-mori/en-uk/Covid-19-and-mental-wellbeing [accessed 2020-11-02]
7. Kessler RC. The effects of stressful life events on depression. Annu Rev Psychol 1997;48:191-214. [doi: 10.1146/annurev.psych.48.1.191] [Medline: 9046559]
8. Mahatme SS, Dhavale HS, Patkar AP. Study of correlation of intensity of symptoms with stressful life events in depressed patients. Indian J Psychiatry 1989 Oct;31(4):288-292 [FREE Full text] [Medline: 21927402]
9. Bebbington PE, Hurry J, Tennant C. Adversity and the symptoms of depression. Int J Soc Psychiatry 1988;34(3):163-171. [doi: 10.1177/002076408803400301] [Medline: 3225138]

10. Taylor SE. Health Psychology, Sixth Edition. New York, NY: McGraw-Hill Higher Education; 2006.
11. Hirvikoski T, Lindholm T, Nordenström A, Nordström AL, Lajic S. High self-perceived stress and many stressors, but normal diurnal cortisol rhythm, in adults with ADHD (attention-deficit/hyperactivity disorder). Horm Behav 2009 Mar;55(3):418-424. [doi: 10.1016/j.yhbeh.2008.12.004] [Medline: 19162030]
12. Lackschewitz H, Hüther G, Kröner-Herwig B. Physiological and psychological stress responses in adults with attention-deficit/hyperactivity disorder (ADHD). Psychoneuroendocrinology 2008 Jun;33(5):612-624. [doi: 10.1016/j.psyneuen.2008.01.016] [Medline: 18329819]
13. Kroenke K, Spitzer RL. The PHQ-9: A New Depression Diagnostic and Severity Measure. Psychiatr Ann 2002 Sep 01;32(9):509-515 [FREE Full text] [doi: 10.3928/0048-5713-20020901-06]
14. 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]
15. Cruz R, Peterson AP, Fagan C, Black W, Cooper L. Evaluation of the Brief Adjustment Scale-6 (BASE-6): A measure of general psychological adjustment for measurement-based care. Psychol Serv 2020 Aug;17(3):332-342. [doi: 10.1037/ser0000366] [Medline: 31169389]
16. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [doi: 10.2307/2136404] [Medline: 6668417]
17. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The Multidimensional Scale of Perceived Social Support. J Pers Assess 1988 Mar;52(1):30-41 [FREE Full text] [doi: 10.1207/s15327752jpa5201_2]
18. de Zwaan M, Gruss B, Müller A, Graap H, Martin A, Glaesmer H, et al. The estimated prevalence and correlates of adult ADHD in a German community sample. Eur Arch Psychiatry Clin Neurosci 2012 Feb;262(1):79-86. [doi: 10.1007/s00406-011-0211-9] [Medline: 21499942]
19. Goniu N, Moreno M. 125. Increased Risk for Anxiety Among College Students With ADHD. J Adolesc Health 2013 Feb;52(2):S81. [doi: 10.1016/j.jadohealth.2012.10.189]
20. Martin A, Rief W, Klaiberg A, Braehler E. Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. Gen Hosp Psychiatry 2006;28(1):71-77. [doi: 10.1016/j.genhosppsych.2005.07.003] [Medline: 16377369]
21. Alpert JE, Maddocks A, Nierenberg AA, O'Sullivan R, Pava JA, Worthington JJ, et al. Attention deficit hyperactivity disorder in childhood among adults with major depression. Psychiatry Res 1996 Jun 01;62(3):213-219. [doi: 10.1016/0165-1781(96)02912-5] [Medline: 8804131]
22. Krause KH, Krause J, Trott GE. Hyperkinetic syndrome (attention deficit-/hyperactivity disorder) in adulthood. Nervenarzt 1998 Jul;69(7):543-556. [doi: 10.1007/s001150050311] [Medline: 9715472]
23. González-Ramírez MT, Rodríguez-Ayán MN, Hernández RL. The perceived stress scale (PSS): normative data and factor structure for a large-scale sample in Mexico. Span J Psychol 2013;16:E47. [doi: 10.1017/sjp.2013.35] [Medline: 23866243]
24. Antshel KM, Biederman J, Spencer TJ, Faraone SV. The Neuropsychological Profile of Comorbid Post-Traumatic Stress Disorder in Adult ADHD. J Atten Disord 2016 Dec;20(12):1047-1055. [doi: 10.1177/1087054714522512] [Medline: 24567364]
25. Demaray MK, Elliott SN. Perceived social support by children with characteristics of Attention-Deficit/Hyperactivity Disorder. Sch Psychol 2001;16(1):68-90 [FREE Full text] [doi: 10.1521/scpq.16.1.68.19156]
26. Heiman T, Olenik-Shemesh D, Eden S. Cyberbullying involvement among students with ADHD: relation to loneliness, self-efficacy and social support. Eur J Spec Needs Educ 2014 Jul 29;30(1):15-29 [FREE Full text] [doi: 10.1080/08856257.2014.943562]

Abbreviations

ADHD: attention deficit/hyperactivity disorder
BASE-6: Brief Adjustment Scale-6
GAD-7: Generalized Anxiety Disorder Scale-7
MSPSS: Multidimensional Scale of Perceived Social Support
PHQ-9: Patient Health Questionnaire
PSS: Perceived Stress Scale

Edited by J Torous; submitted 19.09.20; peer-reviewed by N Kyriakidou, H van Marwijk; comments to author 17.10.20; revised version received 18.10.20; accepted 20.10.20; published 15.12.20.

Please cite as:
Adamou M, Fullen T, Galab N, Mackintosh I, Abbott K, Lowe D, Smith C
Psychological Effects of the COVID-19 Imposed Lockdown on Adults with Attention Deficit/Hyperactivity Disorder: Cross-Sectional Survey Study
JMIR Form Res 2020;4(12):e24430
URL: https://formative.jmir.org/2020/12/e24430
doi:10.2196/24430
PMID:33108312

Original Paper

Psychological Impacts of COVID-19 During the First Nationwide Lockdown in Vietnam: Web-Based, Cross-Sectional Survey Study

Khanh Ngoc Cong Duong[1*], MSc; Tien Nguyen Le Bao[2*], MD, PhD; Phuong Thi Lan Nguyen[3*], BSc; Thanh Vo Van[2,4], MD; Toi Phung Lam[1,5], MD; Anh Pham Gia[6], MD; Luerat Anuratpanich[3], PhD; Bay Vo Van[7], PhD

[1]Mahidol University Health Technology Assessment (MUHTA) Graduate Program, Mahidol University, Bangkok, Thailand
[2]Institute of Orthopaedics and Trauma Surgery, Viet Duc Hospital, Hanoi, Vietnam
[3]Social Economics and Administrative Pharmacy Program, Department of Pharmacy, Faculty of Pharmacy, Mahidol University, Bangkok, Thailand
[4]Department of Surgery, Hanoi Medical University, Hanoi, Vietnam
[5]Health Strategy and Policy Institute, Ministry of Health, Hanoi, Vietnam
[6]Oncology Department, Viet Duc Hospital, Hanoi, Vietnam
[7]Department of Pharmacy, Thong Nhat Hospital, Ho Chi Minh City, Vietnam
[*]these authors contributed equally

Corresponding Author:
Bay Vo Van, PhD
Department of Pharmacy
Thong Nhat Hospital
1 Ly Thuong Kiet Street, Ward 7
Tan Binh District
Ho Chi Minh City, 70000
Vietnam
Phone: 84 988889315
Email: vovanbay2005@yahoo.com.vn

Related Article:

This is a corrected version. See correction statement: https://formative.jmir.org/2021/3/e28357/

Abstract

Background: The first nationwide lockdown due to the COVID-19 pandemic was implemented in Vietnam from April 1 to 15, 2020. Nevertheless, there has been limited information on the impact of COVID-19 on the psychological health of the public.

Objective: This study aimed to estimate the prevalence of psychological issues and identify the factors associated with the psychological impact of COVID-19 during the first nationwide lockdown among the general population in Vietnam.

Methods: We employed a cross-sectional study design with convenience sampling. A self-administered, online survey was used to collect data and assess psychological distress, depression, anxiety, and stress of participants from April 10 to 15, 2020. The Impact of Event Scale-Revised (IES-R) and the Depression, Anxiety, and Stress Scale-21 (DASS-21) were utilized to assess psychological distress, depression, anxiety, and stress of participants during social distancing due to COVID-19. Associations across factors were explored using regression analysis.

Results: A total of 1385 respondents completed the survey. Of this, 35.9% (n=497) experienced psychological distress, as well as depression (n=325, 23.5%), anxiety (n=195, 14.1%), and stress (n=309, 22.3%). Respondents who evaluated their physical health as average had a higher IES-R score (beta coefficient [B]=9.16, 95% CI 6.43 to 11.89), as well as higher depression (B=5.85, 95% CI 4.49 to 7.21), anxiety (B=3.64, 95% CI 2.64 to 4.63), and stress (B=5.19, 95% CI 3.83 to 6.56) scores for DASS-21 than those who rated their health as good or very good. Those who self-reported their health as bad or very bad experienced more severe depression (B=9.57, 95% CI 4.54 to 14.59), anxiety (B=7.24, 95% CI 3.55 to 10.9), and stress (B=10.60, 95% CI 5.56 to 15.65). Unemployment was more likely to be associated with depression (B=3.34, 95% CI 1.68 to 5.01) and stress (B=2.34, 95% CI 0.84 to 3.85). Regarding worries about COVID-19, more than half (n=755, 54.5%) expressed concern for their children aged <18 years, which increased their IES-R score (B=7.81, 95% CI 4.98 to 10.64) and DASS-21 stress score (B=1.75, 95% CI 0.27 to 3.24). The majority of respondents (n=1335, 96.4%) were confident about their doctor's expertise in terms of

COVID-19 diagnosis and treatment, which was positively associated with less distress caused by the outbreak (B=–7.84, 95% CI –14.58 to –1.11).

Conclusions: The findings highlight the effect of COVID-19 on mental health during the nationwide lockdown among the general population in Vietnam. The study provides useful evidence for policy decision makers to develop and implement interventions to mitigate these impacts.

(JMIR Form Res 2020;4(12):e24776) doi:10.2196/24776

KEYWORDS

COVID-19; mental health; psychological distress; depression; anxiety; Vietnam; psychology; distress; lockdown; survey

Introduction

COVID-19 was first reported in December 2019 in China. It was then declared as a pandemic on March 11, 2020 [1]. Globally, there were 58,087 confirmed cases and up to 3670 fatalities, as of March 31, 2020 [2].

Vietnam detected its first two COVID-19 cases on January 22, 2020 [3]. Since then, the country has faced an increased likelihood of infection spread due to its proximity to China and its low-resource health care system. Responding to this urgency, the Vietnamese government declared COVID-19 as an epidemic on February 1, 2020. It implemented proactive actions to confront the spread of the disease [4]. The number of cases, however, increased approximately by 15-folds in 2 weeks from 15 cases on March 15, 2020, to 206 cases on March 31, 2020 [5]. Notably, the epicenter was the largest public hospital in Hanoi, Vietnam [6]. Accordingly, the government imposed the first nationwide lockdown from April 1 to 15, 2020 [7]. It was a partial lockdown that required people to stay at home (with exceptions made for the purchase of critical supplies, health emergencies, and essential workers and businesses) along with personal protective measures such as wearing face masks in public places and hand washing [7]. As the government has clearly and frequently communicated with the public about the dangers of the illness as well as preventive measures through text messages and social media, citizens are aware of the complications of the pandemic, and thus voluntarily abide regulations [8].

Several social distancing measures, such as home quarantine, school closures, and nonessential business closures, have caused negative psychological effects on populations, including depression, anxiety, and posttraumatic distress, which were also found in previous epidemics [9]. Nguyen et al [10] surveyed the early stage of the pandemic from February 14 to March 13, 2020, and found that patients with suspected COVID-19 symptoms were 2.88 times (95% CI 2.18 to 3.80) more likely to experience depression and significantly lower health-related quality of life (HRQOL) (P<.001) as compared with those without suspected COVID-19 symptoms. The study also reported that the prevalence of depression was 7.4% [10]. With dramatically increasing COVID-19 cases and a compulsory nationwide lockdown, it is anticipated that the prevalence of mental health issues in the general population in Vietnam would surge considerably. We therefore conducted this study to investigate the psychological impacts of COVID-19 among the general population in Vietnam during the first national lockdown and to identify the associated factors of these impacts.

Methods

Participants and Procedure

A cross-sectional study design was employed using a respondent-driven sampling method by distributing a self-administered survey through various social media platforms (eg, Facebook, Zalo, etc) from April 10 to 15, 2020, in Vietnam. All Vietnamese residents who were aged ≥18 years and provided informed consent were eligible to participate in the online survey.

The starting point for data collection was 1 week following the implementation of the national lockdown (April 1, 2020) to ensure that participants had been exposed to a sufficient amount of time to social distancing in order to properly measure psychological impacts over the past week [11,12]. A written informed consent was received online before the respondents answered the questionnaire. They clicked the link on the survey platform and voluntarily responded to the survey. Anonymity and confidentiality were ensured throughout the conduct of the survey.

Questionnaire

We employed a self-administered questionnaire, which consisted of three parts and could be completed in 10 minutes. The first part covered the general sociodemographic characteristics of the participants, including age, gender, area of residence, education, marital status, income, occupation, history of chronic disease, household size, current physical health status, and other basic information. The second part was composed of questions related to concerns about the COVID-19 pandemic, which were adapted from a previous study [13]. The third part of the survey included the two psychological and mental health measurement scales described below.

Impact of Event Scale-Revised

The Impact of Event Scale-Revised (IES-R) has 22 items that measure psychological distress after exposure to a crisis. The total IES-R score was categorized as normal (0-23), mild (24-32), moderate (33-36), and severe (≥37) [14]. The IES-R questionnaire has been used in the Vietnamese population in previous studies [14-16]. In this study, a score of ≥24 signified a positive cutoff point to estimate distress due to the COVID-19 lockdown [11,17].

Depression, Anxiety, and Stress Scale-21

Impacts on mental health, including depression, anxiety, and stress, were assessed using the Depression, Anxiety, and Stress

Scale-21 (DASS-21). The questionnaire consisted of 21 items that covered three domains: depression, anxiety, and stress. The score calculation for each domain was adapted from a previous study [18]. The total score range for each domain was 0 to 42. The level of each domain was interpreted as normal (0-9), mild (10-12), moderate (13-20), severe (21-27), and extremely severe (28-42) for the depression subscale; normal (0-6), mild (7-9), moderate (10-14), severe (15-19), and extremely severe (20-42) for the anxiety subscale; and normal (0-10), mild (11-18), moderate (19-26), severe (27-34), and extremely severe (35-42) for the stress subscale. Cutoff scores of 9, 6, and 10 for the depression, anxiety, and stress subscales, respectively, were used to detect depression, anxiety, and stress. The DASS-21 questionnaire has been assessed for reliability and validity in the Vietnamese context previously [19,20].

The data collection tool was piloted in 10 individuals to further develop the questionnaire, and a few questions were adjusted in terms of language and idea expression.

Data Analysis

Descriptive statistics were used to present sociodemographic characteristics. Continuous variables were presented as mean and standard deviation if the data were normally distributed, or median and interquartile range for nonnormal distribution. To identify variables that could be associated with psychological distress, depression, anxiety, and stress, univariate analyses were performed. Subsequently, all variables with a *P*<.20 in the analyses were entered into the multivariate regression model [21]. The association between factors and IES-R score, and DASS-21 depression, anxiety, and stress subscale scores, were reported with beta coefficients (B) and 95% CIs. *P*<.05 was set as the level of statistical significance. All of the statistical analyses were performed using R software, version 3.6.3 (The R Project for Statistical Computing) [22].

This online survey was reported in accordance with CHERRIES (Checklist for Reporting Results of Internet E-Surveys) [23] (Multimedia Appendix 1).

Ethics Statement

The study protocol was approved by the Council of Medical Ethics at Thong Nhat Hospital in Ho Chi Minh City, Vietnam (number 10/BB-BVTN).

Results

A total of 1412 participants filled out the questionnaire. Of these, 1385 (98.1%) were valid for analysis. Reasons for exclusion were invalid age information (n=8), age <18 years (n=10), and respondents living outside Vietnam (n=9).

The majority of the respondents were male (n=880, 63.5%). The median age was 28 years (range 18-70 years), with 82.2% (n=1139) aged 18-39 years. In terms of area of residence, 73.0% (n=1011) lived in urban areas and 30.3% (n=419) lived in provinces or cities with COVID-19 cases. The majority obtained an undergraduate or higher degree (n=1174, 84.8%). As for employment, most were employed during the lockdown (n=895, 64.6%); a few worked from home (n=446, 32.2%) and some were unemployed (n=113, 8.2%). Most respondents (n=1243, 89.7%) had at least one chronic disease. Additional sociodemographic characteristics of participants are presented in Multimedia Appendix 2.

The total IES-R score ranged from 0 to 88 with a median of 17 (IQR 20). The DASS-21 scores ranged from 0 to 42 (median 2, IQR 8) for depression, 0 to 38 (median 0, IQR 4) for anxiety, and 0 to 42 (median 0, IQR 10) for stress. Using the cutoff points, we found that 35.9% (n=497) had psychological distress. In addition, respondents experienced symptoms of depression (n=325, 23.5%), anxiety (n=195, 14.1%), and stress (n=309, 22.3%), respectively. These results are displayed in Multimedia Appendix 3.

In multivariate regression analysis, associations were found between IES-R scores and variables marital status and health status. Respondents who were married, divorced, or widowed had higher IES-R scores compared with those who were single (B=2.60, 95% CI 0.51 to 4.69 vs B=5.23, 95% CI 0.50 to 9.95). Individuals who self-evaluated their physical health as average had higher psychological distress than those who reported good or very good health (B=9.16, 95% CI 6.43 to 11.89) (Table 1).

Table 1. Multivariate linear regression results for the Impact of Event Scale-Revised questionnaire scores and sociodemographic covariates.

Covariates	Coefficient (95% CI)	*P* value
Age group (years) (reference: 18-39)		
40-59	1.23 (–1.13 to 3.58)	.31
≥60	–8.36 (–17.72 to 10)	.08
Marital status (reference: single)		
Married	2.60 (0.51 to 4.69)	.01
Divorced/widowed	5.23 (0.50 to 9.95)	.03
Education level (reference: elementary/secondary)		
High school	3.20 (–4.18 to 10.58)	.40
University/college	5.60 (–1.51 to 12.71)	.12
Postgraduate	4.40 (–2.93 to 11.74)	.24
Occupation (reference: employed)		
Work from home	0.95 (–1.24 to 3.14)	.40
Student	0.98 (–2.15 to 4.11)	.54
Unemployed	3.32 (–0.01 to 6.65)	.051
Other	1.37 (–3.17 to 5.91)	.55
Household size (reference: 1 member)		
2	3.86 (–0.05 to 7.77)	.05
3-5	1.65 (–1.65 to 4.95)	.33
≥6	2.24 (–1.61 to 6.10)	.25
Children <18 years old in the family (reference: no)		
Yes	1.02 (–0.80 to 2.83)	.27
Chronic disease (reference: no)		
Yes	1.38 (–1.23 to 40)	.30
Average income per month (million VND[a]) (reference: no income)		
<1	1.92 (–3.34 to 7.18)	.47
1-5	–0.17 (–3.27 to 2.94)	.91
5-10	1.53 (–1.46 to 4.53)	.32
10-20	–1.12 (–4.34 to 2.10)	.49
>20	–0.22 (–3.75 to 3.32)	.90
Average length of stay (hours) at home during the lockdown (reference: 0-10)		
11-20	–0.94 (–3.53 to 1.65)	.48
21-24	–2.74 (–5.60 to 0.12)	.06
Current health status (reference: good/very good)		
Average	9.16 (6.43 to 11.89)	<.001
Bad/very bad	8.72 (–1.38 to 18.81)	.09

[a] VND: Vietnam Dong.

Age, marital status, occupation, chronic disease, physical health conditions, and residents in areas with infected cases were found to be associated with depression in the multiple regression model. People aged ≥60 years were less depressed than people aged 18-39 years (B=–5.86, 95% CI –10.52 to –1.20). Married respondents had smaller depression scores than their single counterparts (B=–1.02, 95% CI –2.00 to –0.03). Those who were unemployed (B=3.34, 95% CI 1.68 to 5.01), students (B=1.76, 95% CI 0.19 to 3.32), and those who worked from home (B=1.33, 95% CI 0.24 to 2.43) had a higher risk of depression compared to employed participants. Having chronic disease was associated with higher depression (B=1.32, 95% CI 0.02 to 2.63). People with an average, bad, or very bad physical health self-assessment exhibited higher depression scores than those who reported good or very good health (B=5.85, 95% CI 4.49 to 7.21 vs B=9.57, 95% CI 4.54 to 14.59). People residing in provinces or cities with COVID-19 cases had an elevated risk for depression (B=1.69, 95% CI 0.83 to 2.55) compared with those who lived elsewhere (Table 2).

Table 2. Multivariate linear regression results for the Depression, Anxiety, and Stress Scale-21 (depression subscale) with sociodemographic covariates.

Covariates	Coefficient (95% CI)	*P* value
Age group (years) (reference: 18-39)		
40-59	–0.12 (–1.29 to 1.05)	.84
≥60	–5.86 (–10.52 to –1.20)	.01
Marital status (reference: single)		
Married	–1.02 (–2.00 to –0.03)	.04
Divorced/widowed	0.41 (–1.90 to 2.73)	.73
Education level (reference: elementary/secondary)		
High school	–2.74 (–6.41 to 0.94)	.14
University/college	–2.65 (–6.19 to 0.89)	.14
Postgraduate	–2.39 (–6.05 to 1.26)	.20
Occupation (reference: employed)		
Work from home	1.33 (0.24 to 2.43)	.02
Student	1.76 (0.19 to 3.32)	.03
Unemployed	3.34 (1.68 to 5.01)	<.001
Other	2.11 (–0.15 to 4.37)	.07
Chronic disease (reference: no)		
Yes	1.32 (0.02 to 2.63)	.047
Current situation (reference: social distancing)		
Quarantine/isolation	0.62 (–0.91 to 2.15)	.43
Average income per month (million VND [a]) (reference: no income)		
<1	0.16 (–2.46 to 2.78)	.90
1-5	–0.54 (–2.09 to 1.00)	.49
5-10	–0.15 (–1.64 to 1.35)	.84
10-20	–0.79 (–2.40 to 0.82)	.34
>20	–0.08 (–1.85 to 1.70)	.93
Average length of stay (hours) at home during the lockdown (reference: 0-10)		
11-20	–0.65 (–1.94 to 0.64)	.33
21-24	–0.38 (–1.8 to 1.05)	.60
Current health status (reference: good/very good)		
Average	5.85 (4.49 to 7.21)	<.001
Bad/very bad	9.57 (4.54 to 14.59)	<.001
Infected cases in province/city (reference: no)		
Yes	1.69 (0.83 to 2.55)	<.001

[a] VND: Vietnam Dong.

The association between respondents' isolation situation and current health self-assessment with symptoms of anxiety remained associated in the multivariate analysis. Isolated participants (at quarantine centers or at home) had a higher likelihood of experiencing anxiety (B=1.22, 95% CI 0.11 to 2.33). The respondents who rated their health as average (B=3.64, 95% CI 2.64 to 4.63) and bad or very bad (B=7.24, 95% CI 3.55 to 10.90) were more likely to exhibit anxiety than those who reported good or very good health (Multimedia Appendix 4).

Occupation, physical health condition, and COVID-19 cases in place of residence were factors associated with symptoms of stress among respondents. Being unemployed or a student was associated with more stress than employed individuals (B=2.34, 95% CI 0.84 to 3.85 vs B=1.17, 95% CI 0.10 to 2.23). Participants with average (B=5.19, 95% CI 3.83 to 6.56) and very bad or bad (B=10.60, 95% CI 5.56 to 15.65) health evaluations demonstrated a higher risk for stress than those with a good or very good self-reported health status. A higher stress score was recorded among respondents who were living in areas with COVID-19 cases (B=0.92, 95% CI 0.08 to 1.76) (Multimedia Appendix 5).

The results for univariate linear regression for the IES-R and DASS-21 scales and sociodemographic factors are presented in Multimedia Appendix 6.

We found that 96.4% (n=1335) of Vietnamese residents were confident of their doctor's expertise on COVID-19 diagnosis and treatment. Most believed that they were at a low likelihood of contracting COVID-19 (n=841, 63.3%) and perceived that they could survive if they were infected (n=1189, 85.9%). More than half expressed concerns for their children who were under the age of 18 years (n=755, 54.5%) as well as for other family members (n=892, 64.5%). In the multivariate regression analysis, the likelihood of surviving if infected remained statistically associated with the DASS-21 depression and anxiety subscales. Having concerns about children aged <18 years were associated with IES-R score and DASS-21 stress. Respondents who held a particular belief about their likelihood of surviving if infected (ie, somewhat likely, not very likely, and not at all likely) experienced more depression than those who did not know their likelihood (B=2.07, 95% CI 0.59 to 3.55, B=4.10, 95% CI 1.47 to 6.72 vs B=8.67, 95% CI 0.14 to 17.19). In addition, those who believed that they were somewhat likely to survive if infected with COVID-19 were more anxious than those who did not know their likelihood (B=1.21, 95% CI 0.15 to 2.27). Those who were worried about their children aged <18 years had increased IES-R and DASS-21 stress scores than those without children aged <18 years (B=7.81, 95% CI 4.98 to 10.64 vs B=1.75, 95% CI 0.27 to 3.24). Participants who were somewhat worried about children aged <18 years also experienced greater psychological distress (B=3.19, 95% CI 0.61, 5.78). The results for the analysis of the association between concerns related to the COVID-19 pandemic and psychological distress, depression, anxiety, and stress are shown in Multimedia Appendices 7 and 8.

Discussion

Principal Findings

Our study investigated in detail the psychological and mental health impacts of the COVID-19 pandemic on the general population in Vietnam during the first nationwide lockdown. The findings revealed that more than one-third of respondents experienced psychological distress (35.9%), nearly one-quarter exhibited depression and stress (23.5% and 22.3%, respectively), and 14.1% experienced anxiety. The prevalence of the depressed population in the present study was higher than that reported by Nguyen et al [10]; their study investigated outpatients in hospitals and health centers in Vietnam from February 14 to March 13, 2020, and found that 7.4% of the sample exhibited symptoms of depression. There may be some explanations for this discrepancy. First, in February, Vietnam had not implemented compulsory social distancing at the national level while strictly large-scale social distancing substantially increased anxiety, depression, loneliness, and substance use [24]. Second, our questionnaire was distributed via social media platforms, which might involve people who were more concerned about the pandemic and thus there might be a high probability of exposure to misinformation and fake news that could elevate depression rates [25]. Moreover, different measurement scales could be attributed to different prevalence.

Our research findings also suggested factors associated with psychological and mental health impact, which may identify vulnerable populations, such as unemployed individuals, those with a bad or very bad self-reported physical health status, or those living in areas with COVID-19 cases.

Our results confirmed previous research, which found that employment could be a protective factor for mental health during the pandemic [26]. Additionally, participants working from home were more prone to depression than those who returned to their workplaces. This could be due to reduced social interactions and more distractions from pandemic-related news among people working from home, which might increase their depression. Unemployed people were deemed to have the worst psychological and mental health status among occupations in our study, experiencing distress, depression, and anxiety. Indeed, the Vietnamese government has implemented some interventions for this vulnerable population, such as an interim relief package (63 trillion VND) for specific people during the pandemic (eg, furloughed workers, unpaid leave workers, freelancers, the poor, or near-poor individuals, etc) [4]. These short-term interventions could partly mitigate socioeconomic burden. However, to ensure socioeconomic stability during future pandemics, long-term interventions, such as strengthening health care systems and having social protection schemes, should be taken into consideration [4].

Previous studies have shown that poor health status was a negative factor for mental health during the pandemic [10,13,27], which was also seen in our study. Notably speaking, residents in provinces or cities where there were COVID-19 cases were more sensitive to depression and stress. This increase in depression could be influenced by the fear of contracting an infection by living in higher-risk areas. Shi et al [26] showed that people living in higher-risk areas (Hubei) significantly experienced depression, anxiety, insomnia, and acute stress than those living in other areas. This may suggest that public health interventions should be focused on these vulnerable areas.

Nguyen et al [10] revealed that married individuals expressed a lower HRQOL during the COVID-19 pandemic, implying that married participants experienced psychological distress during the COVID-19 lockdown [10]; this was also seen in our study. However, married individuals tended to have lower depression than single individuals during this period. This could be explained by the strong association between loneliness and depression during the lockdown, where married people were less lonely than single people [28].

In terms of concerns related to COVID-19, the majority of participants (n=1189, 85.8%) perceived a high likelihood of surviving if they were infected with COVID-19. This was consistent with a prior study [13]. This can be inferred from their belief in their doctors' capacity to treat COVID-19; in fact, more than 95% (n=1335) of Vietnamese citizens were somewhat or very confident about their health care professionals. This can lead to a decrease in the public's distress level during the pandemic. Another positive finding observed was that those who perceived a reduced likelihood of contracting COVID-19 exhibited less distress during the lockdown. These findings point to the benefits of using contact tracing apps, such as Bluezone [29]. Not only does this intervention effectively support the epidemiological investigation, but it also reduces anxiety among those who were more likely to get exposed to COVID-19 (eg, taxi drivers).

With regards to concerns about family members, people with children younger than 18 years worried more and experience heightened distress, anxiety, and stress. This reflects Vietnamese culture as citizens often live in a three-generation household. Therefore, community interventions that aim to enhance health awareness in order to reduce risk and promote resilience among vulnerable populations are needed. These actions can also help strengthen our long-term response to the pandemic and prepare more effectively for future public health emergencies [30,31]. It suggests that providing reliable and accurate health information through health education or consulting programs, targeted at protecting children and the elderly, could help decrease the mental health impact of the lockdown on family members [32,33].

Strengths and Limitations

This study had some strengths. First, it provided comprehensive information on the prevalence of psychological and mental health impacts of COVID-19 during the first nationwide lockdown. It used IES-R to measure the public crisis impact of the lockdown, and DASS-21 to capture the population's depression, anxiety, and stress levels. Both scales complemented each other to provide a comprehensive picture of the psychological and mental health impact of the first nationwide lockdown in Vietnam. Second, we investigated the public's concerns on COVID-19, which played a useful role in determining which public health interventions might be implemented to mitigate the impact. Third, areas with COVID-19 infections were also pointed out in our findings as vulnerable areas, where citizens may be more likely to experience mental health issues. This suggests that more action plans should be considered in these areas, which could control the disease effectively and reduce psychological impacts.

Several limitations existed in the study as well. First, the online survey with convenience sampling might limit the representativeness of the study. Second, measurement tools were self-assessed by participants, which might possess social bias. Third, the study design was cross-sectional, which reflects the population at the study time point and limits a causal relationship. However, in the context of nationwide social distancing, which restricted face-to-face interviews, and the need to carry out data collection in a short time period, these limitations were inevitable. Regardless, further studies are needed to assess possible long-term psychological and mental health impacts of COVID-19 on the population. In addition, as social distancing during the pandemic has been implemented in many countries and the use of web-based tools to measure psychological impacts has been increasing [34,35], more studies in this field are needed to evaluate these tools.

Acknowledgments

We would like to express our appreciation to all the participants in our study.

Authors' Contributions

KNCD, PTLN, TNLB, TLP, TVV, APG, LA, and BVV were involved in study conceptualization; KNCD, PTLN, TNLB, TLP, LA, and BVV aided with methodology; KNCD, PTLN, and TNLB with the software; KNCD, PTLN, TNLB, TLP, LA, and BVV with validation; KNCD, PTLN, TNLB, TVV, and APG with the formal analysis; KNCD, PTLN, TNLB, TLP, TVV, and BVV with the investigation; KNCD, PTLN, TNLB, TLP, TVV, APG, and BVV with data curation; KNCD, PTLN, and TNLB with writing the original draft; KNCD, PTLN, TNLB, TLP, TVV, APG, LA, and BVV with reviewing and editing; BVV with supervision; and KNCD, PTLN, TNLB, and TLP with project administration. All authors have read and approved the version of the manuscript to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Checklist for Reporting Results of Internet E-Surveys (CHERRIES).

[DOC File , 66 KB - formative_v4i12e24776_app1.doc]

Multimedia Appendix 2

Sociodemographic characteristics of survey respondents.

[DOC File , 62 KB - formative_v4i12e24776_app2.doc]

Multimedia Appendix 3
Prevalence and score of the psychological and mental health impacts of COVID-19 on the general population in Vietnam during the first lockdown.
[DOC File , 31 KB - formative_v4i12e24776_app3.doc]

Multimedia Appendix 4
Multivariate linear regression results for the Depression, Anxiety, and Stress Scale-21 (anxiety subscale) with sociodemographic covariates.
[DOC File , 35 KB - formative_v4i12e24776_app4.doc]

Multimedia Appendix 5
Multivariate linear regression results for the Depression, Anxiety, and Stress Scale-21 (stress subscale) with sociodemographic covariates.
[DOC File , 36 KB - formative_v4i12e24776_app5.doc]

Multimedia Appendix 6
Univariate linear regression results for the Impact of Event Scale-Revised and the Depression, Anxiety, and Stress Scale-21 with sociodemographic covariates.
[DOC File , 90 KB - formative_v4i12e24776_app6.doc]

Multimedia Appendix 7
Univariate linear regression results for the Impact of Event Scale-Revised and the Depression, Anxiety, and Stress Scale-21 with concern-related covariates.
[DOC File , 71 KB - formative_v4i12e24776_app7.doc]

Multimedia Appendix 8
Multivariate linear regression results for the Impact of Event Scale-Revised and the Depression, Anxiety, and Stress Scale-21 with concern-related covariates.
[DOC File , 66 KB - formative_v4i12e24776_app8.doc]

References

1. Timeline of WHO’s response to COVID-19. World Health Organization. 2020 Jun 29. URL: https://www.who.int/news-room/detail/29-06-2020-covidtimeline [accessed 2020-07-21]
2. WHO Coronavirus Disease (COVID-19) Dashboard. World Health Organization. 2020. URL: https://covid19.who.int/ [accessed 2020-07-21]
3. Phan LT, Nguyen TV, Luong QC, Nguyen TV, Nguyen HT, Le HQ, et al. Importation and Human-to-Human Transmission of a Novel Coronavirus in Vietnam. N Engl J Med 2020 Feb 27;382(9):872-874. [doi: 10.1056/nejmc2001272]
4. Tran PB, Hensing G, Wingfield T, Atkins S, Sidney Annerstedt K, Kazibwe J, et al. Income security during public health emergencies: the COVID-19 poverty trap in Vietnam. BMJ Glob Health 2020 Jun 15;5(6):e002504 [FREE Full text] [doi: 10.1136/bmjgh-2020-002504] [Medline: 32540965]
5. COVID-19 databoard. Ministry of Health in Vietnam. 2020. URL: https://ncov.vncdc.gov.vn/pages/viet-nam-1.html?fbclid=IwAR3q2ve1JSS9pKSsOZAEGg_ytCkMO0skq0X4gABnJtrF0psSnrWuXmp_lqA [accessed 2020-07-21]
6. Nguyen TH, Vu DC. The largest epicenter of the coronavirus outbreak in Vietnam. Infect Control Hosp Epidemiol 2020 Aug 13;41(8):984-985 [FREE Full text] [doi: 10.1017/ice.2020.128] [Medline: 32279697]
7. Cong bo dich covid 19 toan quoc. Ministry of Health in Vietnam. 2020. URL: https://ncov.moh.gov.vn/web/guest/-/cong-bo-dich-covid-19-toan-quoc [accessed 2020-07-21]
8. Dinh L, Dinh P, Nguyen P, Nguyen D, Hoang T. Vietnam's response to COVID-19: prompt and proactive actions. J Travel Med 2020 May 18;27(3) [FREE Full text] [doi: 10.1093/jtm/taaa047] [Medline: 32297929]
9. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. The Lancet 2020 Mar;395(10227):912-920. [doi: 10.1016/s0140-6736(20)30460-8]
10. Nguyen HC, Nguyen MH, Do BN, Tran CQ, Nguyen TTP, Pham KM, et al. People with Suspected COVID-19 Symptoms Were More Likely Depressed and Had Lower Health-Related Quality of Life: The Potential Benefit of Health Literacy. J Clin Med 2020 Mar 31;9(4):965 [FREE Full text] [doi: 10.3390/jcm9040965] [Medline: 32244415]

11. Weiss D, Marmar C. The Impact of Event Scale-Revised. In: Wilson JP, Keane TM, editors. Assessing Psychological Trauma and PTSD. New York: Guilford Press; 1997:399-411.
12. Lovibond P, Lovibond S. The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behaviour Research and Therapy 1995 Mar;33(3):335-343. [doi: 10.1016/0005-7967(94)00075-u]
13. 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]
14. Creamer M, Bell R, Failla S. Psychometric properties of the Impact of Event Scale - Revised. Behav Res Ther 2003 Dec;41(12):1489-1496. [doi: 10.1016/j.brat.2003.07.010] [Medline: 14705607]
15. Dao TK, Poritz JMP, Moody RP, Szeto K. Development, reliability, and validity of the Posttraumatic Stress Disorder Interview for Vietnamese refugees: a diagnostic instrument for Vietnamese refugees. J Trauma Stress 2012 Aug 20;25(4):440-445. [doi: 10.1002/jts.21712] [Medline: 22821587]
16. Hendrix CC, Jurich AP, Schumm WR. Validation of the Impact of Event Scale on a sample of American Vietnam veterans. Psychol Rep 1994 Aug 31;75(1 Pt 1):321-322. [doi: 10.2466/pr0.1994.75.1.321] [Medline: 7984745]
17. Asukai N, Kato H, Kawamura N, Kim Y, Yamamoto K, Kishimoto J, et al. Reliability and validity of the Japanese-language version of the impact of event scale-revised (IES-R-J): four studies of different traumatic events. J Nerv Ment Dis 2002 Mar;190(3):175-182. [doi: 10.1097/00005053-200203000-00006] [Medline: 11923652]
18. Le TA, Le MQT, 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]
19. Le MTH, 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. [doi: 10.1371/journal.pone.0180557] [Medline: 28723909]
20. 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(1). [doi: 10.1186/1471-244x-13-24]
21. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. Am J Epidemiol 1993 Dec 01;138(11):923-936. [doi: 10.1093/oxfordjournals.aje.a116813] [Medline: 8256780]
22. R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL: https://www.r-project.org/ [accessed 2020-12-11]
23. 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]
24. Galea S, Merchant RM, Lurie N. The Mental Health Consequences of COVID-19 and Physical Distancing: The Need for Prevention and Early Intervention. JAMA Intern Med 2020 Jun 01;180(6):817-818. [doi: 10.1001/jamainternmed.2020.1562] [Medline: 32275292]
25. Nguyen THD, Vu DC. Impacts of the COVID-19 pandemic upon mental health: Perspectives from Vietnam. Psychol Trauma 2020 Jul;12(5):480-481. [doi: 10.1037/tra0000694] [Medline: 32525376]
26. Shi L, Lu Z, Que J, Huang X, Liu L, Ran M, et al. Prevalence of and Risk Factors Associated With Mental Health Symptoms Among the General Population in China During the Coronavirus Disease 2019 Pandemic. JAMA Netw Open 2020 Jul 01;3(7):e2014053 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.14053] [Medline: 32609353]
27. Liu D, Ren Y, Yan F, Li Y, Xu X, Yu X, et al. Psychological Impact and Predisposing Factors of the Coronavirus Disease 2019 (COVID-19) Pandemic on General Public in China. SSRN Journal 2020:1-16. [doi: 10.2139/ssrn.3551415]
28. Loades ME, Chatburn E, Higson-Sweeney N, Reynolds S, Shafran R, Brigden A, et al. Rapid Systematic Review: The Impact of Social Isolation and Loneliness on the Mental Health of Children and Adolescents in the Context of COVID-19. J Am Acad Child Adolesc Psychiatry 2020 Nov;59(11):1218-1239.e3 [FREE Full text] [doi: 10.1016/j.jaac.2020.05.009] [Medline: 32504808]
29. Remote platform, Bluezone app launched in support of medical treatment, including COVI-19. Vietnamse Government. 2020. URL: http://news.chinhphu.vn/Home/Remote-platform-Bluezone-app-launched-in-support-of-medical-treatment-including-COVI19/20204/39790.vgp [accessed 2020-09-12]
30. 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]
31. Liao Q, Yuan J, Dong M, Yang L, Fielding R, Lam WWT. Public Engagement and Government Responsiveness in the Communications About COVID-19 During the Early Epidemic Stage in China: Infodemiology Study on Social Media Data. J Med Internet Res 2020 May 26;22(5):e18796 [FREE Full text] [doi: 10.2196/18796] [Medline: 32412414]

32. Alsan M, Stantcheva S, Yang D, Cutler D. Disparities in Coronavirus 2019 Reported Incidence, Knowledge, and Behavior Among US Adults. JAMA Netw Open 2020 Jun 01;3(6):e2012403 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.12403] [Medline: 32556260]
33. 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]
34. Ni MY, Yang L, Leung CMC, Li N, Yao XI, Wang Y, et al. Mental Health, Risk Factors, and Social Media Use During the COVID-19 Epidemic and Cordon Sanitaire Among the Community and Health Professionals in Wuhan, China: Cross-Sectional Survey. JMIR Ment Health 2020 May 12;7(5):e19009 [FREE Full text] [doi: 10.2196/19009] [Medline: 32365044]
35. van Agteren J, Bartholomaeus J, Fassnacht D, Iasiello M, Ali K, Lo L, et al. Using Internet-Based Psychological Measurement to Capture the Deteriorating Community Mental Health Profile During COVID-19: Observational Study. JMIR Ment Health 2020 Jun 11;7(6):e20696 [FREE Full text] [doi: 10.2196/20696] [Medline: 32490845]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
DASS-21: Depression, Anxiety, and Stress Scale-21
HRQOL: health-related quality of life
IES-R: Impact of Event Scale-Revised

Edited by J Torous; submitted 07.10.20; peer-reviewed by A Dominguez-Rodriguez, T Hendriks; comments to author 14.11.20; revised version received 29.11.20; accepted 03.12.20; published 15.12.20.

Please cite as:
Ngoc Cong Duong K, Nguyen Le Bao T, Thi Lan Nguyen P, Vo Van T, Phung Lam T, Pham Gia A, Anuratpanich L, Vo Van B
Psychological Impacts of COVID-19 During the First Nationwide Lockdown in Vietnam: Web-Based, Cross-Sectional Survey Study
JMIR Form Res 2020;4(12):e24776
URL: http://formative.jmir.org/2020/12/e24776/
doi:10.2196/24776
PMID:33284778

Original Paper

Factors Influencing Patients’ Initial Decisions Regarding Telepsychiatry Participation During the COVID-19 Pandemic: Telephone-Based Survey

Jennifer Severe[1], MD; Ruiqi Tang[2], BA; Faith Horbatch[2], BSc; Regina Onishchenko[2], BSc; Vidisha Naini[2], BSc; Mary Carol Blazek[1], MD

[1]Department of Psychiatry, University of Michigan, Ann Arbor, MI, United States
[2]University of Michigan Medical School, Ann Arbor, MI, United States

Corresponding Author:
Jennifer Severe, MD
Department of Psychiatry
University of Michigan
4250 Plymouth Road
Rachel Upjohn Building, Floor 1, Door 1
Ann Arbor, MI, 48109
United States
Phone: 1 617 416 8275
Email: severej@med.umich.edu

Related Article:

This is a corrected version. See correction statement: http://formative.jmir.org/2021/1/e27357/

Abstract

Background: Telepsychiatry enables patients to establish or maintain psychiatric care during the COVID-19 pandemic. Little is known about the factors influencing patients’ initial decisions to participate in telepsychiatry in the midst of a public health crisis.

Objective: This paper seeks to examine factors influencing patients’ initial decisions to accept or decline telepsychiatry immediately after the stay-at-home order in Michigan, their initial choice of virtual care modality (video or telephone), and their anticipated participation in telepsychiatry once clinics reopen for in-person visits.

Methods: Between June and August 2020, we conducted a telephone-based survey using a questionnaire comprising 14 quantitative and two qualitative items as part of a quality improvement initiative. We targeted patients who had an in-person appointment date that fell in the first few weeks following the Michigan governor’s stay-at-home order, necessitating conversion to virtual visits or deferment of in-person care. We used descriptive statistics to report individual survey responses and assess the association between chosen visit type and patient characteristics and future participation in telepsychiatry using multivariable logistic regression.

Results: A total of 244 patients whose original in-person appointments were scheduled within the first 3 weeks of the stay-at-home order in Michigan completed the telephone survey. The majority of the 244 respondents (n=202, 82.8%) initially chose to receive psychiatric care through video visits, while 13.5% (n=33) chose telephone visits and 1.2% (n=3) decided to postpone care until in-person visit availability. Patient age correlated with chosen visit type (*P*<.001; 95% CI 0.02-0.06). Patients aged ≥44 years were more likely than patients aged 0-44 years to opt for telephone visits (relative risk reduction [RRR] 1.2; 95% CI 1.06-1.35). Patient sex (*P*=.99), race (*P*=.06), type of insurance (*P*=.08), and number of previous visits to the clinic (*P*=.63) were not statistically relevant. Half of the respondents (132/244, 54.1%) stated theywere likely to continue with telepsychiatry even after in-person visits were made available. Telephone visit users were less likely than video visit users to anticipate future participation in telepsychiatry (RRR 1.08; 95% CI 0.97-1.2). Overall, virtual visits met or exceeded expectations for the majority of users.

Conclusions: In this cohort, patient age correlates with the choice of virtual visit type, with older adults more likely to choose telephone visits over video visits. Understanding challenges to patient-facing technologies can help advance health equity and guide best practices for engaging patients and families through telehealth.

(JMIR Form Res 2020;4(12):e25469) doi:10.2196/25469

KEYWORDS

telepsychiatry; COVID-19; video visit; telephone visit; telehealth; mental health; United States; decision making; virtual care

Introduction

With the outbreak of COVID-19, many health care centers were forced to quickly modify their protocols and accept new modes of health care delivery. In-person routines and nonessential services were cancelled in an effort to protect both patients and frontline health care providers. New telehealth programs were swiftly implemented, and existing programs were expanded to provide ongoing access to care while simultaneously promoting physical distancing [1-3]. Federal and state governments as well as commercial insurers rapidly eased telehealth regulations [4]. The percentage of psychiatrists seeing more than three-quarters of their patient caseload via telehealth reached an unprecedented magnitude during the pandemic (85%) compared to before the pandemic (2.1%) [5]. This accelerated uptake contrasts starkly with the slow rate of engagement noted since the inception of telepsychiatry in 1956 [6-8].

For decades, telepsychiatry has enabled clinicians to deliver an array of mental health services in real time over virtual platforms. It provides health care access to a wider variety of populations residing in areas with limited community clinics, reduces the burden of disease, enhances models of integrated care, eliminates travel costs and logistics, and remains comparable in efficacy to conventional in-person mental health treatment [8,9]. Until recently, the uniform adoption of telepsychiatry by patients and families was limited by several barriers, including technology use, provider-patient relationships, and a lack of awareness of its existence [10].

Despite the rapid expanse of telepsychiatry during the COVID-19 pandemic, for most people, telepsychiatry was a novel and unfamiliar way of engaging in care [11]. Little is known about the factors influencing patients' decisions to participate in telepsychiatry in the midst of a public health crisis. The aim of this paper is to examine the factors that influenced patients' initial decisions to accept or decline telepsychiatry immediately after the issuance of the stay-at-home order in the US state of Michigan on March 23, 2020 [12], including choice of virtual care modality (video or telephone) and anticipated participation in telepsychiatry once clinics reopen for in-person visits. The findings will help provide actionable insights for further engagement of patients and families in telepsychiatry services.

Methods

Survey Setting and Participants

The Outpatient Psychiatry Clinics at the University of Michigan health care system, known as Michigan Medicine, provide approximately 67,000 psychiatry outpatient visits per year. In September 2019, the Outpatient Psychiatry Clinics launched a pilot telepsychiatry program with six providers; however, it failed to gain traction. By February 2020, only 26 virtual visits had been conducted via the Epic electronic medical record (EMR)–integrated platform. On March 23, 2020, in response to the COVID-19 pandemic, Governor Gretchen Whitmer issued a stay-at-home order (Executive Order 2020-21) [12] and Michigan Medicine called for the closure of all nonurgent ambulatory care clinics. Patients were offered a choice of video or telephone visits as well as the option to wait for an in-person visit when the clinics reopened. By June 2020, the Outpatient Psychiatry Clinics had converted 100% of in-person visits to virtual care, and the Psychiatry Department had become the lead department in telehealth for Michigan Medicine, followed by Neurology (62%) and Family Medicine (45%).

In April 2020, a multidisciplinary team from the Psychiatry Department and the Office of Patient Experience designed a telephone-based survey as part of a quality improvement (QI) project to evaluate the rapid scaleup of telepsychiatry and better understand patient experiences. The survey data were obtained as part of a strict QI initiative and as such did not require Institutional Review Board oversight. The telephone survey was conducted between June and August 2020 and targeted patients who had an in-person appointment date that fell in the first few weeks following the Michigan governor's stay-at-home edict, necessitating conversion to virtual visits or deferment of in-person care. A total of 1030 patients who had an appointment scheduled between March 23 and April 13, 2020, were called, and 431 patients answered the call. Among these, 65 declined to participate in the survey, 113 asked to be recontacted at a more convenient time, and 9 were wrong numbers. A total of 244 patients participated in and completed the survey. This accounts for a response rate of 56.6% (231/244), which matches the overall response rates of Michigan Medicine (~50%) and the state (47.3%) for telephone surveys [13].

Recruitment

The choice of telephone to contact participants offered a valuable opportunity to reach patients who lack the technological access required for a web-based or email survey. Seven volunteer interviewers, four medical students and three undergraduate students, completed a web-based Health Insurance Portability and Accountability Act (HIPAA) training module. They were instructed to follow Office of Patient Experience patient telephone call guidelines, which capped the maximum number of calls to each patient at two. They received a script to read at the beginning of each telephone call to give respondents the opportunity to make an informed decision and accept or decline the survey. Parents or proxies were interviewed on behalf of patients who were <11 years of age. In addition, adults were given the option to have a proxy answer the questionnaire on their behalf.

Telephone-Based Questionnaire

The work group generated a 16-item questionnaire comprising 14 quantitative and 2 qualitative items, including the participant's initial choice of visit type after the stay-at-home order, factors influencing this choice, digital platform use, experience with digital technology, future participation in telepsychiatry when clinics reopen for in-person visits, and readiness to resume in-person visits during the pandemic. We defined telepsychiatry as synchronous live mental health services delivered via video or telephone by a mental health clinician, such as a physician, nurse practitioner, psychologist, or social worker. An electronic version of the questionnaire was created using Qualtrics survey software. We adopted a quantitative approach by creating multiple-choice survey questions with predetermined answer options as well as Likert scale questions. Two open-ended questions were included to gather elaborated comments about respondents' overall expectations of virtual visits and to welcome any additional reflections at the end of the survey. The participants' demographics and number of previous clinic visits were extracted from the Epic EMR. The survey guide is included in Multimedia Appendix 1.

Analysis

We used descriptive statistics to report individual survey responses and their relative percentages. We examined the association between chosen visit type and patient characteristics (age, sex, race, type of insurance, and number of previous visits to the clinic) and future virtual care participation using multivariable logistic regression through JMP 15 software (SAS Institute). Qualitative analysis of the free comments (questions 10 and 16) was not performed to allow for timely dissemination of our findings; however, narrative summaries of the comments are provided in the Results and Discussion sections.

Results

Survey Participants

Survey data were collected for a total of 244 patients whose original in-person appointments were scheduled in the first three weeks following the COVID-19 Michigan stay-at-home order. The majority of the 244 survey participants (n=212, 86.9%) were patients themselves; the remainder of respondents (n=32, 13.1%) were parents or proxies. In this sample of 244 respondents, 149 (61.1%) were adults aged 18-64 years, 49 (20.1%) were minors aged <18 years, and 45 (18.4%) were older people aged ≥65 years. Of the 244 respondents, 68.4% (n=167) identified themselves as female, 77.5% (n=189) as White, 10.7% (n=26) as Black, and 4.5% (n=11) as Asian. According to medical records, most survey respondents were established patients before the clinic closures who had attended 1-6 previous appointments (133/244, 54.5%) or >6 appointments (n=82, 33.6%) between July 1, 2019, and March 22, 2020. A total of 20/244 patients (8.2%) had no prior appointment with the clinic. Managed Medicaid and Medicare patients each represented 38 (15.6%) of the 244 visitors surveyed, 92 (37.7%) were insured with Blue Care Network (ie, Premier Care for Michigan Medicine employees), and 67 (27.5%) were insured with Blue Cross Blue Shield. In Michigan, straight Medicaid patients receive their behavioral health care through Community Mental Health clinics as part of their contract and are not routinely served by Michigan Medicine Outpatient Psychiatric Clinics. The patient characteristics are included in Multimedia Appendix 2.

Telephone-Based Questionnaire

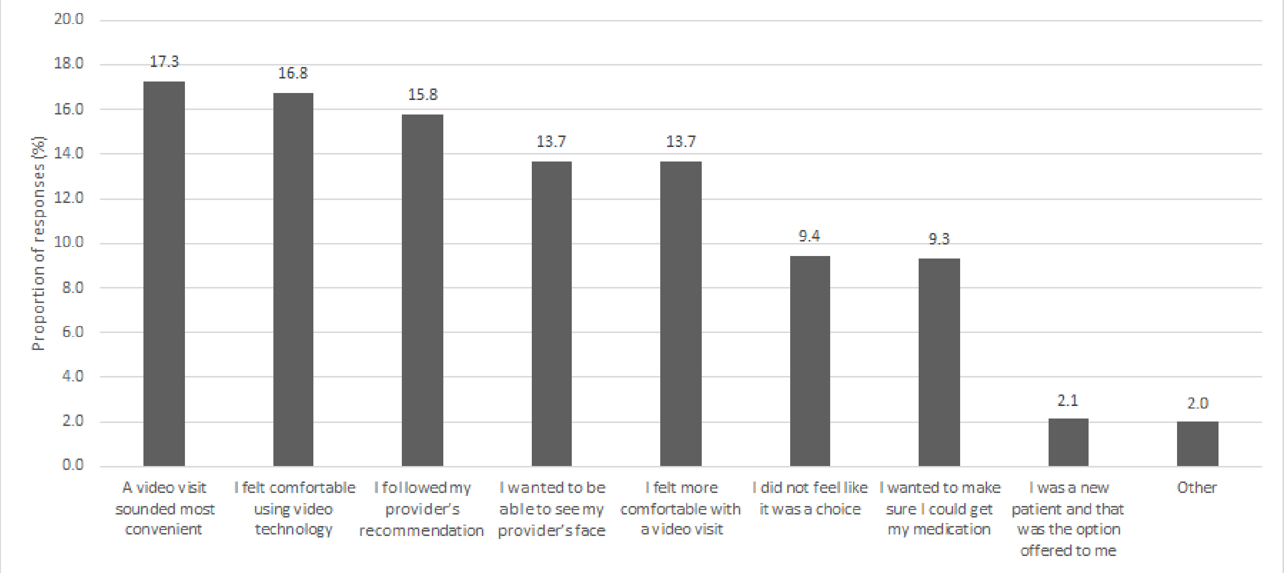
Participants' Initial Choice of Visit Type After the Stay-at-Home Order

Patients chose to receive psychiatric care through video visits with their provider (202/244, 82.8%), to receive psychiatric care through telephone visits (33/244, 13.5%), or to postpone care until in-person visit availability (3/244, 1.2%). Although some patients may have switched modalities for subsequent visits, survey respondents were asked to answer questions based on their experience with the initial modality they selected. Patient age correlated with initial choice in visit type (P=.002), and the magnitude of the association remained after adjustment for sex, race, type of insurance, and number of previous clinic visits (P<.001; 95% CI 0.02-0.06). Patients aged ≥44 years were more likely than patients aged 0-44 years to opt for telephone visits (relative risk ratio [RRR] 1.2; 95% CI 1.06-1.35). A slightly higher percentage of Asian respondents surveyed (10/11, 90.9%) chose to receive psychiatric care through video visits, followed by White respondents (159, 84.1%) and Black respondents (21/26, 80.7%); this difference was not statistically significant (P=.06). The patients' sex (P=.99), race (P=.06), type of insurance (P=.08), and number of previous clinic visits (P=.63), delineating new and pre-existing patients, were not statistically relevant.

Factors Influencing Initial Choice of Video Visits

Out of the 805 responses gathered, the main factors influencing the respondents' decision to attempt a video visit included finding video visits more convenient (n=139, 17%), feeling comfortable with the technology (n=135, 17%), and feeling comfortable with a video visit in general (n=110, 14%). The influencing factors indicated by the respondents and their proportions are shown in Figure 1.

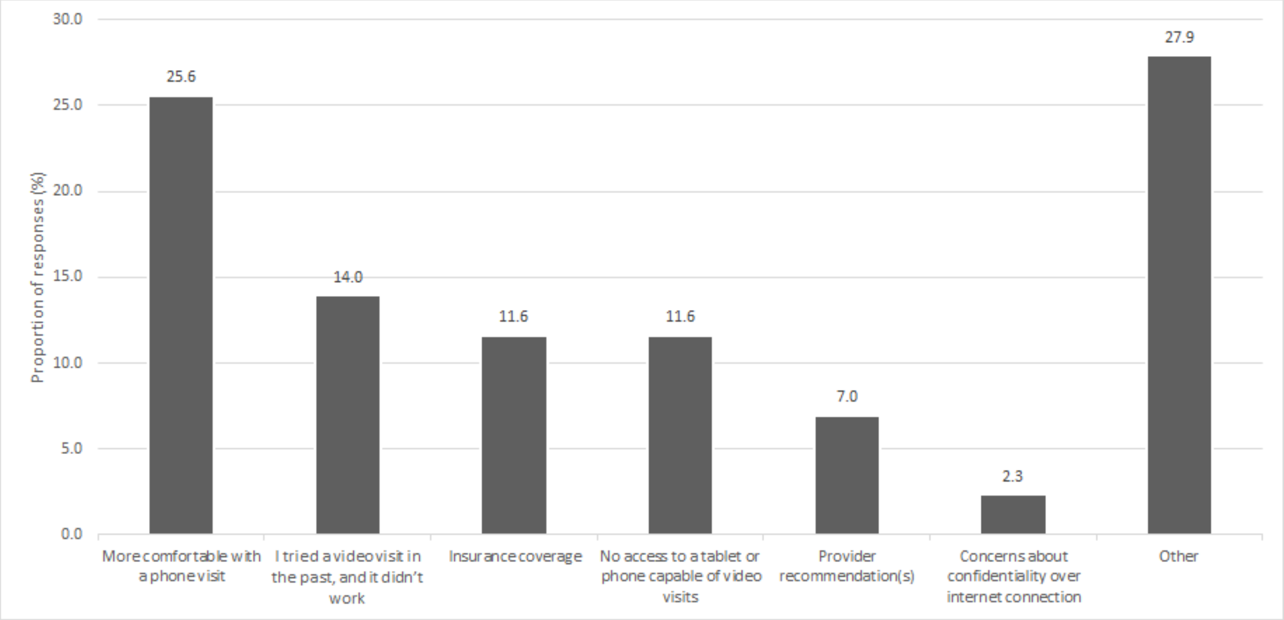
Figure 1. Factors influencing patients' decisions to receive care by video visit.



Factors Influencing Initial Choice of Telephone Visits

A total of 43 responses were collected. Most respondents who chose telephone visits indicated that they felt more comfortable with the telephone (n=11, 26%), had complications with video visits in the past (n=6, 14%), or did not have appropriate technology for video (n=5, 12%) The influencing factors indicated by the respondents and their proportions are shown in Figure 2.

Figure 2. Factors influencing patients' decisions to receive care by telephone visit.



Factors Influencing Initial Choice of In-Person Visits

The 3 patients who opted to wait for in-person visits indicated a preference for face-to-face visits, with 1 participant (33%) also reporting lack of comfort with the video visit technology.

Participant Experiences With Virtual Visits

Among the 235 patients who elected to receive psychiatric care virtually (either video or telephone visits), 220 (93.6%) reported that virtual visits met or exceeded expectations ("as expected," (n=126, 53.6%), "somewhat better than expected" (n=42, 17.9%), and "much better than expected" (n=52, 22.1%), with many participants indicating initial hesitation with virtual visits that were ultimately resolved with experience.

Video Visits

Video visits were generally conducted via a Michigan Medicine Patient Portal application (149/202, 73.7%), although several other platforms were used. Use of the video technology was regarded by the 202 respondents as being either "extremely easy" (n=115, 56.9%), "somewhat easy" (n=41, 20.3%), "neither easy nor difficult" (n=16, 7.9%), "somewhat difficult" (n=23, 11.4%), or "extremely difficult" (n=7, 3.5%). Parents of pediatric patients commented that they preferred the video platform for virtual care because it was more engaging than telephone visits.

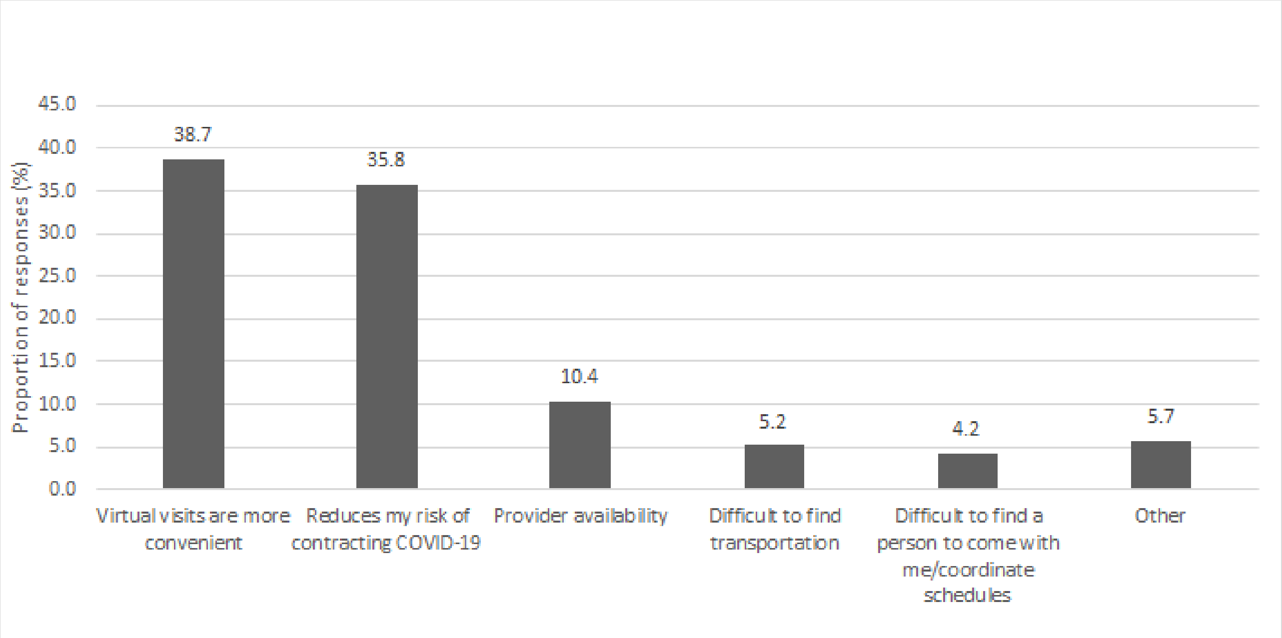
Telephone Visits

Some patients commented that telephone visits were particularly advantageous when their internet connection was too unstable to access video technology or when the video platforms were otherwise problematic.

Anticipation of Virtual Care Participation After Clinics Open for In-Person Visits

Approximately half of the respondents (132/244, 54.1%) stated that they were likely to continue with telepsychiatry visits even after in-person visits were made available. Patients who selected telephone visits were less likely than video visit users to anticipate participation in virtual care in the future (RRR 1.08; 95% CI 0.97-1.2). Factors influencing the decision to continue with virtual visits generated 235 answers and encompass convenience (n=149, 39%), reduced ability to contract COVID-19 (n=138, 36%) especially for those who reported having underlying medical conditions, and provider availability (n=40, 10%). Most patients who found the video technology extremely easy to use (115/202, 56.9%) indicated that they were either extremely likely (50/115, 43.5%) or somewhat likely (20/115, 17.4%) to continue with telepsychiatry visits after the clinics opened. Respondents who did not want to continue with virtual visits expressed a preference for face-to-face visits and a lack of comfort with the digital technology as the top two contributing factors. Parents of pediatric patients commented that their child had difficulty focusing and building a therapeutic relationship with their provider during virtual visits. The influencing factors indicated by the respondents and their proportions are shown in Figure 3.

Figure 3. Factors influencing anticipated participation in virtual care after clinics open for in-person visits.



Comfort With Returning to In-Person Visits During the COVID-19 Pandemic

All survey respondents were asked about their comfort with returning to in-person visits during the pandemic. Almost half of the patients (114/244, 46.7%) indicated that they were comfortable (ie, “extremely comfortable” or “somewhat comfortable”), while 105/244 (43.0%) were not. Among factors influencing patient comfort, perceived precautions by clinics to ensure COVID-19 protection accounted for 21% of the responses (182/885) followed by provider (140/885, 16%), state (105/885, 12%) or federal (71/885, 8%) recommendations regarding return to clinic, and provider availability (121/885, 14%). Several patients commented that the availability of a COVID-19 vaccine would influence their comfort with returning to in-person visits during the pandemic, while others wanted evidence of declining infection.

Discussion

In this paper, we sought to examine factors influencing patients’ initial decisions to accept or decline telepsychiatry and choices of virtual care modality (telephone or video) after the Michigan COVID-19 stay-at-home order. Our findings showed that 235 of the 244 respondents (96.3%) rapidly decided to receive psychiatric care virtually, and among those, more than three-quarters (202/235, 86.0%) opted for video visits as their initial choice. Of the demographic factors explored, patient age correlated significantly with the initial choice of type of visit, while sex, race, insurance type, and number of previous visits did not.

A vast body of literature prior to the pandemic shows that virtual care cohorts to be largely younger adults, commonly below mid-forties [14-16], with a preference for video-based telehealth [16-18]. We observed a similar trend in our sample, with the majority of patients selecting video visits were aged 0-44 years and the majority of patients opting for telephone visits being aged ≥45 years. As previously published and as illustrated in our survey, key challenges to older patients' willingness and readiness to partake in video visits encompass the capacity to use and access digital technology, low self-efficacy, and lack of support and facilitating conditions [17-19]. At Michigan Medicine, the GET Access (Geriatric Education for Telehealth Access) program serves as a model to address these challenges. Tailored virtual education successfully increased access to, comfort with, and participation in video visits for older adults.

One study showed that male patients were 1.6 times more likely than female patients with identical characteristics to use video than telephone calls for urgent care services [20]. Studies quantifying virtual care use and preference compared to traditional in-person care have yielded no consistent correlations with sex [14-16]. With the seriousness of the pandemic, preferences for a psychiatric visit type based on sex may have had less impact; this may also be true for race. Compared to other studies [16,21], our sample revealed no divide in selection of visit type with regard to patient race. This may stem from our cohort being predominantly White. Although health insurance type did not impact the initial choice of visit type, Medicare patients regardless of age are more likely to lack access to high-speed internet and smartphones with a data plan [22]. We advocate addressing these barriers with personalized support at different levels of use.

This survey was conducted at an unusual time, during a public health crisis, when many of the traditional barriers to scaling outpatient telehealth were significantly reduced. In a 2-year study conducted looking at primary care patients' choices between video, telephone, and in-person visits before the pandemic (N=1.1 million), there were only 14% scheduled telehealth visits, with the majority (93%) tallied as telephone visits [16]. The pandemic may have served as a catalyst for more rapid acceptance of virtual care out of necessity. It has also opened doors for a blended care model bridging conventional psychiatric face-to-face sessions with telepsychiatry visits as deemed appropriate.

Our results yielded high anticipation of continuous participation in telepsychiatry even after in-person visits become available. Convenience was the most commonly cited reason to continue virtual care, exceeding but followed closely by reduced ability to contract COVID-19. This enthusiasm should compel state, federal, and private entities to strongly consider ongoing funding and regulatory models to support telepsychiatry both during the ongoing pandemic and beyond. In states such as Colorado and New Hampshire, lawmakers have already passed bills permanently supporting all or parts of the telehealth expansions they adopted during the COVID-19 pandemic [23]. To our surprise, anticipation of future participation in telepsychiatry was lower among patients who selected telephone visits compared to video visits, although the telephone seems to offer an easier communication medium for health care delivery. With three-quarters of our video visits taking place on the Epic EMR-integrated platform, use of video visits is contingent upon comfort not only with the video technology but also with the patient portal system as part of a broader telehealth platform.

Our initiative had several limitations. First, our sample lacked diversity with regard to race and insurance type, which may have concealed possible nuances in choice of visit type. Second, the respondents were not stratified by ZIP code. Individuals living in lower socioeconomic status neighborhoods or rural areas may favor telephone visits over video visits due to limited access to appropriate devices and high-speed internet. Finally, we surveyed patients three months after their appointment change. Their perceptions of virtual visits may have changed over time.

As we look toward the future and apply lessons learned regarding telepsychiatry, we recommend that payments and regulations match patient needs and preferences. Understanding challenges to patient-facing technologies, including patient attitudes and perceptions, will help advance health equity and guide best practices for engaging patients and families through telehealth. Further studies evaluating disadvantaged populations, such as straight Medicaid beneficiaries, would further inform health policy and should be designed to capture patients who do not have access to the technology required for conventional web-based or email surveys.

Acknowledgments

The authors thank Jennifer Sullivan, Caitlin Hanna, and Darnysus Jackson II for their support and guidance and Daniele Littera, PhD for his contribution to the statistical analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Patient survey guide regarding virtual visits.
[DOCX File , 25 KB - formative_v4i12e25469_app1.docx]

Multimedia Appendix 2
Patient characteristics.
[DOCX File , 22 KB - formative_v4i12e25469_app2.docx]

References

1. Hong Y, Lawrence J, Williams D, Mainous I. Population-Level Interest and Telehealth Capacity of US Hospitals in Response to COVID-19: Cross-Sectional Analysis of Google Search and National Hospital Survey Data. JMIR Public Health Surveill 2020 Apr 07;6(2):e18961 [FREE Full text] [doi: 10.2196/18961] [Medline: 32250963]
2. Chen JA, Chung W, Young SK, Tuttle MC, Collins MB, Darghouth SL, et al. COVID-19 and telepsychiatry: Early outpatient experiences and implications for the future. Gen Hosp Psychiatry 2020;66:89-95 [FREE Full text] [doi: 10.1016/j.genhosppsych.2020.07.002] [Medline: 32750604]
3. Joshi AU, Lewiss RE, Aini M, Babula B, Henwood PC. Solving Community SARS-CoV-2 Testing With Telehealth: Development and Implementation for Screening, Evaluation and Testing. JMIR Mhealth Uhealth 2020 Oct 16;8(10):e20419 [FREE Full text] [doi: 10.2196/20419] [Medline: 33006942]
4. Hare N, Bansal P, Bajowala SS, Abramson SL, Chervinskiy S, Corriel R, et al. Work Group Report: COVID-19: Unmasking Telemedicine. J Allergy Clin Immunol Pract 2020 Sep;8(8):2461-2473.e3 [FREE Full text] [doi: 10.1016/j.jaip.2020.06.038] [Medline: 32603900]
5. Psychiatrists Use of Telepsychiatry During COVID-19 Public Health Emergency: Policy Recommendations. Washington, DC. Published online: June 2020. American Psychiatric Association. 2020 Jun. URL: https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/Telepsychiatry/APA-Telehealth-Survey-2020.pdf [accessed 2020-10-15]
6. Choi S, Wilcock AD, Busch AB, Huskamp HA, Uscher-Pines L, Shi Z, et al. Association of Characteristics of Psychiatrists With Use of Telemental Health Visits in the Medicare Population. JAMA Psychiatry 2019 Jun 01;76(6):654-657. [doi: 10.1001/jamapsychiatry.2019.0052] [Medline: 30892583]
7. Ramtekkar U, Bridge JA, Thomas G, Butter E, Reese J, Logan E, et al. Pediatric Telebehavioral Health: A Transformational Shift in Care Delivery in the Era of COVID-19. JMIR Ment Health 2020 Sep 18;7(9):e20157 [FREE Full text] [doi: 10.2196/20157] [Medline: 32525485]
8. Chakrabarti S. Usefulness of telepsychiatry: A critical evaluation of videoconferencing-based approaches. World J Psychiatry 2015 Sep 22;5(3):286-304 [FREE Full text] [doi: 10.5498/wjp.v5.i3.286] [Medline: 26425443]
9. Langarizadeh M, Tabatabaei M, Tavakol K, Naghipour M, Rostami A, Moghbeli F. Telemental Health Care, an Effective Alternative to Conventional Mental Care: a Systematic Review. Acta Inform Med 2017 Dec;25(4):240-246 [FREE Full text] [doi: 10.5455/aim.2017.25.240-246] [Medline: 29284913]
10. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: A systematic review. J Telemed Telecare 2018 Jan;24(1):4-12 [FREE Full text] [doi: 10.1177/1357633X16674087] [Medline: 29320966]
11. O'Brien M, McNicholas F. The use of telepsychiatry during COVID-19 and beyond. Ir J Psychol Med 2020 May 21:1-6 [FREE Full text] [doi: 10.1017/ipm.2020.54] [Medline: 32434596]
12. Executive Order 2020-21 (COVID-19). Office of Governor Gretchen Whitmer. 2020 Mar 24. URL: https://www.michigan.gov/whitmer/0,9309,7-387-90499_90705-522626--,00.html [accessed 2020-08-03]
13. Behavioral Risk Factor Surveillance System: Summary Data Quality Report. US Centers for Disease Control and Prevention. 2018 Jun 13. URL: https://www.cdc.gov/brfss/annual_data/2017/pdf/2017-sdqr-508.pdf [accessed 2020-10-01]
14. Gonzalez F, Cimadevila B, Garcia-Comesaña J, Cerqueiro S, Andion E, Prado J, et al. Telephone consultation in primary care. J Health Organ Manag 2018 Apr 09;32(2):321-337 [FREE Full text] [doi: 10.1108/JHOM-08-2017-0201] [Medline: 29624133]
15. McGrail KM, Ahuja MA, Leaver CA. Virtual Visits and Patient-Centered Care: Results of a Patient Survey and Observational Study. J Med Internet Res 2017 May 26;19(5):e177. [doi: 10.2196/jmir.7374] [Medline: 28550006]
16. Reed ME, Huang J, Graetz I, Lee C, Muelly E, Kennedy C, et al. Patient Characteristics Associated With Choosing a Telemedicine Visit vs Office Visit With the Same Primary Care Clinicians. JAMA Netw Open 2020 Jun 01;3(6):e205873. [doi: 10.1001/jamanetworkopen.2020.5873] [Medline: 32585018]
17. Uscher-Pines L, Sousa J, Raja P, Mehrotra A, Barnett ML, Huskamp HA. Suddenly Becoming a "Virtual Doctor": Experiences of Psychiatrists Transitioning to Telemedicine During the COVID-19 Pandemic. Psychiatr Serv 2020 Nov 01;71(11):1143-1150. [doi: 10.1176/appi.ps.202000250] [Medline: 32933411]
18. Cheung L, Leung TI, Ding VY, Wang JX, Norden J, Desai M, et al. Healthcare Service Utilization under a New Virtual Primary Care Delivery Model. Telemed J E Health 2019 Jul;25(7):551-559 [FREE Full text] [doi: 10.1089/tmj.2018.0145] [Medline: 30192211]
19. van Houwelingen CT, Ettema RG, Antonietti MG, Kort HS. Understanding Older People's Readiness for Receiving Telehealth: Mixed-Method Study. J Med Internet Res 2018 Apr 06;20(4):e123 [FREE Full text] [doi: 10.2196/jmir.8407] [Medline: 29625950]
20. Khairat S, Liu S, Zaman T, Edson B, Gianforcaro R. Factors Determining Patients' Choice Between Mobile Health and Telemedicine: Predictive Analytics Assessment. JMIR Mhealth Uhealth 2019 Jun 08;7(6):e13772 [FREE Full text] [doi: 10.2196/13772] [Medline: 31199332]
21. Campos-Castillo C, Anthony D. Racial and Ethnic Differences in Self-Reported Telehealth Use during the COVID-19 Pandemic: A Secondary Analysis of a U.S. Survey of Internet Users from Late March. J Am Med Inform Assoc 2020 Sep 07 [FREE Full text] [doi: 10.1093/jamia/ocaa221] [Medline: 32894772]

22. Roberts ET, Mehrotra A. Assessment of Disparities in Digital Access Among Medicare Beneficiaries and Implications for Telemedicine. JAMA Intern Med 2020 Oct 01;180(10):1386-1389. [doi: 10.1001/jamainternmed.2020.2666] [Medline: 32744601]
23. Pitsor J, Enlund S. States Turn to Telehealth During the Pandemic. National Conference of State Legislatures. 2020 Oct 29. URL: https://www.ncsl.org/research/health/states-turn-to-telehealth-during-the-pandemic-magazine2020.aspx [accessed 2020-11-01]

Abbreviations

EMR: electronic medical record
GET Access: Geriatric Education for Telehealth Access
HIPAA: Health Insurance Portability and Accountability Act
QI: quality improvement
RRR: relative risk reduction

Edited by G Eysenbach; submitted 03.11.20; peer-reviewed by T Jamieson, V Strotbaum; comments to author 24.11.20; revised version received 03.12.20; accepted 12.12.20; published 22.12.20.

Please cite as:
Severe J, Tang R, Horbatch F, Onishchenko R, Naini V, Blazek MC
Factors Influencing Patients' Initial Decisions Regarding Telepsychiatry Participation During the COVID-19 Pandemic: Telephone-Based Survey
JMIR Form Res 2020;4(12):e25469
URL: https://formative.jmir.org/2020/12/e25469
doi:10.2196/25469
PMID:33320823

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org