JMIR Formative Research

Impact Factor (2023): 2.0
Volume 5 (2021), Issue 3 ISSN 2561-326X Editor in Chief: Amaryllis Mavragani, PhDc

Contents

Original Papers

Viewpoints

Review

Corrigenda and Addenda

Original Paper

Delivering Perinatal Health Information via a Voice Interactive App (SMILE): Mixed Methods Feasibility Study

Lisa Militello[1*], MPH, PhD; Emre Sezgin[2*], PhD; Yungui Huang[2], PhD; Simon Lin[2], MBA, MD

[1]Martha S Pitzer Center for Women, Children & Youth, College of Nursing, The Ohio State University, Columbus, OH, United States

[2]Research Information Solutions and Innovation, The Abigail Wexner Research Institute, Nationwide Children's Hospital, Columbus, OH, United States

[*]these authors contributed equally

Corresponding Author:
Lisa Militello, MPH, PhD
Martha S Pitzer Center for Women, Children & Youth
College of Nursing
The Ohio State University
1585 Neil Ave
145 Newton Hall
Columbus, OH, 43210
United States
Phone: 1 3128135007
Email: militello.14@osu.edu

Abstract

Background: Perinatal health care is critically important for maternal health outcomes in infants. The United States fares considerably worse than comparable countries for maternal and infant mortality rates. As such, alternative models of care or engagement are warranted. Ubiquitous digital devices and increased use of digital health tools have the potential to extend the reach to women and infants in their everyday lives and make a positive impact on their health outcomes. As voice technology becomes more mainstream, research is prudent to establish evidence-based practice on how to best leverage voice technology to promote maternal-infant health.

Objective: The aim of this study is to assess the feasibility of using voice technology to support perinatal health and infant care practices.

Methods: Perinatal women were recruited from a large Midwest Children's Hospital via hospital email announcements and word of mouth. Owing to the technical aspects of the intervention, participants were required to speak English and use an iPhone. Demographics, patterns of technology use, and technology use specific to perinatal health or self-care practices were assessed at baseline. Next, participants were onboarded and asked to use the intervention, Self-Management Intervention–Life Essentials (SMILE), over the course of 2 weeks. SMILE provided users with perinatal health content delivered through mini podcasts (ranging from 3 to 8 minutes in duration). After each podcast, SMILE prompted users to provide immediate verbal feedback to the content. An exit interview was conducted with participants to gather feedback on the intervention and further explore participants' perceptions of voice technology as a means to support perinatal health in the future.

Results: In total, 19 pregnant women (17 to 36 weeks pregnant) were consented. Themes identified as important for perinatal health information include establishing routines, expected norms, and realistic expectations and providing key takeaways. Themes identified as important for voice interaction include customization and user preferences, privacy, family and friends, and context and convenience. Qualitative analysis suggested that perinatal health promotion content delivered by voice should be accurate and succinctly delivered and highlight key takeaways. Perinatal health interventions that use voice should provide users with the ability to customize the intervention but also provide opportunities to engage family members, particularly spouses. As a number of women multitasked while the intervention was being deployed, future interventions should leverage the convenience of voice technology while also balancing the influence of user context (eg, timing or ability to listen or talk versus nonvoice interaction with the system).

Conclusions: Our findings demonstrate the short-term feasibility of disseminating evidence-based perinatal support via podcasts and curate voice-captured data from perinatal women. However, key areas of improvement have been identified specifically for perinatal interventions leveraging voice technology. Findings contribute to future content, design, and delivery considerations of perinatal digital health interventions.

(JMIR Form Res 2021;5(3):e18240) doi:10.2196/18240

KEYWORDS

perinatal care; infant mortality; health education; mobile health; feasibility studies; family; mobile phone; webcasts as topic; user-computer interface

Introduction

Background

Nearly 60% of maternal deaths are preventable [1], and infant mortality rates are approximately 71% higher in the United States than in other comparable countries [2]. Race/ethnicity, low income, and chronic stress are associated with pregnancy-related complications and poor maternal and infant mortality statistics [3-6]. To set the stage for long-term health and well-being of the mother and the infant [7-9], the American College of Obstetricians and Gynecologists (ACOG) and the American Academy of Pediatrics emphasize the importance of maternal perinatal care and infant preventive care. The benefits of quality perinatal health care are well established, reducing the risk of pregnancy complications, rates of low birth weight infants, and infant mortality rates [5,10,11]. However, the United States fares much worse in preventing pregnancy-related complications than most other developed countries, despite spending more than any other country on hospital-based maternity care [12,13]. Growing concerns regarding maternal-infant health outcomes, patient satisfaction, access to quality prenatal care, and costs have increased interest in alternative models of prenatal care [14]. Given the broad and ubiquitous nature of technology, digital health tools have the potential to advance perinatal care and empower women to engage in the provision of care while maintaining expert recommended standards of care [15,16].

Evidence shows that pregnant women and those with young children are accustomed to readily available information using digital technologies and desire better access to information offered by health professionals [17]. Earlier efforts to supplement perinatal care with digital health tools have demonstrated variable levels of technological complexity. One of the most notable public health campaigns for perinatal health is Text4baby. With more than 250 messages tailored to pregnant women and new mothers, Text4baby represents one of the first empirically supported mobile health campaigns to reach over 685,000 mothers through text messaging [18]. Similarly designed for national scalability, Expect With Me follows the same schedule as individual prenatal care from week 14 of pregnancy and follows ACOG recommendations for clinical practice implemented through group prenatal care supplemented with information technology [15,19]. The obstetric OB Nest program proposes a reduced number of prenatal visits (ie, 8 onsite obstetric appointments; 6 virtual nurse visits) for low-risk pregnant women by leveraging technology (eg, fetal heartbeat and blood pressure home monitoring devices; web-based social support) to demedicalize the pregnancy experience and provide care within patients' daily lives [20,21]. Similar efforts to reach women in their daily lives, researchers found it feasible to use an embodied conversational agent (ie, animated conversational character simulating face-to-face interaction) accessed over the web, *Gabby*, to promote preconception health, healthy eating, and stress management [22,23]. Collectively, a review of perinatal care and telemedicine/eHealth suggests that digital tools may be beneficial in empowering patients and promoting value-based health care, yet ongoing efforts are needed to provide evidence specific to health outcomes, satisfaction, and cost and reflect a constantly evolving digital landscape [24]. As a digital health intervention tool, voice technology has recently been explored as a modality for delivering information to support health and well-being [25]. For the purposes of this paper, we define voice technologies as digital tools and devices that enable bidirectional communication of information through speech (eg, conversational agents, dialog systems using audio content or text-to-speech over smart speakers, smartphone voice assistants, voice-based apps). Voice technology interventions that rely on listening and speaking interactions differ from visual intervention predecessors and warrant further research to understand how users interact and consume information [26].

Aims of This Study

The primary aim of this study is to assess the feasibility of delivering perinatal health education via voice among a group of perinatal women. To explore the potential of voice technology in maternal-infant health, we aimed to assess the feasibility of a voice technology mobile app prototype, Self-Management Intervention–Life Essentials (SMILE), among a group of perinatal women. Following expert recommendations [27], we defined feasibility through 4 general domains: (1) acceptability, (2) demand, (3) practicality, and (4) adaptation. We examined the ability of the application to retrieve and deliver perinatal health information through spoken words (eg, podcasts) and prompt and audio-capture participant reactions to intervention content immediately following the podcasts. Before efficacy testing, we sought to understand participants' tolerability of the platform, appropriateness and interest in the content and delivery, and how participants used the system.

Intervention: SMILE

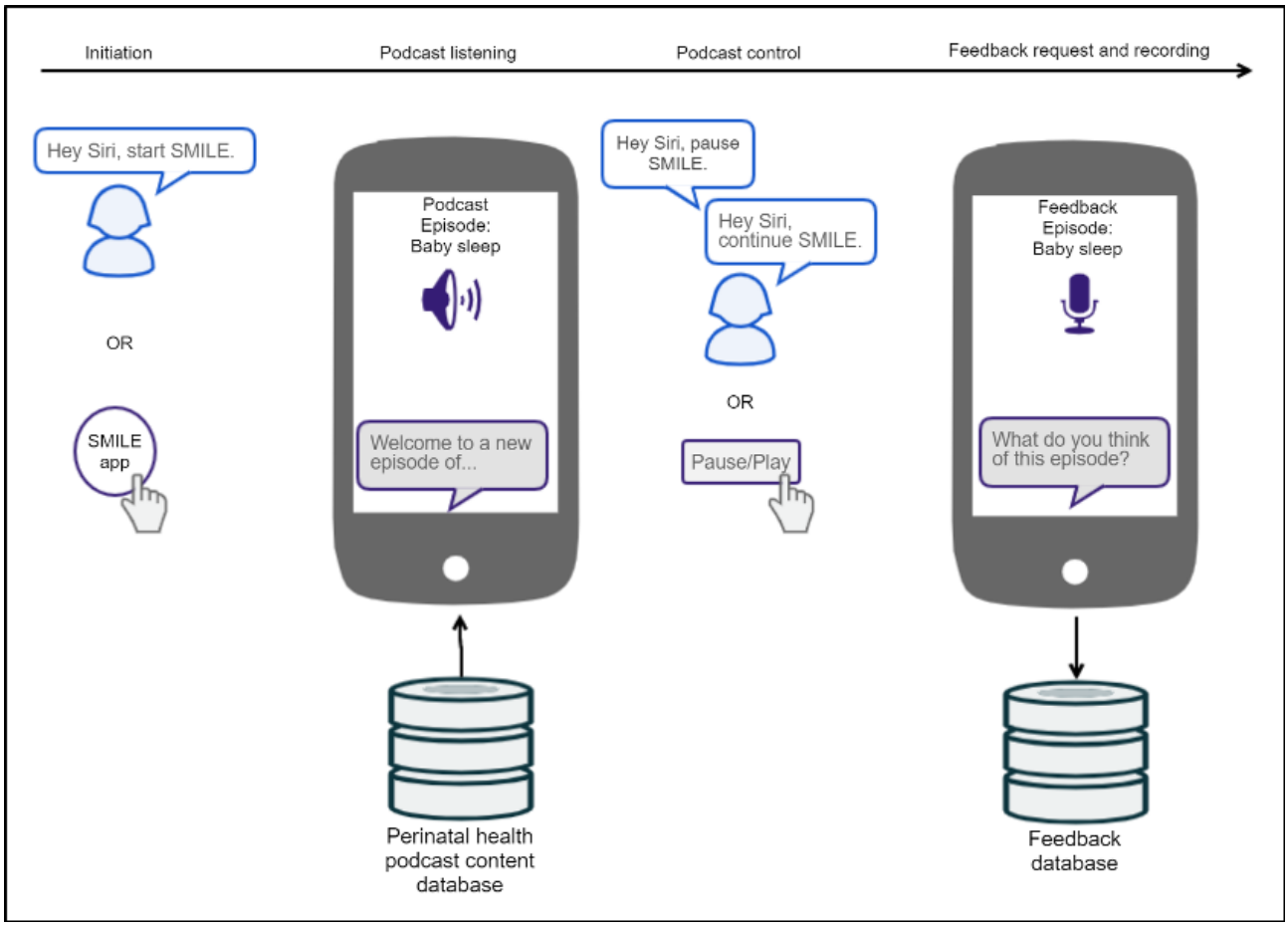
As an initial prototype, SMILE was created using the input from the literature. From pregnancy through a baby's first birthday, the literature collectively identifies the following categories necessary in perinatal education: information regarding infant/babies' needs, postpartum care and postpartum depression, baby's feeding/breastfeeding, strategies to manage the couple's relationship, mobile/digital resources with links to reliable documents, and a list of useful contacts/professional resources [28,29]. In addition to traditional resources such as family and close friends, new parents use alternative contemporary channels to find perinatal (pregnancy/parenting) information to include mobile and internet-based resources [30]. Specifically, digital health interventions have demonstrated the ability to provide women with perinatal support when they most need it (ie, immediate) and/or when they have opportunities to

access content (ie, support is more readily accessible than clinic visits alone) [17,28].

Podcasts are increasingly being used for education, both to providers and patients, with demonstrated feasibility, acceptability, and reach [31,32]. We leveraged the convenience of podcasts to deliver SMILE content. SMILE was developed to retrieve content from the long-standing, evidence-based Dr Mike PediaCast program affiliated with the clinical setting for the study [33]. PediaCast is a parent-facing podcast that provides relevant information and news to parents by answering listener questions, interviewing pediatric experts, offering overviews of research, and providing the latest news on pediatric health–related topics.

For initial testing and because of time and monetary constraints, SMILE leveraged existing voice technology services (ie, Siri, podcasts) and was designed as a mobile app for use on the iPhone platform. SMILE could be initiated through voice on the user's phone (eg, "Hey Siri, start SMILE") or by launching the app through touch/tap, either method would then start a preselected podcast. Intervention podcast topics were selected based on their relevance to infant care practices, prenatal care practices related to improved infant outcomes, and the duration of podcasts. Although evidence suggests options for users to tailor/personalize the intervention, podcasts were delivered sequentially to gauge user perceptions of various topics and durations. Users were able to listen to podcast in the background (ie, screen off), with an option to play/pause with voice command through Siri ("start SMILE" to initiate the app and it automatically starts the podcast, "Pause SMILE" to pause episode, and "Continue SMILE" to play where left off). Finally, upon completing each podcast episode, the app (through spoken language) asked users to provide feedback by answering (verbally) 2 brief questions. User feedback was collected by the app through voice recordings. Figure 1 highlights SMILE functionalities and how participants could use the app during the study.

Figure 1. Self-Management Intervention–Life Essentials interactivity highlights.



Theory

The underlying tenets of the proposed innovation are theoretically grounded in cognitive load theory [34] and the technology acceptance model (TAM) [35]. Cognitive load theory asserts that when experience overloads working memory capacity, learning is impaired [34]. Learning is better supported when content is broken down into smaller, more manageable pieces. As such, SMILE podcasts ranged from 3- to 8-minute chunks, each episode slightly longer, but allowed users the option to stop after one episode or continue listening. Qualitative aspects of the study were guided by the TAM (discussed in detail in the *Data Collection* section).

Methods

Design

We conducted a 2-week within-subject feasibility study with 19 perinatal women.

Recruitment, Sample, and Setting

Recruitment occurred over a 1-month period (April to May 2019) at the Nationwide Children's Hospital, Columbus, Ohio. Participants were invited to participate through hospital email announcements and word of mouth. Interested persons were screened before enrollment. To advance the health of women and infants, a two-generation approach acknowledging the interrelated health between mother and infant is critical. Therefore, the main study inclusion criterion was perinatal women ≥18 years of age, either pregnant or having an infant less than 1 year of age. Owing to technical aspects of the intervention, participants were also required to be English speaking and an iPhone user. In appreciation for participant time and feedback, compensation was provided for baseline survey completion and downloading the app (US $10), field testing the app, and participating in an exit interview (US $20) for a possible US $30 total.

Data Collection

Upon written informed consent, participants completed baseline surveys that captured demographics, patterns of technology use, and technology use specific to perinatal health and/or self-care practices (Multimedia Appendix 1). Participants were then onboarded and instructed to download and use the app over a 2-week period. To prevent nonrecruited users from downloading and using the app, participants were provided with a link to download SMILE along with an assigned entry code necessary to launch the app. To broadly assess app usage, data were collected from app store analytics (ie, number of active devices, number of impressions). Participants' individual app use (eg, demand) was also collected (ie, podcast number, duration of app used) to understand how participants progressed through the intervention. Qualitative data reflective of participants' acceptability, perceptions, and attitudes toward the intervention were obtained through 2 channels. First, immediately following each podcast, participants were asked 2 questions and responses were audio-captured and recorded through the app. Second, participants were invited to participate in a semistructured exit interview following the 2-week intervention field test. Blending formal structured and unstructured interviewing techniques, semistructured techniques are widely used in formative research studies that provide researchers with information on the acceptability of intervention components [36]. The exit interview was informed by the TAM [35] and a scale to assess burden [37] to gather feedback and attitudes toward the technology (Multimedia Appendix 2). Although scheduled and drop-in sessions were offered for exit interviews within the hospital setting, 25% (3/12) of participants completed face-to-face interviews, whereas the rest opted to complete interviews via audio/videoconferencing. Qualitative feedback was audio recorded, deidentified, transcribed verbatim, and verified against actual recordings by study staff. Field notes taken during the interviews were used to supplement the transcripts.

Data Analysis

Quantitative data were analyzed using descriptive statistics. Two researchers performed thematic analysis of qualitative data [38,39]. Owing to the study design, 2 forms of qualitative data were evaluated. First, qualitative data collected immediately following each podcast episode queried participants about the intervention content. Second, qualitative data collected via exit interviews provided participant feedback on the overall intervention. For all qualitative data, researchers first became familiar with the qualitative data, which involved multiple readings of transcripts, but did not code any data. Third, key concepts were identified. Color-coding strategies, both manually and with Excel spreadsheets, were employed to highlight various concepts and generate initial codes. Coded data were reviewed by research team members who compared and contrasted their independent findings. The initial codes were iteratively modified in the process of open coding to capture information relevant to the research question. Identified thematic findings were reviewed and modified within the context of the larger data set to ensure that the themes were cohesive, yet distinct [39]. Final, prominent themes were discussed between the 2 authors until a consensus was reached. To gauge user perceptions of episode content and delivery, an additional sentiment analysis was independently performed for app-collected feedback. Participant feedback responses were coded −1 for negative, 0 for neutral, and 1 for positive. Disagreements were resolved through discussion and consensus.

Ethical Consideration

Ethics approval was obtained from the participating hospital, Nationwide Children's Hospital Internal Review Board (IRB #00000159). Participation was strictly voluntary, and participants were informed of their right to withdraw at any time without penalty. Participant data were deidentified and stored on a secure server.

Results

Demographics

Collectively, 19 participants (17 to 36 weeks pregnant) were consented and completed baseline surveys, 18 downloaded the SMILE app, 17 used the app, and 12 participated in the exit interviews. The sample was predominantly White (15/19, 79%), married (19/19, 100%), between 25 and 34 years of age (16/19, 84%), and pregnant with their first child (12/19, 63%). Although education level and occupation were not formally assessed, some of the participants self-identified as nurses or social workers or in *admin* during the exit interviews.

At baseline, the top 3 resources participants used for pregnancy-related information included calling their health care provider (7/30, 23%), searching the web (7/30, 23%), or using smartphone apps (6/30, 20%). The top 2 most cited apps used to support perinatal health were Ovia (6/42, 14%) and What to Expect (5/42, 12%). Resources least used included an information packet provided by a health care provider (2/30, 7%), calling a friend (2/30, 7%), or using a doula (1/30, 3%). Calendar was the most cited (6/14, 43%) app used to support everyday life. More than half of our participants reported using

voice assistants before this study (11/19, 58%). Table 1. Participant-reported voice technology data are highlighted in

Table 1. Participant-reported voice technology engagement (baseline).

Category	Value, n (%)
Smartphone-specific voice technology engagement	
Duration of voice use (n=11 participants)	
1-3 months	2 (18)
3-12 months	1 (9)
1-3 years	6 (55)
>3 years	2 (18)
Top 3 interactions (n=21 responses)[a]	
Weather	4 (14)
Phone calls	3 (10)
Timer	3 (10)
Date/time	2 (7)
Information seeking	2 (7)
Texting	2 (7)
Music	2 (7)
Reminders	2 (7)
Map	1 (5)
Smart speaker–enabled or voice-enabled technology engagement	
Duration of voice use (n=10 participants)	
1-3 months	2 (20)
3-12 months	2 (20)
1-3 years	5 (50)
More than 3 years	1 (10)
Top 3 interactions (n=26 responses)[a]	
Weather	6 (23)
Playing music	6 (23)
Smart house (lights, switches, and thermostat)	5 (19)
Timer	4 (15)
Information seeking	2 (8)
Cooking recipes	1 (4)
Conversion or calculation	1 (4)
Smart television access (search movie)	1 (4)

[a]Participants offered more than one response.

App Usage

Unfortunately, the system failed to capture data related to how the app was initiated (touch/tap or voice), time of day when the app was being used, or app usage duration. Therefore, app usage was determined from participant feedback captured immediately after each podcast episode listened to by the participants. Using these data, SMILE was able to deploy 239 podcasts across 17 participants in a 2-week timeframe (Table 2).

Table 2. Sentiment ratings for podcast episodes grouped by category.

Category	Proportion of podcast episodes available by category[a] (n=23), n (%)	Proportion of listens by category[a] (n=239), n (%)	Frequency of response for question 1 ("What did you think of this content?")	Content positively rated (question 1), n (%)	Frequency of response for question 2 (Interest in hearing more)	Interest in hearing more (question 2), n (%)
All	23 (100)	239 (100.0)	237	205 (86)	238	209 (86)
Sleep	8 (35)	81 (33.9)	81	71 (88)	80	67 (84)
Pregnancy	7 (30)	72 (30.1)	71	59 (83)	72	66 (92)
Parenting	7 (30)	72 (30.1)	75	71 (95)	74	60 (81)
Breastfeeding	4 (17)	36 (15.1)	35	32 (91)	36	33 (92)

[a]Some percentages will sum to >100%, as some categories overlap.

Participants were required to listen to the podcasts in order. The response rate to the 2 questions prompted after each podcast was 60.1% (475/782). The least amount of words voice captured from participant feedback was 2 (17 characters), while the most amount of words voice captured from participant feedback was 134 (725 characters). Participants most often used 18 words (12/239, 5.0%) in voice feedback, with a median of 34 words per response (IQR 37.3, min-max 20-57.3). Participant feedback was coded positive (1, *good*, *helpful*), neutral (0, *not sure*), or negative (−1, *not relevant*, *not helpful*). The full podcast episode list and ratings are given in Multimedia Appendix 3. From the 475 responses collected by the system, 87.2% (414/475) comments were positive. The average sentiment rating for question 1 was favorable at 86.8%, and the average interest in hearing more was positive (85.8%). Despite the positive feedback, the number of users per episode was reduced by nearly half during episode 15. Each category was covered by at least two podcasts during the first 15 episodes. Therefore, the notable attrition does not necessarily reflect loss of interest in a category but may reflect decreased use of the intervention over time. Furthermore, it is difficult to say with certainty that missing data (eg, no user feedback for an episode) were truly indicative of missing data or technical glitches. From the app-captured participant feedback, at least three instances were noted where the technology failed to capture participants' initial responses:

> *I already answered questions for this episode, and it didn't save.* [Participant 5, episode 14]

Podcast Qualitative Feedback

Participant feedback captured by voice recordings following each podcast was analyzed qualitatively. Sentiment was gauged from 0% (poor/negative) to 100% (good/positive). Themes identified from content feedback include (1) establishing/transitioning routines, (2) expected norms and tempered expectations, and (3) key takeaways. The themes and representative quotes are given in Table 3.

Table 3. Examples of qualitative feedback to podcast episodes by theme.

Theme/subtheme	Example (feedback on podcast)	Interested in hearing more
Routines	"Like it's very relevant to a new mom because I currently have a 16-month-old and pregnant and I did not know how difficult getting a sleep routine down was. So, I think hearing I didn't know that routine made such a difference and once I started routine. It was very helpful." (Participant 17, episode 8)	"How and when to start sleep routines with younger babies" (Participant 8, episode 9)
Expected norms/realistic expectations	"I've had a lot of friends tell me that it's good that I'm considering breastfeeding, but not to feel so ashamed if it doesn't happen the way that I want it to go and it's okay." (Participant 15, episode 14)	"I think it might be helpful to hear more about what to look for, how to treat it, when to call the doctor." (Participant 17, episode 1)
Troubleshooting (subtheme)	"I like this episode a lot because he went through several different reasons for why the baby was crying before sleep and kind of troubleshooting to figure out what the issue was. It was nice to hear him describe several options that could've been the case. Yeah, I liked it a lot. I thought it was useful." (Participant 10, episode 8)	"... what are some techniques to help soothe them rather than feed them and what are some cues as a mom to figure out that they really need fed or they just want attention." (Participant 17, episode 19)
Provide objective, key takeaways	"I thought this episode was pretty detailed and gave good information. I would have liked to also hear what to look for when looking for signs of suffocation and then maybe have him defined supervision a little bit MORE ... very helpful to have an acronym to remember safe practices." (Participant 15, episode 6)	"As I mentioned previously it would be helpful for parents to have more information about specific examples." (Participant 11, episode 15)

Theme 1: Establishing/Transitioning Routines

Content about establishing routines and transitioning to having routines were of interest to the participants. For example, the ninth podcast, *Infant Sleep Problems*, was the fourth time that typical infant sleep patterns were discussed. However, participants' feedback indicated that they would be interested in hearing more about the timing and strategies to establish sleep routines. Collectively, participants were interested in

learning *when* it was safe to establish or transition their baby and strategies on *how* to go about it.

Theme 2: Expected Norms and Tempered Expectations

Participants were eager to hear information that validated their prior knowledge and experiences to gain *peace of mind*. Participants favored information that showed the pros and cons of various parenting techniques and infant care strategies. Anecdotal experiences can be perceived as positive if supported by accurate medical data:

> *More real-life examples of breast feeding and issues that go along with it.* [Participant 6, episode 14]

When caring for their newborn, participants wanted information about expected behaviors and strategies to determine infant needs, specifically infant *cues*. Feedback indicated that participants desired to learn more about alternative solutions to common parenting challenges.

Theme 3: Provide Key Takeaways

Participants favored podcasts that provided objective data to dispel myths. Often, information reinforced with statistics, data, or research has been favorably received. However, this information must be tempered by presenting the information in terms of the user's understanding:

> *...it had a lot of good data, but it was almost too data heavy to stay focused on*
>
> *what it was saying because it was statistic after statistic after statistic.* [Participant 8, episode 2]

Information tended to be favorably received if it provided succinct tips and was not *all over the place* or *lots of information* as in too general or too much.

Exit Interview Findings

Exit interviews were focused on participant feedback and attitudes toward voice technology. Exit interview qualitative data were analyzed separately from app feedback qualitative data. Themes that emerged from coded exit interviews included (1) customization of user preferences, (2) privacy concerns, (3) family and friends, and (4) convenience and context.

Theme 4: Customize and User Preferences

Participants expressed their preferences for seeking information and learning. In exit interviews, participants spoke to the benefits of having information presented in more than one way, with at least three women self-identifying as visual learners:

> *If it's something in depth you know, and I want to be able to... I am a visual learner to see the chart or the you know to comprehend it. I need both.* [Participant 3]
>
> *When I was looking at My Chart (patient portal), there have been times that I didn't even remember what the test was called that they did. So, it would have been hard to recall that and then go through all of those tests... I like to see things versus hearing. I like a little bit of both, but I lean towards more visual so I like to be able to click through.* [Participant 7]

Prior experience and user familiarity with technology have emerged as relevant. Some described themselves as *not very tech savvy* and preferred more traditional methods such *writing things down*:

> *I prefer reading because it can, um, allow me to go at my own pace and I don't miss information. Um, and then I can use that as notes. Like, I can take a snapshot of it, uh, and use that to reference back upon because when you hear stuff, it doesn't necessarily, like you might not hear it the right way. Or, um, you know, you can't remember exactly what was said.* [Participant 2]

Finally, participants advocated for expanding content to allow for more customized material and tailoring options. Thus, the creation of user profiles may be beneficial:

> *... being able to like choose their own content, um, not necessarily have to listen to the same thing multiple times in a row and same in different ways, you know? If there's more options, then people can pick their own thing. I think that that will have more benefit and possible interaction with voice commands and that kind of stuff.* [Participant 1]
>
> *How, in my mind, I'm thinking when you go into Netflix, and you get a profile, and you pick, you know, my husband has a profile, I have a profile, and when you log onto that, they recommend certain shows to watch based off your past history. So, I guess if there's a way that you could log into your Alexa as far as like your portal or your profile, and then it kind of like tracks your trends of things you ask or interests like music that you play.* [Participant 8]

Theme 5: Privacy

Privacy was brought up in an array of contexts. Privacy of voice technology was discussed within the context of information control and information sharing.

One participant stated this about digital privacy:

> *I don't think I have a lot to hide... I'm not nervous about it. But I also feel like I probably should be a little bit more aware of it. Cuz I feel like on my phone it's like you can be talking about something and then you get on your Facebook page.* [Participant 2]

Privacy was also mentioned with regard to context of voice:

> *I'm rarely in a quiet and alone place where I really can speak, to give a command to a phone, even if it were to recognize me more sometimes either maybe a privacy thing, or I don't want everybody around me to know and I'm I just don't want to be broadcasting is what I'm learning about.* [Participant 11]

Theme 6: Family and Friends

Exit interviews captured voice-enabled devices used in the home setting (eg, smart speakers, smart switches, smart thermometers) other than smartphones alone. Voice technology use was

perceived differently depending on the family member using it.

Several participants spoke of the influence their spouses, particularly their husbands, had on technology uptake. In most cases, spouses promoted the uptake of technology; however, they could also pose as a barrier.

Responses suggested that husbands were the drivers of technology within the household:

> *Yes, absolutely. Yeah. If he wasn't super techno savvy, then I probably honestly wouldn't have an Alexa and I wouldn't have done the smart lights or things like that. He's definitely the driver behind because he likes being the new and cutting edge and trying it out and that type of stuff.* [Participant 9]

> *Yeah we don't use Alexa at home. Cuz mostly my husband is a security person, he doesn't like it. I'd use it myself a lot more if it were not for him.* [Participant 10]

Husbands could also be engaged in the perinatal process through voice:

> *I could absolutely see my husband going like Alexa, tell me about my baby today and she's rattling off like your baby should be about this weight and should be eating this much and all this like he would totally absolutely use it. I would probably use it as well if the information is valuable.* [Participant 9]

Participants voiced similar concerns about their children's safety, as infants/siblings learn to speak and role model others, exposure to inappropriate material, and becoming addicted to technology. Others have suggested a need for parental control on voice-enabled devices:

> *I'd rather the kids wait until they're older to use it, especially unsupervised. I have a baby that's not even two right now. And she already yells that, you know, she says, Hey, Google play a song. So, it's just playing random songs at her all day. And if I'm in the bathroom or something, I feel like I'm not in control of what she's doing or listening to.* [Participant 4]

> *Our daughter can say, Mom, Dad, and Alexa.* [Participant 2]

Participant feedback highlighted that voice technology used in the home setting provided instances to engage multiple users simultaneously, such as family and friends. However, the content needed to be appropriate for diverse audiences:

> *I mean, in terms of the content of the episodes you had us listen to... it would be great for my husband to hear and I'm fine with my kids hearing that kind of content. I didn't think that it was inappropriate at all.* [Participant 5]

One participant reported not using technology to manage her stress but suggested a case of joint media engagement:

> *As far as technology, no. My husband will sometimes. I don't know what the app is called, but he uses it like a meditation app sometimes when we go to bed. He'll play it and it's just like breathing.…it's like breathing techniques we'll do just to like, relax to go to bed. I make fun of him about it sometimes, but [laugh] it is very beneficial. It's just sometimes I'm like I don't like doing this I just want to go to bed.* [Participant 2]

Although not as frequent, some have also reported using voice for group entertainment. In reference to the voice assistant led trivia:

> *it's pretty much music trivia and they'll play a snippet of a music. Yes, and it's just fun if people are over, it's like a pastime kind of little game that doesn't, it's not intimidating people the whole group can play.* [Participant 2]

> *Yes, we have at friends' houses before we don't do it regularly though … but we should utilize the game more often because it was fun.* [Participant 3]

Theme 7: Context and Convenience

Context played a major role in how women perceived their interactions with voice technology. Women also discussed how voice interventions could be weaved into their daily routines:

> *in like, like a convenient setting like if I could have it, so one of the things that our Alexa does in the morning is like a morning news update and they'll just give us like a 5 or 10 minute update about like, what's happening in the news something like that could happen like today in your pregnancy and or like this might be a topic if you're interested in that would be really cool then I could just listen to it while I'm getting ready in the morning.* [Participant 4]

Women favored voice interactions, which were convenient. A number of women reported that they did not regularly use the voice assistant on their phones (ie, Siri) because of frustration or a lack of understanding. General sentiments suggested that voice interaction could be challenging:

> *I mean, it's come a long way compared to what Siri used to be used to be even worse, but in terms of what you can do, or how you ask a question matters a lot. What exactly you say matters a lot as to what type of response you're gonna get.* [Participant 9]

Outside of the home setting, no voice use was preferred over Siri. In the home setting, Alexa-enabled devices were favored over Siri-enabled smartphones:

> *I find that Siri doesn't work as well for me, my husband's much better at it. I am much more comfortable using Alexa. I find that it responds better.*

> *Yeah, well, I mean, like we use that because everything in our house is tied into Alexa … but like if I'm somewhere else, like, I won't use Siri on my phone.* [Participant 9]

> *I do use Alexa at home and I don't really use Siri that much on my phone though.* [Participant 4]

However, the value of voice-enabled smartphones increased when participants needed to be hands-free, particularly in cars.

Even if a participant reported not using voice routinely on their phone, they did in the car:

> *No, I don't use Siri. I've found that she doesn't really understand what I say. So, I don't tend to use it. But I do have, like, voice recognition in my car that I use a lot.* [Participant 5]

> *Especially thinking about being able to use it in the car, you know, to be able to get in and ask the app to you know, play me episodes about, you know hygiene or whatever.* [Participant 6]

Finally, women stated that it would be beneficial to have voice assistants to help them communicate with their health care providers and/or receive laboratory results. However, users wanted to have control over voice technology-communicated medical information:

> *I think it's If it becomes convenient, and there are less hoops to jump through to be able to communicate to healthcare providers, then heck yeah, I'll use it.* [Participant 6]

> *If I had lab results and it came through and Alexa was like, Hey – you have new lab results, do you want to hear them? I would love that…If I'm having a dinner party, I'd probably say no not right now. You know what I mean? If there's people over, I could see why I would want to defer it.* [Participant 3]

Specifically, women indicated that it was more convenient to listen to the intervention compared with providing feedback, particularly relevant with regard to user context. One participant reported the following:

> *I actually didn't really like it. I didn't really want to like give the vocal feedback when I was out and about, like in my car listening to it or like in an airport... So if somebody were to give, if this were to be an app, and someone was trying to give feedback, it could be definitely a distraction thing if they're driving and trying to give feedback.* [Participant 5]

Discussion

Principal Findings

Maternal child health remains to be a public health priority, and practices to improve outcomes are urgently needed. As an adjunct to clinical care, digital health interventions have the potential to broaden opportunities to intervene when patients may be most receptive to support [40]. The use of voice technology in digital health is nascent, particularly among perinatal populations. Thus, the goal of this study is to address the question, "Can it work?" [27] and provide evidence to support or negate the use of voice technology in perinatal digital health interventions. To address the primary research aim and determine whether voice technology is an appropriate medium to leverage in perinatal health, feasibility is framed in key focus areas: (1) acceptability, (2) demand, (3) practicality, and (4) adaptation [36].

Acceptability

Acceptability was examined by how the individuals reacted to the intervention, perceived ease of use, perceived helpfulness, and intention to try podcast recommendations. At baseline, users were more familiar with common smartphone features such as text messaging, calendars, and apps compared with voice technology. However, no users required additional assistance beyond onboarding to use the intervention, nor did any users provide feedback that the voice-based app was difficult to use. Of the 239 podcasts listened to by participants, sentiment ratings suggested that the content was favorably received, particularly on the topic of breastfeeding. Exit surveys indicated that women were receptive to using voice technology as a potential platform to support health. The findings also showed that participants perceived the advantages/disadvantages of voice technology depending on the device. On smartphones, voice assistants were perceived to have quick response rates, yet users complained that the technology often did not understand them or provided unsatisfactory answers. On smart speakers/devices, voice assistants were perceived as convenient, but unnecessary and unfavorable because of costs.

Demand

Demand was assessed by documenting intervention activities and self-reported use of technology. As a percentage, overall response rates to intervention activities (approximately 61%) suggest a fair ability/interest in completing the study tasks. Similar to other research [28], the participants in our study were technologically capable of reporting interactions with various technologies before SMILE. However, at baseline, only 1 participant reported using a smartphone to listen to podcasts, and during exit interviews, less than half of the participants reported listening to podcasts outside of the intervention. Just over half of our sample reported using voice assistants (11/19, 57.9%), slightly higher than the Pew survey data that reported 46% of Americans reported using digital voice assistants [41]. Of those with previous voice experience, 72.7% (8/11) had ≥1 year experience with smartphone voice technology and 60% (6/10) had ≥1 year experience with voice-enabled devices in the home setting. Baseline survey data and exit interviews aligned, highlighting that hands-free, convenient activities were a strength of voice technology. However, baseline surveys highlighted case-specific uses for voice technology (eg, setting a timer, checking the weather), whereas exit interviews revealed broad uses for voice technology (eg, games, communication to more than one person, answering questions). Across all sources of data collected, inability for voice technology to recognize what the user was saying was a primary reason for nonuse.

Practical

Participants did not find the podcasts *practical* when they were asked to refer to other podcasts for the background. Similarly, participants did not find the podcasts practical when content was not perceived as relevant to them. Participants desired content that was perceived as sensible, objective, and poised in a manner that tempered prior knowledge with new information. These findings support the use of cognitive load theory as an intervention guide, emphasizing the importance of balancing processing capacity relative to cognitive load. Regarding the

practicality of intervention delivery, the voice was highly sensitive to the user context. Several participants discussed instances when intervention activities were either facilitated or constrained (ie, listening/speaking in public vs private). For instance, intervention delivery was favorably perceived during multitasking situations when users desired to be hands-free (eg, driving). Conversely, intervention activities were negatively perceived when in public settings or when with friends or colleagues outside of the woman's immediate perinatal social support system (ie, friend vs spouse). To build upon these strengths and limitations, intervention activities should be available on demand and able to satisfy user needs at the moment. For example, the podcast duration should be able to accommodate activities such as driving, both short (10-15 minutes) and long commutes (3-4 hours).

Adaptations

As such, a number of *adaptations* are recommended for voice interventions to best align with user situations. Our data suggest that voice intervention activities that deliver personal or sensitive information or require quick response/feedback may be better suited for delivery on smartphones. Conversely, voice intervention activities that are more general may be better suited for delivery via smart speakers, which may also be an opportunity to engage other members of the family or social support. Although participants did not directly discuss using SMILE with others, many did see an opportunity to engage spouses and extended family members (eg, other children/siblings). In joint media engagement, more than 1 user jointly may engage with a technology and, consequently, support prescribed activities [42], which is a potential strength for voice technology. Depending on delivery and users, voice technology has the ability to support stealth health promotion, wherein efforts are perceived as an activity spent with family or friends, and the target of the intervention (eg, stress management) is a side effect but not the primary motivator of participation [43,44].

Voice Interaction and Podcast Intervention Use

Similar to attrition found among digital health interventions [45], SMILE data suggest user attrition over time. As a prototype, SMILE was rapidly designed for this population to gauge user responses to voice technology in terms of listener preferences for duration, timing, and interaction. The sequence, number, and duration of SMILE content were fixed. After the 14th podcast, the episode duration was slightly longer (ie, further away from 5 minutes, closer to 6-8 minutes in duration). As the intervention was delivered via smartphone, users could see the title of the podcast and duration, which could bias them to listen or not. User preferences for learning may also have contributed to attrition. Some women reported personal preferences for learning (eg, *like to see things versus hearing, go at my own pace*), which may have been in contrast to how the intervention was designed. Given the documented need for perinatal support and value placed on digital health tools by perinatal women, particularly those tools that are multifunctional [17,28,30,46], we believe there is value in pursuing both voice and visual interventions in perinatal health. Specifically, our findings align with other research [26,47], suggesting a need to address the inefficiencies of voice navigation (eg, pace, duration, ability to choose from a list of options).

Podcasts were purposefully selected based on recommendations from the literature; therefore, our findings reaffirm previous literature and the importance of topics such as breastfeeding, infant sleep, parenting skills, and pregnancy self-care. Similarly, we also found that women prefer tailored content, which is relevant to their needs, practical, unbiased/objective, and available on demand [17,28]. However, we discovered that the potential benefit of delivering perinatal health information via voice is the ability of women to digest perinatal health content while multitasking. Thus, the intervention could be woven into daily life. However, voice interventions are highly sensitive to context. As such, design and interactivity must be agile in the user context. For example, a number of women reported listening to podcasts while driving. Voice interventions must be careful not to place individuals in harm during use (eg, distracted driving). Similarly, a number of women discussed/envisioned listening to podcasts during *3 am feedings*. Thus, postpartum voice interventions, in particular, would benefit from adaptive volume control, such as Amazon's Alexa whisper mode.

Digital tools to support management of daily activities were common among our sample, specifically calendar use. Maternal-infant interventions deployed/accessed through voice-enabled devices in a shared setting (eg, home, car) have the potential to reach beyond the woman alone, to include social support members (eg, spouses, siblings). Perinatal interventions that aim to engage users may want to leverage such information, as research has shown that the value of something is increased if the activity can serve more than one need [44,48]. Qualitative findings from this pilot highlight opportunities to expand perinatal health promotion efforts beyond individual women to include spouses/partners using voice technology. Irrefutably, evidence shows that *pregnant women need the support of caring family members, friends, and health professionals* [49]. Research and interventions are required to provide partners of pregnant women with evidence-based information and support whole families during the perinatal period [50]. Evidence shows that men who attend antenatal care express concerns about being excluded and left feeling disappointed [30,51]. Programs that support new fathers need to help form realistic expectations, provide information ahead of time, and provide information about the possible changes in their conjugal relationship and how to develop related coping strategies. Our findings suggest that voice interventions are strengthened by individual user profiles. Voice research also suggests that users tend to explore less and choose higher-ranked items, which could be a potential limitation of individualized voice content unless options are personalized, yet diverse to expose users to broad options [26]. In exit interviews, Alexa was the most commonly mentioned smart speaker. When the study was conducted, Alexa did not support individual user profiles. However, both Google Assistant and Alexa now support individual users through voice recognition technology.

Privacy and Security

Privacy of voice technology is complex and serves as an indicator of voice intervention acceptability. The findings reflected participant concerns over health data privacy and patient control over information sharing (ie, sharing data when and with whom). However, recent findings from a systematic review found that few voice assistant research studies reported privacy or security concerns associated with voice assistants and no studies refer to proprietary challenges that can arise when using commercial devices [25]. Therefore, we strongly advocate for transparent research and reporting when using voice, raising standards accountable to the scientific community and the participants they serve. Such methods may include lay language to explain what data are being captured, how data are shared and their intended use, potential for security breaches, and options for participants to participate fully or partially (eg, delete partial transcripts).

Although some women were enthusiastic about using voice to receive and share medical data with their health care provider, at this time, there is not enough evidence to support the safe and effective use of voice between patients and providers. Reliance on conversational assistants for actionable medical information represents a safety risk for patients, and in some instances, may pose harm [52,53]. Further research is necessary to forge confidence in voice technology and explore methods to mitigate safety risks. Developers and health technology experts should explore opportunities to broaden voice technology use; however, transparency about partnerships and data use is ethically prudent as device capabilities expand [52,54].

Contributions to the Literature and Implications for Future Research

Findings from this feasibility study suggest a role for voice technology in maternal-infant health efforts; however, the size of the role has yet to be determined. A systematic review of voice assistant technology used in behavioral health research found that, from a limited number of studies, voice interventions were in the early stages of development with limited efficacy testing [24]. With the proliferation of voice-activated devices (eg, Apple Siri, Google Assistant, Amazon Alexa), there is substantial opportunity for empirically supported voice-enabled health solutions. Owing to dramatic improvements in voice recognition accuracy and intelligent conversational agents, voice assistant technology allows for true hands-free operation and conversation, increasing flexibility and efficiency.

Our findings further demonstrate the feasibility of capturing data from perinatal populations through voice. Voice technology is a novel strategy to collect data and interact with perinatal populations beyond the clinic. The benefit of using voice is that speaking/talking is naturalistic across diverse populations. Voice technology may help to reduce barriers associated with literacy (eg, spelling errors, mistyped words), support formative assessments, and engage social support beyond just the patient. In our future efforts, we plan to explore whether voice-captured data differ from other just-in-time data collection strategies (ie, do participants ramble or provide longer voice responses; how do voice data capture differ/align with text message data capture). Although improving, challenges to voice include errors in transcription (eg, tone, rates of speech), understanding various accents and medical terminology (eg, mispronunciation or misuse), and deducing user intent from context, that is, intent schemas to facilitate custom interactions with users [55]. Research in this area is critical to avoid user frustration and potential abandonment if the system does not understand or does not reflect user intent (ie, does not respond appropriately).

We believe that the proposed project is a vital first step necessary to elucidate how digital health, particularly voice activation, may be leveraged to promote positive perinatal health behaviors and reduce maternal/infant morbidity and mortality rates. This formative evaluation provides evidence to suggest that a perinatal educational support program using voice technology is acceptable among a group of pregnant women and has the potential to engage spousal and familial support. The findings will be iteratively combined with our team's current efforts to partner with key stakeholders (ie, low-income families, community health workers, health care professionals) in the development of a perinatal digital health platform with voice-enabled capabilities. Future efforts focused on voice should explore the feasibility, usability, and effects of voice interventions delivered through different voice-enabled devices (eg, smartphones vs smart speakers).

Limitations

SMILE represents the pilot work necessary to guide voice technology intervention use among perinatal populations. As such, the study was limited by a small, predominantly White, and married (2-parent) convenience sample and short study length. Similar to other digital health research in this population, moving forward, it will be necessary to engage women from socioeconomically disadvantaged backgrounds and rural locations to determine if intervention needs differ from this sample [17,54]. Detailed assessments of sociodemographic data and gauges of health and technology literacy were not captured in this study. In future research, lessons learned from this pilot will be used to conduct longer studies across diverse populations, to also include assessments of health literacy and technology proficiency. Despite efforts to reach postpartum women, our study only included pregnant women. Evidence suggests that reaching postpartum women is challenging; however, web-based recruitment strategies have been shown to increase interest/screening of postpartum women in health promotion research [56]. Other strategies worth exploring include recruiting women, infants, and children, health care providers, and mother-baby groups [57,58]. Another limitation was the basic functionality of the tested technology (voice-only interactions via mobile phones without multimodal and/or tailored content). However, the study and intervention were conducted to address the feasibility of using voice in perinatal populations and to provide a jumping off point for future research. As such, the effects of the app on perinatal health education and health outcomes were not measured.

Conclusions

This study is one of the first attempts to develop and evaluate the feasibility of a voice technology app to promote positive self-management skills during the perinatal period using

evidence-based podcasts. Our findings suggest that how pregnant women use digital health interventions differ not only between visual and voice but also between the type of device used. In addition, we collected feedback through voice interactions. The findings support further development and usability testing of voice technology to promote maternal-infant health outcomes. Given trajectory and market growth, the necessity for hands-free interaction with interactive voice devices will be a growing industry across all customer segments. The development of empirically supported interactive voice solutions should be a priority to address this inevitable need.

Acknowledgments

The authors would like to thank Kayla Finch for her assistance in data analysis and Michael Patrick (Dr Mike) and Stephanie Cannon for their support in providing and curating the podcast content. The SMILE app was developed in partnership with Duet Health. They would also like to thank Healthcare Information and Management Systems Society for providing funding support.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Self-Management Intervention–Life Essentials baseline survey.
[PDF File (Adobe PDF File), 751 KB - formative_v5i3e18240_app1.pdf]

Multimedia Appendix 2
Self-Management Intervention–Life Essentials exit interview guide.
[DOCX File , 24 KB - formative_v5i3e18240_app2.docx]

Multimedia Appendix 3
Self-Management Intervention–Life Essentials list of podcasts grouped by category.
[DOCX File , 492 KB - formative_v5i3e18240_app3.docx]

References

1. Petersen EE, Davis NL, Goodman D, Cox S, Mayes N, Johnston E, et al. Vital signs: pregnancy-related deaths, United States, 2011-2015, and strategies for prevention, 13 states, 2013-2017. MMWR Morb Mortal Wkly Rep 2019 May 10;68(18):423-429 [FREE Full text] [doi: 10.15585/mmwr.mm6818e1] [Medline: 31071074]
2. Sawyer B, Gonzales S. Kaiser Family Foundation: How does infant mortality in the U.S. compare to other countries? Peterson-Kaiser Health System Tracker. URL: https://www.healthsystemtracker.org/chart-collection/infant-mortality-u-s-compare-countries/ [accessed 2018-10-05]
3. Cohen S, Janicki-Deverts D. Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006, and 2009. J Appli Soc Psychol 2012:1334. [doi: 10.1111/j.1559-1816.2012.00900.x]
4. Giurgescu C, Zenk SN, Engeland CG, Garfield L, Templin TN. Racial discrimination and psychological wellbeing of pregnant women. MCN Am J Matern Child Nurs 2017;42(1):8-13 [FREE Full text] [doi: 10.1097/NMC.0000000000000297] [Medline: 27749288]
5. Howell EA. Reducing disparities in severe maternal morbidity and mortality. Clin Obstet Gynecol 2018 Jun;61(2):387-399 [FREE Full text] [doi: 10.1097/GRF.0000000000000349] [Medline: 29346121]
6. Mohamoud YA, Kirby RS, Ehrenthal DB. Poverty, urban-rural classification and term infant mortality: a population-based multilevel analysis. BMC Pregnancy Childbirth 2019 Jan 22;19(1):40 [FREE Full text] [doi: 10.1186/s12884-019-2190-1] [Medline: 30669972]
7. American Academy of Pediatrics. Engaging patients and families: periodicity schedule. 2019. URL: https://www.aap.org/en-us/professional-resources/practice-transformation/managing-patients/Pages/Periodicity-Schedule.aspx [accessed 2019-01-31]
8. Screening recommendations. American Academy of Pediatrics. URL: https://www.aap.org/en-us/advocacy-and-policy/aap-health-initiatives/Screening/Pages/Screening-Recommendations.aspx [accessed 2019-01-31]
9. American College of Obstetricians and Gynecologists. Redefining postpartum care task force. ACOG Postpartum Toolkit for Health Care Providers. URL: https://www.acog.org/About-ACOG/ACOG-Departments/Toolkits-for-Health-Care-Providers/Postpartum-Toolkit?IsMobileSet=false [accessed 2019-06-01]
10. Perinatal and infant health. Maternal & Child Health Topics. 2018. URL: https://mchb.hrsa.gov/maternal-child-health-topics/perinatalinfant-health [accessed 2018-12-06]
11. Tessema J, Jefferds ME, Cogswell M, Carlton E. Motivators and barriers to prenatal supplement use among minority women in the United States. J Am Diet Assoc 2009 Jan;109(1):102-108. [doi: 10.1016/j.jada.2008.10.013] [Medline: 19103329]

12. Pregnancy risk assessment monitoring system - reproductive health. Centers for Disease Control and Prevention. 2018. URL: https://www.cdc.gov/prams/index.htm [accessed 2021-02-05]
13. Maternal health in the United States. Maternal Health Task Force. 2015. URL: https://www.mhtf.org/topics/maternal-health-in-the-united-states/ [accessed 2019-06-01]
14. The American College of Obstetricians and Gynecologists. Group versus conventional antenatal care for women. Women's Health Care Physicians. 2018. URL: https://www.acog.org/Clinical-Guidance-and-Publications/Committee-Opinions/Committee-on-Obstetric-Practice/Group-Prenatal-Care?IsMobileSet=false [accessed 2020-01-06]
15. Cunningham SD, Lewis JB, Thomas JL, Grilo SA, Ickovics JR. Expect with me: development and evaluation design for an innovative model of group prenatal care to improve perinatal outcomes. BMC Pregnancy Childbirth 2017 May 18;17(1):147 [FREE Full text] [doi: 10.1186/s12884-017-1327-3] [Medline: 28521785]
16. van den Heuvel JF, Groenhof TK, Veerbeek JH, van Solinge WW, Lely AT, Franx A, et al. eHealth as the next-generation perinatal care: an overview of the literature. J Med Internet Res 2018 Jun 05;20(6):e202 [FREE Full text] [doi: 10.2196/jmir.9262] [Medline: 29871855]
17. Lupton D. The use and value of digital media for information about pregnancy and early motherhood: a focus group study. BMC Pregnancy Childbirth 2016 Jul 19;16(1):171 [FREE Full text] [doi: 10.1186/s12884-016-0971-3] [Medline: 27435182]
18. Wellpass. Text4baby. URL: https://text4baby.org/ [accessed 2014-01-16]
19. Cunningham SD, Lewis JB, Shebl FM, Boyd LM, Robinson MA, Grilo SA, et al. Group prenatal care reduces risk of preterm birth and low birth weight: a matched cohort study. J Womens Health (Larchmt) 2019 Jan;28(1):17-22. [doi: 10.1089/jwh.2017.6817] [Medline: 30256700]
20. Butler Tobah YS, LeBlanc A, Branda ME, Inselman JW, Morris MA, Ridgeway JL, et al. Randomized comparison of a reduced-visit prenatal care model enhanced with remote monitoring. Am J Obstet Gynecol 2019 Dec;221(6):638.e1-638.e8. [doi: 10.1016/j.ajog.2019.06.034] [Medline: 31228414]
21. de Mooij MJM, Hodny RL, O'Neil DA, Gardner MR, Beaver M, Brown AT, et al. OB nest: reimagining low-risk prenatal care. Mayo Clin Proc 2018 Apr;93(4):458-466. [doi: 10.1016/j.mayocp.2018.01.022] [Medline: 29545005]
22. Gardiner PM, McCue KD, Negash LM, Cheng T, White LF, Yinusa-Nyahkoon L, et al. Engaging women with an embodied conversational agent to deliver mindfulness and lifestyle recommendations: a feasibility randomized control trial. Patient Educ Couns 2017 Sep;100(9):1720-1729 [FREE Full text] [doi: 10.1016/j.pec.2017.04.015] [Medline: 28495391]
23. Gardiner P, Hempstead MB, Ring L, Bickmore T, Yinusa-Nyahkoon L, Tran H, et al. Reaching women through health information technology: the Gabby preconception care system. Am J Health Promot 2013;27(3 Suppl):eS11-eS20 [FREE Full text] [doi: 10.4278/ajhp.1200113-QUAN-18] [Medline: 23286652]
24. Sezgin E, Militello L, Huang Y, Lin S. A scoping review of patient-facing, behavioral health interventions with voice assistant technology targeting self-management and healthy lifestyle behaviors. SSRN Journal 2019. [doi: 10.2139/ssrn.3381183]
25. Halili L, Liu R, Hutchinson KA, Semeniuk K, Redman LM, Adamo KB. Development and pilot evaluation of a pregnancy-specific mobile health tool: a qualitative investigation of SmartMoms Canada. BMC Med Inform Decis Mak 2018 Nov 12;18(1):95 [FREE Full text] [doi: 10.1186/s12911-018-0705-8] [Medline: 30419896]
26. Slomian J, Emonts P, Vigneron L, Acconcia A, Glowacz F, Reginster JY, et al. Identifying maternal needs following childbirth: a qualitative study among mothers, fathers and professionals. BMC Pregnancy Childbirth 2017 Jul 03;17(1):213 [FREE Full text] [doi: 10.1186/s12884-017-1398-1] [Medline: 28673272]
27. Shorey S, Ng ED. Evaluation of mothers' perceptions of a technology-based supportive educational parenting program (part 2): qualitative study. J Med Internet Res 2019 Feb 13;21(2):e11065 [FREE Full text] [doi: 10.2196/11065] [Medline: 30758295]
28. Cho D, Cosimini M, Espinoza J. Podcasting in medical education: a review of the literature. Korean J Med Educ 2017 Dec;29(4):229-239 [FREE Full text] [doi: 10.3946/kjme.2017.69] [Medline: 29207454]
29. Parga-Belinkie J, Merchant RM. Voices in evidence-based newborn care: a how-to-guide on developing a parent-facing podcast. JMIR Pediatr Parent 2019 Dec 20;2(2):e16335 [FREE Full text] [doi: 10.2196/16335] [Medline: 31859674]
30. Mike Patrick. PediaCast - a pediatric podcast for parents. URL: https://www.pediacast.org/ [accessed 2020-01-06]
31. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
32. Suh H, Shahriaree N, Hekler E, Kientz J. Developing and validating the user burden scale: a tool for assessing user burden in computing systems. Computer Human Interaction. 2016. URL: https://www.hcde.washington.edu/files/news/Suh-UserBurdenScale-CHI2016.pdf [accessed 2016-12-01]
33. Ayala GX, Elder JP. Qualitative methods to ensure acceptability of behavioral and social interventions to the target population. J Public Health Dent 2011;71 Suppl 1:S69-S79 [FREE Full text] [doi: 10.1111/j.1752-7325.2011.00241.x] [Medline: 21656958]
34. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res in Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
35. Maguire M, Delahunt B. Doing a thematic analysis: a practical, step-by-step guide for learning and teaching scholars. 2017. URL: https://ojs.aishe.org/index.php/aishe-j/article/view/335 [accessed 2021-02-16]

36. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
37. Usability Testing. Usability.gov. URL: https://www.usability.gov/how-to-and-tools/methods/usability-testing.html [accessed 2020-01-14]
38. Pew Research Center. Nearly half of Americans use digital voice assistants, mostly on their smartphones. 2017. URL: https://www.pewresearch.org/fact-tank/2017/12/12/nearly-half-of-americans-use-digital-voice-assistants-mostly-on-their-smartphones/ [accessed 2019-03-21]
39. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
40. Yang L, Sobolev M, Tsangouri C, Estrin D. Understanding user interactions with podcast recommendations delivered via voice. In: Proceedings of the 12th ACM Conference on Recommender Systems.: Association for Computing Machinery; 2018 Presented at: 12th ACM Conference on Recommender Systems; 2018; Vancouver, British Columbia, Canada p. 190-194. [doi: 10.1145/3240323.3240389]
41. Turner-McGrievy G, Kalyanaraman S, Campbell MK. Delivering health information via podcast or web: media effects on psychosocial and physiological responses. Health Commun 2013;28(2):101-109 [FREE Full text] [doi: 10.1080/10410236.2011.651709] [Medline: 22420785]
42. Ickovics JR, Lewis JB, Cunningham SD, Thomas J, Magriples U. Transforming prenatal care: multidisciplinary team science improves a broad range of maternal-child outcomes. Am Psychol 2019 Apr;74(3):343-355. [doi: 10.1037/amp0000435] [Medline: 30945896]
43. Takeuchi L, Stevens R. The new coviewing: designing for learning through joint media engagement. Sesame workshop. 2011. URL: http://joanganzcooneycenter.org/publication/the-new-coviewing-designing-for-learning-through-joint-media-engagement/ [accessed 2021-02-05]
44. Wearing JR, Nollen N, Befort C, Davis AM, Agemy CK. iPhone app adherence to expert-recommended guidelines for pediatric obesity prevention. Child Obes 2014 Apr;10(2):132-144 [FREE Full text] [doi: 10.1089/chi.2013.0084] [Medline: 24655230]
45. Militello L, Hanna N, Nigg C. Leveraging family to promote digital health: findings from the retrospective pokémon go family study. JMIR Pediatrics and Parenting 2018 Aug. [doi: 10.2196/preprints.10679]
46. Kruglanski AW, Shah JY, Fishbach A, Friedman R, Chun WY, Sleeth-Keppler D. A theory of goal systems. In: Advances in experimental social psychology. Amsterdam: Elsevier; 2002:331-378.
47. Bickmore TW, Trinh H, Olafsson S, O'Leary TK, Asadi R, Rickles NM, et al. Patient and consumer safety risks when using conversational assistants for medical information: an observational study of Siri, Alexa, and Google Assistant. J Med Internet Res 2018 Sep 04;20(9):e11510 [FREE Full text] [doi: 10.2196/11510] [Medline: 30181110]
48. Miner AS, Milstein A, Schueller S, Hegde R, Mangurian C, Linos E. Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. JAMA Intern Med 2016 May 01;176(5):619-625 [FREE Full text] [doi: 10.1001/jamainternmed.2016.0400] [Medline: 26974260]
49. Alagha EC, Helbing RR. Evaluating the quality of voice assistants' responses to consumer health questions about vaccines: an exploratory comparison of Alexa, Google Assistant and Siri. BMJ Health Care Inform 2019 Nov;26(1) [FREE Full text] [doi: 10.1136/bmjhci-2019-100075] [Medline: 31767629]
50. East CE, Biro MA, Fredericks S, Lau R. Support during pregnancy for women at increased risk of low birthweight babies. Cochrane Database Syst Rev 2019 Apr 01;4:CD000198 [FREE Full text] [doi: 10.1002/14651858.CD000198.pub3] [Medline: 30933309]
51. Suto M, Takehara K, Yamane Y, Ota E. Effects of prenatal childbirth education for partners of pregnant women on paternal postnatal mental health and couple relationship: a systematic review. J Affect Disord 2017 Mar 01;210:115-121. [doi: 10.1016/j.jad.2016.12.025] [Medline: 28024222]
52. Entsieh AA, Hallström IK. First-time parents' prenatal needs for early parenthood preparation- a systematic review and meta-synthesis of qualitative literature. Midwifery 2016 Aug;39:1-11 [FREE Full text] [doi: 10.1016/j.midw.2016.04.006] [Medline: 27321714]
53. Baldwin S, Malone M, Sandall J, Bick D. Mental health and wellbeing during the transition to fatherhood: a systematic review of first time fathers' experiences. JBI Database System Rev Implement Rep 2018 Nov;16(11):2118-2191 [FREE Full text] [doi: 10.11124/JBISRIR-2017-003773] [Medline: 30289768]
54. Guendelman S, Broderick A, Mlo H, Gemmill A, Lindeman D. Listening to communities: mixed-method study of the engagement of disadvantaged mothers and pregnant women with digital health technologies. J Med Internet Res 2017 Jul 05;19(7):e240 [FREE Full text] [doi: 10.2196/jmir.7736] [Medline: 28679489]
55. Kumah-Crystal YA, Pirtle CJ, Whyte HM, Goode ES, Anders SH, Lehmann CU. Electronic Health Record Interactions through Voice: A Review. Appl Clin Inform 2018 Jul;9(3):541-552 [FREE Full text] [doi: 10.1055/s-0038-1666844] [Medline: 30040113]
56. Leach LS, Butterworth P, Poyser C, Batterham PJ, Farrer LM. Online Recruitment: Feasibility, Cost, and Representativeness in a Study of Postpartum Women. J Med Internet Res 2017 Mar 08;19(3):e61-552 [FREE Full text] [doi: 10.2196/jmir.5745] [Medline: 28274906]

57. Haste A, Adamson AJ, McColl E, Araujo-Soares V, Bell R. Problems recruiting and retaining postnatal women to a pilot randomised controlled trial of a web-delivered weight loss intervention. BMC Res Notes 2018 Mar 27;11(1):203 [FREE Full text] [doi: 10.1186/s13104-018-3305-x] [Medline: 29587868]
58. Silfee VJ, Lopez-Cepero A, Lemon SC, Estabrook B, Nguyen O, Wang ML, et al. Adapting a Behavioral Weight Loss Intervention for Delivery via Facebook: A Pilot Series Among Low-Income Postpartum Women. JMIR Form Res 2018 Sep 10;2(2):e18 [FREE Full text] [doi: 10.2196/formative.9597] [Medline: 30684423]

Abbreviations

ACOG: American College of Obstetricians and Gynecologists
SMILE: Self-Management Intervention–Life Essentials
TAM: technology acceptance model

Edited by G Eysenbach; submitted 13.02.20; peer-reviewed by T Muto, T Bickmore, J Ayre, M Nomali; comments to author 20.04.20; revised version received 10.06.20; accepted 17.01.21; published 01.03.21.

Please cite as:
Militello L, Sezgin E, Huang Y, Lin S
Delivering Perinatal Health Information via a Voice Interactive App (SMILE): Mixed Methods Feasibility Study
JMIR Form Res 2021;5(3):e18240
URL: https://formative.jmir.org/2021/3/e18240
doi:10.2196/18240
PMID:33646136

Original Paper

A Fitness App for Monitoring Walking Behavior and Perception (Runkeeper): Mixed Methods Pilot Study

Justin B Hollander[1], PhD; Sara C Folta[2], PhD; Erin Michelle Graves[3], PhD; Jennifer D Allen[4], PhD; Minyu Situ[5], MSc

[1]Department of Urban and Environmental Policy and Planning, School of Arts and Sciences, Tufts University, Medford, MA, United States

[2]Friedman School of Nutrition Science and Policy, Tufts University, Medford, MA, United States

[3]Federal Reserve Bank of Boston, Boston, MA, United States

[4]Department of Community Health, Tufts University, Medford, MA, United States

[5]Graduate Program of Community and Regional Planning, School of Architecture, University of Texas Austin, Austin, TX, United States

Corresponding Author:
Justin B Hollander, PhD
Department of Urban and Environmental Policy and Planning
School of Arts and Sciences
Tufts University
97 Talbot Avenue
Medford, MA, 02155
United States
Phone: 1 6176273394
Email: justin.hollander@tufts.edu

Abstract

Background: Physical activity has a strong positive impact on both physical and mental health, and public health interventions often encourage walking as a means to promote physical activity. Social connectivity, such as that among spouses, families, friends, and colleagues, highly influences physical activity. Although technology-based interventions have some influence on human behavior, they have not been fully implemented and evaluated for their influence on walking through social connectivity.

Objective: We aimed to pilot-test the organization of neighborhood walking clubs and use of a mobile app (Runkeeper) to encourage social connectedness and neighborhood cohesion, as well as to increase physical activity.

Methods: We used a convenience sampling method to recruit 46 adults from an urban location in Greater Boston, Massachusetts. We assigned participants to teams based on their geographic location and neighborhood and required them to use the app (Runkeeper). Participants completed 2 self-administered web-based surveys before and after the intervention period. The surveys included standard measures to evaluate physical activity, social connectedness, perceived social support, and neighborhood cohesion (Buckner Neighborhood Cohesion Scale) before and after the intervention. Following the intervention, we randomly selected 14 participants to participate in postintervention, in-depth phone interviews to gain an understanding of their experiences.

Results: This study was approved by the institutional review board in June 2018 and funded in January 2018. Recruitment started in May 2019 and lasted for 2 months. Data were collected from July 2019 to January 2020. In this study, Runkeeper was of limited feasibility as an app for measuring physical activity or promoting social connectedness. Data from the app recorded sparse and uneven walking behaviors among the participants. Qualitative interviews revealed that users experienced difficulties in using the settings and features of the app. In the questionnaire, there was no change between pre-post assessments in walking minutes (b=−.79; 95% CI −4.0 to 2.4; P=.63) or miles (b=−.07; 95% CI −0.15 to 0.01; P=.09). We observed a pre-post increase in social connectedness and a decrease in neighborhood cohesion. Both quantitative and qualitative results indicated that the psychosocial aspects of walking motivated the participants and helped them relieve stress. Interview results showed that participants felt a greater virtual connection in their assigned groups and enhanced connections with friends and family members.

Conclusions: Our study found that Runkeeper created a virtual connection among walking group members and its data sharing and ranking motivated walking. Participants felt that walking improved their mental health, helped to relieve stress, and made them feel more connected with friends or family members. In future studies, it will be important to use an app that integrates with a wearable physical activity device. There is also a need to develop and test intervention components that might be more effective in fostering neighborhood cohesion.

(JMIR Form Res 2021;5(3):e22571) doi:10.2196/22571

KEYWORDS

physical activity; smartphone; mobile app; sense of belongingness; community cohesion

Introduction

Physical Activity

Physical activity (PA) has been found to be significantly associated with not only physical health and chronic disease risk reduction [1] but also with psychological health and well-being [2,3]. The World Health Organization calls for interventions to promote PA as a means of lowering the global burden of chronic disease, with group walking thought to potentially increase the motivation to engage in PA [4,5]. Despite the known health benefits of PA, just over half of the adults in United States (51%) meet the 2008 federal guidelines for PA [6]. Many practitioners have focused on walking as a form of PA that is inexpensive, accessible, and well accepted among adults [7-9]. Walking is the most common form of PA in the United States [4]. It has a low risk of injury [10], making it safe for all age groups and feasible even in the presence of other health concerns or logistical barriers to exercise. Interventions to increase moderately intense walking among Americans could help meet the PA guidelines [11].

In general, walking has multiple positive effects on health and well-being. It improves physical health by improving blood pressure control [12-14], weight loss [12,15], and prevention of obesity and cardiovascular diseases [16-18]. Research has shown that walking can affect other dimensions of well-being, such as reducing depression [19], lowering physiological stress [20], and stabilizing cognitive functioning for those at risk of dementia [21]. Sedentary people can benefit from modest increases in walking, especially for otherwise healthy, middle-aged or older adults.

Group Walking

There is some evidence that walking in groups influences the motivation to walk. Ball et al [22] found that the likelihood of an individual walking for physical exercise increases when they are doing so with another person. People also prefer and enjoy walking with others more than walking alone [8]. The benefits of PA are realized through group walking. Meta-analyses of walking group interventions showed statistically significant improvements in blood pressure, resting heart rate, and body mass index, among other measures [12,23]. Most of the studies report on the physical benefits of walking in groups [23], with some research measuring emotional well-being. These generally show a reduction in depression scores [12,24], although one recent study found that individual walking is more effective at reducing depression [25]. Marselle et al [11] found that walking in groups contributed more to personal emotional and mental well-being than walking individually.

Walking clubs can potentially improve social connectivity [26]. Chen and Pu [26] reported that walking in teams promotes social connections with friends. Several recent studies also found positive effects for participating in walking clubs, including among specific populations including postpartum women [27], adults with diabetes, and older adults [28-30]. A number of researchers have reported that the social dimensions of a walking group contribute to people's interest in the initialization of walking, maintaining participation, and increasing walking behavior [31,32]. Others raise concern that walking clubs may disproportionately serve traditionally advantaged subpopulations, although walking clubs in the United States may be relatively more inclusive than other countries [33].

Although walking programs have the potential to improve *social* well-being, defined as a sense of belonging and interdependence with others [34], the impact of group walking on social well-being is understudied relative to the larger literature on physical well-being. Qualitative research suggests that group walks have a positive effect on social well-being [35,36]. In their meta-analysis, Meads and Exley [23] report that *walking in groups tended to increase quality of life measures and may increase social connectedness, but the evidence for this was uncertain* [23]. Thus, although the literature supports the conclusion that walking clubs offer a number of individual health and psychological benefits, the social benefits are less well established.

Recent technological interventions have been shown to increase social connectivity and PA. However, few studies have explored the use of technology to facilitate neighborhood walking clubs as a PA intervention. Walsh et al [37] conducted a study to analyze the effect of mobile phone app intervention on increasing daily walking steps among youth. Their study pointed out that PA could be enhanced through specific settings in mobile phone apps, such as self-monitoring of walking steps and setting of personal walking goals. Chen and Pu [26] used a mobile app called *Healthy Together*, which enabled users to participate in physical activities together by sending each other messages and earning badges. Their research goal is to compare different social incentives in mobile fitness apps and to provide new angles of design implications for mobile fitness apps. They found that building users' performance with their team members promotes not only individuals' PA levels but also social connections with friends. They recommended that an app design for physical activities should consider adopting social interaction as the key motive for user involvement [26]. Some recent innovative electronics such as *FitBit* and *Jawbone* are wearable PA tracking technology, which enable users to connect their data of physical activities to their phone or app automatically [38]. Comparison to others could be perceived as motivating if people looked to those with greater PA success as positive role models, which is an upward comparison [38]. Women who had a strong interest in upward comparisons also presented significant increases in PA [39].

However, most of the above studies mainly focused on the benefit and motivation of walking at the personal level, instead of focusing on whether the walking clubs impact neighborhood social cohesion.

This study aims to examine the potential impact of walking clubs on neighborhood social cohesion. The goal of this study is to gather preliminary information on technologically enhanced walking clubs to encourage social connectedness and neighborhood cohesion and to increase PA to help guide future studies that will advance public health not only at the personal level but also at the community level.

Methods

In this mixed methods pilot study, we used a pre-post evaluation design to assess the potential impact of the intervention Runkeeper on walking distance and duration, belongingness, neighborhood cohesion, motivation for PA, and self-reported PA. Qualitative methods were used to contextualize quantitative results. This study was approved by the institutional review board in June 2018 and was funded in January 2018. Recruitment started in May 2019 and lasted for 2 months. Data collection started in July 2019 and ended in January 2020.

Sample and Setting

The inclusion criteria were living in an urban location in Greater Boston, Massachusetts, aged 18 years or older and ability to speak and read English. We used geographic information systems and census track-level demographic and socioeconomic data to identify neighborhoods for the walking clubs, with the aim of reaching a diverse set of participants across income, race, ethnicity, and age. The final locations were determined based on the number of potentially eligible participants and their proximity to Tufts University.

The procedure involves 4 main steps. First, participants were recruited and administered a mandatory baseline survey (May-June 2019); the second step was to monitor participants' walking data coming from the app for 3 months (July-October 2019); the third step was conducting a mandatory exit survey to gain participant feedback (October 2019); and the last step was interviewing participants in terms of the data reflection (January 2020). The baseline survey and exit survey took participants 30 min to complete, and the questions in the baseline survey were identical to all questions in the exit survey, and they were both 9-page surveys and each page had 4 questionnaire items.

Participants were recruited through Facebook advertisements and door-to-door invitations in the neighborhoods of interest. Recruitment took 1.5 months, which involved recruiting participants and administering a baseline survey (May 21 to June 30, 2019). When potential participants responded to our Facebook advertisement, we would confirm their home addresses first to see whether they were in our desired location before sending them informed consent agreement. If they agreed in person (during door-to-door recruitment), they signed a hard copy of the informed consent agreement and provided an email address for further contact. No personal information was collected. Upon study entry, each participant was given a unique study identification number, which was only for data organization, management, and analysis. This identification number was associated with all data records from both surveys and apps. Participants' names or other identifying information was kept separate from this identification number. The survey was hosted and all online data were stored on a secure server, and the hard copy of the consent form was stored in a locked office.

After recruitment, our researchers were in charge of inviting participants to take both baseline and exit surveys by sending emails. We used an open survey approach where the survey links sent to participants were the same, and the technical functionality of the electronic questionnaire was tested before fielding the survey via Qualtrics. Participants were able to change their answers through a back button, which displays a summary of the responses. The survey never displayed a second time once the participants had completed it. In addition, only completed surveys were analyzed, and the corresponding participants were allowed to continue accepting the walking club intervention.

During the baseline survey, all eligible participants received the walking club intervention. We purposefully divided them into 3 groups, of which the geographic radius of each group was within a walkable distance (10-min walking time). Each group had an average of 15 participants. Each group was asked for a volunteer to serve as a captain. The captain was responsible for serving as a liaison with the Tufts research team and organizing several walks per month for the group. Recent research has proved that the frequency of communication across teams is significantly associated with maintaining walking steps [40]. A participant volunteered to be captain in two of the groups, and the study coordinator served as the captain in the third group. In addition, participants received a US $100 gift card for completing both pre- and postsurveys and for 3-month monitoring of walking, as tracked by Runkeeper. Those who participated in phone interviews following the intervention were provided with an additional US $50 gift card.

Walking clubs were facilitated using Runkeeper, a free mobile fitness app that allows users to keep track of distance traveled, calories burned, and time spent engaging in physical activities. The app separates statistics by week, month, and year so that people are able to track their short- and long-term progression. In addition, Runkeeper has social features that set it apart from other fitness trackers. It is possible to create groups within the app in which the members can interact. Users can scroll through a feed and leaderboard that details their group mates' PA, and they can like, post, and comment. A feature of the app is that it is necessary to manually start and stop exercise occasions for the app to log them. Participants were required to log walks on Runkeeper at least twice per month (6 times in total) to be considered adherent to the intervention.

Quantitative Data Collection

Runkeeper uses GPS technology to record walking distance in miles (not steps). Participants manually started and stopped the app for each PA occasion by selecting a start function on the app to track walking time data. The research team had access to app data on the aggregate minutes and miles walked per participant for each of the 12 intervention weeks. Data were analyzed using a mixed effects regression model with participant ID as the random effect. The survey data were analyzed using paired sample one-tailed *t* tests. Stata version 15 was used for

all analyses. We also had data from the app on the number of walks that occurred each week. These data were used to help classify participants as adherent or not, defined as at least two walks per month. Sensitivity analyses were conducted using data from adherent participants.

Before and after the 3-month intervention period, all participants received a web-based mandatory survey (Qualtrics). Self-reported PA was measured using the International Physical Activity Questionnaire (IPAQ), long form, which provides an estimate of walking, moderate-intensity, and vigorous-intensity activity within each of the following domains: work, transportation, domestic chores and gardening (yard), and leisure time [41]. Metabolic equivalent (MET) minutes per week were compiled according to the IPAQ's Guidelines for Data Processing and Analysis [42].

The pre-post survey included questions about motivation for PA using the Physical Activity and Leisure Motivation Scale [43]. Its 8 subscales are (1) competition or ego, (2) appearance, (3) other's expectations, (4) affiliation, (5) physical condition, (6) psychological condition, (7) mastery, and (8) enjoyment. Each subscale has 4 items, measured on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Higher scores reflected higher motivation levels.

Social factors were also included in the pre-post survey. Social connectedness was measured using the General Belongingness Scale [44]. This scale has 2 domains, each with 6 items: acceptance or inclusion and rejection or exclusion. The rejection or exclusion subscale is reverse-coded so that higher scores reflect greater levels of perceived acceptance or social connectedness. Neighborhood cohesion was measured using the Buckner Neighborhood Cohesion Scale [45]. This validated scale has 18 items that measure 3 constructs: attraction to neighborhood, neighboring, and psychological sense of community [46,47]. *Attraction to neighborhood* measures resident's degree of willingness to remain in the neighborhood, *neighboring* is the perceived degree of interaction between residents in the neighborhood, and *psychological sense of community* generally refers to the subjective sense of community that residents felt within the neighborhood [45].

We ran tests of significance to ensure internal validity of the data collected from both surveys and apps. First, we ran descriptive statistics and all variables of interest. Then, we ran the inferential statistics (regression) to explore the cause of the relationship between the independent and dependent variables.

Qualitative Data Collection

We used postintervention qualitative interviews to help assess engagement with the intervention, usability of the app, and contextualize the quantitative results. On the basis of the 3-months monitoring and final data collected from the Runkeeper and surveys, we randomly chose 14 participants to conduct a phone call interview. The interviewees included 2 captains, 2 active (met adherence criteria) and 2 inactive (did not meet adherence criteria) participants in each group. The interviews were conducted in January 2020 and averaged 25 min in length. All participants were asked to conduct a phone interview with our research assistants through the FreeConferenceCall app [48]. The audio content of each interview was recorded with the permission of the interviewees.

Our recorded audio files of the interviews were transcribed verbatim using GoTranscript [49]. The data were analyzed using qualitative content analysis [50,51]. The initial codebook was mainly based on the questioning route and was refined based on team discussion. The final themes were built based on the patterns and frequencies found in the data, and all final codes were established at 90% agreement. NVivo (version 12, QSR International) was used to assist in the process of coding and analysis.

Results

Final Sampling

In total, we recruited 46 participants: 34 from the Facebook advertisement, 5 from door-to-door recruitment, and 7 from friends' recommendations. A total of 7 were identified as males and 39 as females (Multimedia Appendix 1). The participation rate of the baseline survey was 100% (43/43), and the completion rate of the exit survey was 100% (43/43).

PA

On the basis of the data collection from the app, the results indicated no increase in either minutes (b=−.79; 95% CI −4.0 to 2.4; P=.63) or miles (b=−.07; 95% CI −0.15 to 0.01; P=.09) walked over the intervention weeks. According to the IPAQ, there was a significant decrease in moderate MET minutes per week from pre- to postintervention; no other intensity categories were significant (Multimedia Appendix 2). In the sensitivity analysis, pre-post total and walking MET minutes per week were positive but remained nonsignificant. The pre-post change in psychological condition as a form of motivation (engaging in PA to relax and cope with stress) was positive and significant (Multimedia Appendix 2). There were no significant changes in any of the other types of motivation. Changes in physical condition and affiliation (doing it with others) motivations changed direction in the sensitivity analysis and remained nonsignificant.

Qualitative Analysis

Themes from the interviews provide further insight into walking and motivation (Multimedia Appendix 3). Although not reflected in the survey data, one theme suggests that participants were motivated by the competitive aspect of the walk ranking on the app. Participants also reported that being organized in groups to walk enhanced feelings of accountability, competitiveness, and peer pressure.

> *I felt a little positive peer pressure.*
>
> *I feel like when I had to be accounted for it, I did more, but when I wasn't being accounted for it, I did less.*

A theme from the interviews supports and provides context for the change in psychological condition as a form of motivation. Many considered walking to be the best way of releasing stress, especially after their daily work.

> *Anytime I've had maybe an argument with someone or just had a stressful day, like going for a walk does help.*

In addition, many participants described walking as *meditative* and described inner peace as the most meaningful benefit of walking.

> *It was nice because it's very peaceful. Walking, to me, is almost meditative. You get to be in the moment. I guess you can be mindful.*

Themes from the interviews further suggest that participants were avid walkers but not avid app users. Many said that they walked more frequently than suggested by the app and often forgot to manually start it.

> *I probably used the app 25% of the time while walking.*

This was a particularly strong theme among participants who did not meet the adherence criteria.

> *I would forget about [using] it, that would be the honest answer...Then I would realize, what I'm sending you guys is probably just a small percentage of what I actually do because the app didn't catch my walk accurately.*

Although many participants appreciated the tracking features of the app, most were frustrated by the manual operations of it.

> *I would forget to turn it on or turn it on midway or turn it off for one part and then I would stop walking and then forget to turn it back on.*

Some said they preferred wearable PA monitors that did not require this. Themes related to the functions and operation of the app were similar between the adherent and nonadherent groups. However, nonadherent participants described more barriers to PA, such as work and family issues.

> *I would say that I tried to walk every week, or at least every two weeks...I am just too busy sometimes with the two kids to fit in that, to fit in a walk. I guess it would be more of a time thing more than a not wanting to.*

General belongingness increased significantly from pre- to postintervention, whereas neighborhood cohesion decreased (Multimedia Appendix 2). The total scores and all subscores were significant for both measures. There were no changes in direction or significance based on the sensitivity analysis.

In the interviews, a theme from both adherent and nonadherent participants was that they enjoyed walking by themselves. Nonadherent participants did not mention walking with others. Adherent participants said that they preferred to walk with friends or family members if they needed companionship.

> *I would walk with my mom.*

> *I walked with that one friend every time.*

When they walked with others, it was described as an opportunity to enhance these established relationships.

> *It was nice because our girls could catch up on what's going on with our lives...I think it made us more connected because usually when we get together, we go out once a week, but this time, we were getting together two, three times a week...*

Regarding neighborhood cohesion, participants mostly expressed that walking enhanced their appreciation for the physical and natural features of their neighborhood.

> *I like looking at people's houses, backyards or front yards and things like that. So, I do feel a greater sense of belonging to my neighborhood or connection to my neighborhood.*

Some participants also felt more virtually connected with their neighbors because of the chat platform on the app.

> *I feel a little bit more connected to the neighborhood because the people that were in my group that we spoke during tech-wise lived around you.*

These participants also revealed a lack of belongingness to their physically existing community.

> *I also feel like I might be in a little bit of a different mindset since I'm a college student and don't really view [city] as my home, home. Whereas some of the people in the study, this is their full-time residence and they know their neighbors and things like that.*

Discussion

In this study, we deepened the insight into the social and health benefits derived from participating in a neighborhood walking club facilitated by a technical intervention.

According to the questionnaire, there was no change between the pre-post assessments in walking minutes or miles. However, both quantitative and qualitative results indicated that the psychosocial aspects of walking motivated them and helped relieve stress. The interview results showed that they considered the most meaningful benefit of walking outside to be that it allowed them to relieve daily stress and find some form of inner peace. Their feedback is consistent with recent research that outdoor walking is a way of restoring mood and releasing stress [52,53].

Walking Patterns and Technology

Neither app data nor self-report (IPAQ) indicated an increase in PA. However, based on interviews, the need to manually start and stop the app resulted in inaccurate logging of activity time. This finding is aligned with suggestions raised in other studies that an app is best designed to integrate with wearable PA tracking devices [26,38]. In addition, interviewees described themselves as avid walkers, suggesting a ceiling effect that would make it difficult for the intervention to further increase walking behavior. Preintervention MET minutes of activity indicated a highly active cohort. Even *nonadherence* likely refers to nonuse of the app rather than inactivity, based on the qualitative results.

Social Connectedness and Technology

Research suggests that with respect to external interventions that try to modify human habits, people are generally interested in changing their behavior in a particular environment only if

they find that their energy and time are not wasted and that there are more benefits or motivations generated. In our sensitivity analysis, the motivation of the competition and ego did not change, but our qualitative results showed how their motivation for competition played a role in their participation in walking. Interviewees' motivation to increase walking was stimulated by the sharing of ranking data in their corresponding group on the app, which generated feelings of competitiveness. This is consistent with the findings of Chen and Pu [26] that linking users' performance with their teammates promotes PA. The feature of ranking in the app provides a solution in dealing with *the absence of effective tools in motivating members in walking clubs* [54].

Interviewees' reflections on walking data sharing and ranking are consistent with the literature that sharing PA data and ranking among groups in apps stimulates social comparison [55-57]. In addition, according to the findings of Arigo et al [39], women's strong interest in upward comparison motivates their physical activities. Our interviewees were all female, and some of them who had a feeling of competition were from our active app users. Comparing individual ranking data in a group increases individuals' likelihood of PA, which provides another possibility of PA motivation: that social support comes from not only those they are familiar with [58] but also from social cohesion through community connections [59].

In our quantitative analysis, neighborhood cohesion significantly decreased from the beginning of the study to the end, whereas general belongingness increased. On the basis of the qualitative analysis, our interviewees did not have a strong feeling of connectedness to their physical communities, although they felt a greater virtual connection in their assigned groups. The qualitative results provide an explanation for these results. Although our entire walking club was built at a geographic community around Tufts University, group members relied on the virtual connection generated from the chat feature on Runkeeper rather than meeting in person for walks. This is consistent with the argument raised by Gusfield [60] that community can also be defined by *human relationship* rather than geographic boundaries. The feeling of virtual connection in our study should be noted as a new *sense of community*, which is related to the findings from McMillan and Chavis [61] that a *sense of community* would be shaped by feelings of membership and shared emotional connection. In addition, the walking incentive generated by the ranking competition is consistent with the argument raised by Wood et al [62] that walking for purpose was associated with a sense of community. The walking clubs built in our study and the mechanism of ranking in the app may create a feeling of membership and shared competitive connection for our participants, which might also result in the *sense of online community* coming up. However, as members know that their online community was also all residents of the same geographic communities, it is not possible to know if the sense of online community would have developed without the underlying knowledge of shared geography.

Interviewees indicated that they felt a physical connection with their neighborhood, which would not be particularly well captured by the Neighborhood Cohesion Scale. None of our interviewees indicated feeling physically closer to their fellow club members who were strangers to them before the study. Rather, they did feel close ties with friends or family members in real life after this study, even though their friends or family members did not participate. In our sensitivity analysis, the motivation of the affiliation (ie, walking with others) motivations changed direction. This reflects that participants tend to stay with their familiar connections, which validates the research conducted by Chen and Pu [26], who recruited people and their friends together to increase participants' motivation for involvement. Chen and Pu [26] pointed out that technical interventions for physical activities should consider adopting social interaction. On the basis of our interviewees' experience, social interaction in PA should be built upon established relationships rather than the interactions of strangers, a point that has been overlooked in previous literature.

Limitations

Given that the number of participants was limited in this pilot study, our findings cannot be generalized, and selection bias cannot be ruled out. Although our chosen mobile app has fully accessible open data, it had complicated settings and no auto-tracking, which led to inaccurate data. In addition, the intervention period lasted for 3 months during summer, and 3 months may not be long enough to observe a measurable impact on behavior. However, the fact that the intervention took place in summer may overestimate walking behavior, as people are more active in the summer. The long-term effects of community-level cohesion and participants' various walking patterns and motivations throughout the year could not be tracked in this study.

Conclusions

This pilot study aims to gather preliminary information on technology-based interventions for walking to promote social connectedness and neighborhood cohesion and to increase PA.

According to participants' feedback, walking benefits were not exclusive to physical health, as mental health and inner peace were addressed as well. Given that not only our study but also other scholars have observed that walking has an impact on mental health improvement, the effective promotion of walking clubs should be related to the latest topics in mental health and stress release. The number of participants recruited from Facebook was much larger than that from on-site recruitment, which indicates that traditional recruitment was not as effective as the internet-based method. The mechanism of recruitment conducted by Chen and Pu [26], asking participants to invite at least one person to join them, had some success in this study.

By adopting the intervention of a mobile app, the *sense of online community* appeared in our findings; it appeared that place-based connections that might enhance neighborhood cohesion were not formed. The feeling of competitiveness derived from online data sharing was found in our study as a new incentive for walking. However, it was not the direct motivation for social connections among group members. The mobile app created a virtual connection among walking group members, and its data sharing and ranking generated new motivations for individual walking. However, it was difficult to build social connections

effectively in walking clubs. On the basis of our qualitative analysis, the ties between participants and their friends were enhanced because they walked together, but they did not develop connections with other participants who were strangers to them before the study.

Technology-based interventions should be tailored to the majority's needs. Participants in our study expressed their need to use wearable auto-tracking fitness app or electronic and push notifications. The mobile app should be provided with communication timeliness and user convenience, which would increase the chance of changing human behavior of walking and social contact regarding technology-based interventions. Ogilvie et al [63] concluded that people would be able to walk more if interventions were targeted at their needs or motivations.

In summary, the technology-based intervention for walking effectively boosts individual-level well-being by creating virtual connections and feelings of competitiveness. Most participants discovered that walking helped to improve their mental health, release stress, and promote bonds with their friends or family members. However, different strategies are likely needed to promote community-level cohesion. One such mechanism might be asking participants to involve acquaintances in the technology-based intervention to help bridge the gap between virtual and physical connections.

The lessons learned from this study suggest areas for future research. To achieve the goal of neighborhood cohesion, it may be necessary to redesign the intervention or to build features that are more closely tied to the physical environment. There is clearly a need to try different apps to achieve our goals, most likely those that are compatible with a wearable fitness device. Future studies may focus on the user experience of wearing fitness devices, from which we can examine how social perceptions are derived from a technological intervention as well as any difference between those perceptions in the physical and virtual worlds. Our future research will also aim to investigate in-depth interactions between walking behavior and the corresponding mobile app or devices to discover or modify the interventions that best meet the needs of walking group participants.

Acknowledgments

This study was supported by a grant from the Office of the Vice Provost for Research at Tufts University and by the Federal Reserve Bank of Boston. The views expressed in this article are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of Boston or the Federal Reserve System.

The authors would like to thank the study participants in the Medford community. The authors would also like to thank Atrey Bhargava, Ruoyang Li, and Cyrus Hastings Miceli for their assistance with recruitment and data monitoring and Rachel Herman and Jennifer Han for their assistance with the interviews.

Authors' Contributions

JH conceptualized and designed the study, interpreted the data, drafted the manuscript, and critically reviewed the manuscript. SF made substantial contributions to the design of the interview process, quantitative analysis of survey data, and manuscript drafts and reviews. EG made substantial contributions to the interview design, interview process, and manuscript review. JA critically reviewed the manuscript. MS was responsible for recruitment, collecting data, interviewing participants, organizing data, interpreting the qualitative data, and drafting the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Participant characteristics preintervention.
[DOCX File , 14 KB - formative_v5i3e22571_app1.docx]

Multimedia Appendix 2
Pre-post survey results: neighborhood cohesion, belonging, physical activity motivation, and self-reported physical activity.
[DOCX File , 16 KB - formative_v5i3e22571_app2.docx]

Multimedia Appendix 3
Summary of interview findings.
[DOCX File , 17 KB - formative_v5i3e22571_app3.docx]

References

1. U.S. Department of Health and Human Services. 2018 Physical Activity Guidelines Advisory Committee Scientific Report. U.S. Department of Health and Human Services 2018:1-779 [FREE Full text]

2. Mao Z, Zhao X. The effects of social connections on self-rated physical and mental health among internal migrant and local adolescents in Shanghai, China. BMC Public Health 2012 Feb 03;12(1):97 [FREE Full text] [doi: 10.1186/1471-2458-12-97] [Medline: 22299776]
3. Penedo FJ, Dahn JR. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. In: Curr Opin Psychiatry LWW. Netherlands: Wolters Kluwer; 2005:189-193.
4. CDC VitalSigns - More people walk to better health internet. Cent Dis Control Prev. 2013. URL: https://www.cdc.gov/vitalsigns/walking/index.html [accessed 2020-10-17]
5. World Health Organization. More active people for a healthier world: global action plan on physical activity 2018-2030. In: World Health Organization. Geneva: World Health Organization; 2018.
6. Centers for Disease Control Prevention. Leisure-time physical activity. Centers for Disease Control and Prevention 2015:11. [doi: 10.1007/springerreference_308179]
7. Diehr P, Hirsch C. Health benefits of increased walking for sedentary, generally healthy older adults: using longitudinal data to approximate an intervention trial. In: J Gerontol A Biol Sci Med Sci. Health Benefits of Increased Walking for Sedentary, Generally Healthy Older Adults: Using Longitudinal Data to Approximate an Intervention Trial. J Gerontol Ser A Oxford Academic; Sep 2010:982-989.
8. Johansson M, Hartig T, Staats H. Moderation by company and outdoor environment. In: Appl Psychol Health Well-Being. New Jersey: John Wiley & Sons, Ltd; Nov 1, 2011:261-280.
9. Simpson M, Serdula M, Galuska D, Gillespie C, Donehoo R, Macera C, et al. Walking trends among U.S. adults. Am J Preven Med 2003 Aug;25(2):95-100. [doi: 10.1016/s0749-3797(03)00112-0]
10. US Department of Health and Human Services. Physical Activity Guidelines for Americans, 2nd edition. US Department of Health and Human Services 2018:1-118 [FREE Full text]
11. Marselle M, Irvine K, Warber S. Walking for well-being: are group walks in certain types of natural environments better for well-being than group walks in urban environments? Int J Environ Res Public Health 2013 Oct 29;10(11):5603-5628 [FREE Full text] [doi: 10.3390/ijerph10115603] [Medline: 24173142]
12. Hanson S, Jones A. Is there evidence that walking groups have health benefits? A systematic review and meta-analysis. Br J Sports Med 2015 Jun;49(11):710-715 [FREE Full text] [doi: 10.1136/bjsports-2014-094157] [Medline: 25601182]
13. Lee L, Watson MC, Mulvaney CA, Tsai C, Lo S. The effect of walking intervention on blood pressure control: a systematic review. Int J Nurs Stud 2010 Dec;47(12):1545-1561. [doi: 10.1016/j.ijnurstu.2010.08.008] [Medline: 20863494]
14. Murphy MH, Nevill AM, Murtagh EM, Holder RL. The effect of walking on fitness, fatness and resting blood pressure: a meta-analysis of randomised, controlled trials. Prev Med 2007 May;44(5):377-385. [doi: 10.1016/j.ypmed.2006.12.008] [Medline: 17275896]
15. Richardson CR, Newton TL, Abraham JJ, Sen A, Jimbo M, Swartz AM. A meta-analysis of pedometer-based walking interventions and weight loss. Ann Fam Med 2008 Jan 01;6(1):69-77 [FREE Full text] [doi: 10.1370/afm.761] [Medline: 18195317]
16. Hamer M, Chida Y. Walking and primary prevention: a meta-analysis of prospective cohort studies. Br J Sports Med 2008 Apr;42(4):238-243. [doi: 10.1136/bjsm.2007.039974] [Medline: 18048441]
17. Lee I, Buchner D. The importance of walking to public health. Med Sci Sports Exerc 2008 Jul;40(7 Suppl):512-518. [doi: 10.1249/MSS.0b013e31817c65d0] [Medline: 18562968]
18. Pucher J, Buehler R. Walking and cycling for healthy cities. Built Environment; Alexandrine Press 2010 Dec 05;36(4):391-414. [doi: 10.2148/benv.36.4.391]
19. Robertson R, Robertson A, Jepson R, Maxwell M. Walking for depression or depressive symptoms: a systematic review and meta-analysis. Database Abstr Rev Eff DARE Qual-Assess Rev Internet. London: Centre for Reviews and Dissemination (UK); 2012. URL: https://www.ncbi.nlm.nih.gov/books/NBK99429/ [accessed 2021-02-09]
20. Hartig T, Evans GW, Jamner LD, Davis DS, Gärling T. Tracking restoration in natural and urban field settings. J Env Psychol 2003 Jun;23(2):109-123. [doi: 10.1016/s0272-4944(02)00109-3]
21. Smith DH, Johnson VE, Stewart W. Chronic neuropathologies of single and repetitive TBI: substrates of dementia? Nat Rev Neurol 2013 Apr 5;9(4):211-221 [FREE Full text] [doi: 10.1038/nrneurol.2013.29] [Medline: 23458973]
22. Ball K, Bauman A, Leslie E, Owen N. Perceived environmental aesthetics and convenience and company are associated with walking for exercise among Australian adults. Prev Med 2001 Nov;33(5):434-440. [doi: 10.1006/pmed.2001.0912] [Medline: 11676585]
23. Meads C, Exley J. A systematic review of group walking in physically healthy people to promote physical activity. Int J Technol Assess Health Care 2018 Jan 17;34(1):27-37. [doi: 10.1017/s0266462317001088]
24. Hanson S, Cross J, Jones A. Promoting physical activity interventions in communities with poor health and socio-economic profiles: a process evaluation of the implementation of a new walking group scheme. Soc Sci Med 2016 Nov;169:77-85 [FREE Full text] [doi: 10.1016/j.socscimed.2016.09.035] [Medline: 27693974]
25. Mau M, Rasmussen KW, Jacobsen M, Roessler KK. Walking against depression: a brief report. SDU Research Portal 2020:6 [FREE Full text]

26. Chen Y, Pu P. HealthyTogether: exploring social incentives for mobile fitness applications. In: Proceedings of the Second International Symposium of Chinese CHI. 2014 Presented at: Chinese CHI '14: The Second International Symposium of Chinese CHI; April, 2014; Toronto Ontario Canada. [doi: 10.1145/2592235.2592240]
27. Keller C, Ainsworth B, Records K, Todd M, Belyea M, Vega-López S, et al. A comparison of a social support physical activity intervention in weight management among post-partum Latinas. BMC Public Health 2014 Sep 19;14(1):971 [FREE Full text] [doi: 10.1186/1471-2458-14-971] [Medline: 25233867]
28. Hughes J, Riddlesworth T, DiMeglio L, Miller K, Rickels M, McGill J, T1D Exchange Clinic Network. Autoimmune diseases in children and adults with Type 1 diabetes from the T1D Exchange Clinic Registry. J Clin Endocrinol Metab 2016 Dec;101(12):4931-4937 [FREE Full text] [doi: 10.1210/jc.2016-2478] [Medline: 27676394]
29. Valenti G, Bonomi AG, Westerterp KR. Walking as a contributor to physical activity in healthy older adults: 2 week longitudinal study using accelerometry and the doubly labeled water method. JMIR Mhealth Uhealth 2016 Jun 07;4(2):56 [FREE Full text] [doi: 10.2196/mhealth.5445] [Medline: 27268471]
30. Ehn M, Eriksson LC, Åkerberg N, Johansson A. Activity monitors as support for older persons' physical activity in daily life: qualitative study of the users' experiences. JMIR Mhealth Uhealth 2018 Feb 01;6(2):34 [FREE Full text] [doi: 10.2196/mhealth.8345] [Medline: 29391342]
31. Kassavou A, Turner A, French DP. Do interventions to promote walking in groups increase physical activity? A meta-analysis. Int J Behav Nutr Phys Act 2013 Feb 06;10(1):18 [FREE Full text] [doi: 10.1186/1479-5868-10-18] [Medline: 23388115]
32. Wensley R, Slade A. Walking as a meaningful leisure occupation: the implications for occupational therapy. Br J Occup Therapy 2012 Feb;75(2):85-92. [doi: 10.4276/030802212x13286281651117]
33. Stylianou M, van der Mars H, Kulinna P, Adams M, Mahar M, Amazeen E. Before-school running/walking club and student physical activity levels: an efficacy study. In: Res Q Exerc Sport. Before-School Running/Walking Club and Student Physical Activity Levels: An Efficacy Study. Res Q Exerc Sport Routledge; Dec 2016:342-353.
34. Thompson K. Suicide. In: Emile Durkheim. Oxfordshire, United Kingdom: Taylor & Francis; 2002.
35. Hynds H, Allibone C. What motivates people to participate in organised walking activity. Nat Engl Res Rep 2009:28 [FREE Full text]
36. Villalba VDL, Cacace M, Nolte E, Sach T, Fordham R, Suhrcke M. Costing the walking for health programme. Natural England 2012 [FREE Full text]
37. Walsh J, Corbett T, Hogan M, Duggan J, McNamara A. An mhealth intervention using a smartphone app to increase walking behavior in young adults: a pilot study. JMIR Mhealth Uhealth 2016 Sep 22;4(3):109 [FREE Full text] [doi: 10.2196/mhealth.5227] [Medline: 27658677]
38. Arigo D. Promoting physical activity among women using wearable technology and online social connectivity: a feasibility study. Health Psychology and Behavioral Medicine 2015 Dec 31;3(1):391-409. [doi: 10.1080/21642850.2015.1118350]
39. Arigo D, Schumacher LM, Pinkasavage E, Butryn ML. Addressing barriers to physical activity among women: a feasibility study using social networking-enabled technology. Digit Health 2015 Nov 24;1 [FREE Full text] [doi: 10.1177/2055207615583564] [Medline: 29942539]
40. Hamamatsu Y, Ide H, Kakinuma M, Furui Y. Maintaining physical activity level through team-based walking with a mobile health intervention: cross-sectional observational study. JMIR Mhealth Uhealth 2020 Jul 03;8(7):e16159 [FREE Full text] [doi: 10.2196/16159] [Medline: 32618576]
41. Hagströmer M, Oja P, Sjöström M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. Public Health Nutr 2006 Sep;9(6):755-762. [doi: 10.1079/phn2005898] [Medline: 16925881]
42. IPAQ Group. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ). International Physical Activity Questionnaire. URL: http://www.ipaq.ki.se
43. Molanorouzi K, Khoo S, Morris T. Validating the Physical Activity and Leisure Motivation Scale (PALMS). BMC Public Health 2014 Sep 03;14:909 [FREE Full text] [doi: 10.1186/1471-2458-14-909] [Medline: 25182130]
44. Malone GP, Pillow DR, Osman A. The General Belongingness Scale (GBS): assessing achieved belongingness. Personality and Individual Differences 2012 Feb;52(3):311-316. [doi: 10.1016/j.paid.2011.10.027]
45. Buckner JC. The development of an instrument to measure neighborhood cohesion. Am J Community Psychol 1988;16(6):791. [doi: 10.1007/bf00930892]
46. Li C, Hsu P, Hsu S. Assessing the application of the neighborhood cohesion instrument to community research in East Asia. J Community Psychol 2011 Oct 11;39(8):1031-1039. [doi: 10.1002/jcop.20448]
47. Robinson D, Wilkinson D. Sense of community in a remote mining town: validating a neighborhood cohesion scale. Wiley Online Library 1995;23(1):137-148. [doi: 10.1007/bf02506926]
48. FreeConferenceCall. FreeConferenceCall. 2020. URL: https://www.freeconferencecall.com/ [accessed 2020-10-17]
49. GoTranscript: transcription services; transcribe audio/video to text internet. https://gotranscript.com/. 2020. URL: https://gotranscript.com/ [accessed 2020-10-17]
50. Miles MB, Huberman AM, Saldaña J. Qualitative Data Analysis: a sourcebook of new methods. Thousand Oaks, California, United States: SAGE Publications Ltd; 2013:1-408.
51. Saldaña J. The Coding Manual for Qualitative Researchers. Thousand Oaks, California: SAGE Publications Ltd; 2015.

52. Olafsdottir G, Cloke P, Schulz A, van Dyck Z, Eysteinsson T, Thorleifsdottir B, et al. Health benefits of walking in nature: a randomized controlled study under conditions of real-life stress. Environment and Behavior 2018 Sep 28;52(3):248-274. [doi: 10.1177/0013916518800798]
53. Song C, Ikei H, Park B, Lee J, Kagawa T, Miyazaki Y. Correction: Song, C.; Ikei, H.; Park, B.J.; Lee, J.; Kagawa, T.; Miyazaki, Y. Psychological benefits of walking through forest areas. 2018, 15, 2804. Int J Environ Res Public Health 2020 Feb 18;17(4) [FREE Full text] [doi: 10.3390/ijerph17041316] [Medline: 32085636]
54. Nguyên MN, Gauvin L, Martineau I, Grignon R. Sustainability of the impact of a public health intervention: lessons learned from the Laval walking clubs experience. In: Health Promot Pract. Sage Publications: Thousand Oaks, CA; Jan 2005:44-52.
55. Arigo D, Suls JM, Smyth JM. Social comparisons and chronic illness: research synthesis and clinical implications. Health Psychol Rev 2014;8(2):154-214. [doi: 10.1080/17437199.2011.634572] [Medline: 25053133]
56. Festinger L. A theory of social comparison processes. Human Relations 2016 Apr 22;7(2):117-140. [doi: 10.1177/001872675400700202]
57. Taylor M, Vlaev I, Maltby J, Brown G, Wood A. Improving social norms interventions: rank-framing increases excessive alcohol drinkers' information-seeking. Health Psychol 2015 Dec;34(12):1200-1203. [doi: 10.1037/hea0000237] [Medline: 26192384]
58. Carron A, Hausenblas H, Mack D. Social influence and exercise: a meta-analysis. In: J Sport Exerc Psychol. United States: Human Kinetics, Inc; 1996:A.
59. Zhang W, Ta VM. Social connections, immigration-related factors, and self-rated physical and mental health among Asian Americans. Soc Sci Med 2009 Jun;68(12):2104-2112. [doi: 10.1016/j.socscimed.2009.04.012] [Medline: 19427087]
60. Gusfield J. A critical response. New York: Harper & Row New York; 1975.
61. McMillan DW, Chavis DM. Sense of community: a definition and theory. J Community Psychol 1986 Jan;14(1):6-23. [doi: 10.1002/1520-6629(198601)14:1<6::aid-jcop2290140103>3.0.co;2-i]
62. Wood L, Frank LD, Giles-Corti B. Sense of community and its relationship with walking and neighborhood design. Soc Sci Med 2010 May;70(9):1381-1390. [doi: 10.1016/j.socscimed.2010.01.021] [Medline: 20189699]
63. Ogilvie D, Foster CE, Rothnie H, Cavill N, Hamilton V, Fitzsimons CF, Scottish Physical Activity Research Collaboration. Interventions to promote walking: systematic review. Br Med J 2007 Jun 09;334(7605):1204 [FREE Full text] [doi: 10.1136/bmj.39198.722720.BE] [Medline: 17540909]

Abbreviations

IPAQ: International Physical Activity Questionnaire
MET: metabolic equivalent
PA: physical activity

Edited by G Eysenbach; submitted 16.07.20; peer-reviewed by C Janney, T Yano; comments to author 09.10.20; revised version received 22.10.20; accepted 17.01.21; published 01.03.21.

Please cite as:
Hollander JB, Folta SC, Graves EM, Allen JD, Situ M
A Fitness App for Monitoring Walking Behavior and Perception (Runkeeper): Mixed Methods Pilot Study
JMIR Form Res 2021;5(3):e22571
URL: https://formative.jmir.org/2021/3/e22571
doi:10.2196/22571
PMID:33646132

Original Paper

Web-Based Dietary Intake Estimation to Assess the Reproducibility and Relative Validity of the EatWellQ8 Food Frequency Questionnaire: Validation Study

Balqees Alawadhi[1], BSc, MSc, PhD; Rosalind Fallaize[1,2], BSc, PhD; Rodrigo Zenun Franco[3], BSc, MSc, PhD; Faustina Hwang[3], BEng, MEng, PhD; Julie Lovegrove[1], BSc, PhD

[1]Hugh Sinclair Unit of Human Nutrition and Institute for Cardiovascular and Metabolic Research, Department of Food and Nutritional Sciences, University of Reading, Reading, United Kingdom

[2]School of Life and Medical Science, University of Hertfordshire, Hertfordshire, United Kingdom

[3]Biomedical Engineering Section, School of Biological Sciences, University of Reading, Reading, United Kingdom

Corresponding Author:
Julie Lovegrove, BSc, PhD
Hugh Sinclair Unit of Human Nutrition and Institute for Cardiovascular and Metabolic Research, Department of Food and Nutritional Sciences
University of Reading
Whiteknights
Reading, RG6 6AP
United Kingdom
Phone: 44 0118 378 6418
Email: j.a.lovegrove@reading.ac.uk

Abstract

Background: The web-based EatWellQ8 food frequency questionnaire (FFQ) was developed as a dietary assessment tool for healthy adults in Kuwait. Validation against reliable instruments and assessment of its reproducibility are required to ensure the accuracy of the EatWellQ8 FFQ in computing nutrient intake.

Objective: This study aims to assess the reproducibility and relative validity of the EatWellQ8 146-item FFQ, which included images of food portion sizes based on *The Composition of Foods* by McCance and Widdowson and food composition tables from Kuwait and the Kingdom of Bahrain, against a paper-based FFQ (PFFQ) and a 4-day weighed food record (WFR).

Methods: Reproducibility of the EatWellQ8 FFQ was assessed using a test-retest methodology. Participants were required to complete the FFQ at 2 time points, 4 weeks apart. To assess the relative validity of the EatWellQ8 FFQ, a subset of the participants were asked to complete a PFFQ or a 4-day WFR 1 week after the administration of the EatWellQ8 FFQ. The level of agreement between nutrient and food group intakes was estimated by repeated EatWellQ8 FFQ administration. The EatWellQ8 FFQ, PFFQ, and 4-day WFR were also evaluated using the Bland-Altman methodology and classified into quartiles of daily intake. Crude unadjusted correlation coefficients were also calculated for nutrients and food groups.

Results: A total of 99 Kuwaiti participants (64/99, 65% female and 35/99, 35% male) completed the study—53 participated in the reproducibility study and the 4-day WFR validity study (mean age 37.1 years, SD 9.9) and 46 participated in the PFFQ validity study (mean age 36.2 years, SD 8.3). Crude unadjusted correlations for repeated EatWellQ8 FFQs ranged from 0.37 to 0.93 (mean r=0.67, SD 0.14; 95% CI 0.11-0.95) for nutrients and food groups (P=.01). Mean cross-classification into *exact agreement plus adjacent* was 88% for nutrient intakes and 86% for food groups, and Bland-Altman plots showed good agreement for energy-adjusted macronutrient intakes. The association between the EatWellQ8 FFQ and PFFQ varied, with crude unadjusted correlations ranging from 0.42 to 0.73 (mean r=0.46, SD 0.12; 95% CI −0.02 to 0.84; P=.046). Mean cross-classification into *exact agreement plus adjacent* was 84% for nutrient intake and 74% for food groups. Bland-Altman plots showed moderate agreement for both energy and energy-controlled nutrient intakes. Crude unadjusted correlations for the EatWellQ8 FFQ and the 4-day WFR ranged from 0.40 to 0.88 (mean r=0.58, SD 0.13; 95% CI 0.01-0.58; P=.01). Mean cross-classification into *exact agreement plus adjacent* was 85% for nutrient intake and 83% for food groups. Bland-Altman plots showed moderate agreement for energy-adjusted macronutrient intakes.

Conclusions: The results indicate that the web-based EatWellQ8 FFQ is reproducible for assessing nutrient and food group intake and has moderate agreement compared with a PFFQ and a 4-day WFR for measuring energy and nutrient intakes.

(JMIR Form Res 2021;5(3):e13591) doi:10.2196/13591

KEYWORDS

web-based; Kuwait; weighed food record; app; food frequency questionnaire; validation; dietary assessment

Introduction

Background

According to the World Health Organization, noncommunicable diseases (NCDs) remain to be the main cause of global premature mortality [1]. Diets rich in energy and saturated fat and low in fruits and vegetables have been associated with the development of NCDs [2,3]. Inaccurate dietary assessment methods may be a serious obstacle in understanding the impact of dietary factors on disease [4]. Several dietary assessment methods are available, including the food frequency questionnaire (FFQ), diet history, weighed food record (WFR), and 24-hour dietary recall [5]. FFQs require respondents to state the frequency of intake of a predefined list of foods over a specified period and are one of the most commonly used tools to assess the relationship between diet, health, and disease [6].

With the widespread availability of the internet, there has been a growing interest in using the web to assess dietary intake and deliver health-related messages. Traditional dietary assessment methods have been customized for internet use in research as they allow for the direct storage of data and automatic generation of nutrition outputs [7,8]. In addition, web-based dietary assessment methods may be more cost-effective and can include photographs of food portion sizes, increasing the ease of use for respondents, and can be designed to be user-friendly and tailored toward a specific target group [9,10].

This study is part of the EatWellQ8 study, which aims to investigate whether web-based personalized nutrition (PN; based on dietary intake and anthropometrics) is as effective as face-to-face communication of PN in Kuwait. Kuwait currently has the highest adult obesity levels in the Gulf region [11]. The latest findings indicate that around 78% of adult men and 82% of women in Kuwait are either overweight or obese [12].

The novel EatWellQ8 FFQ was developed to assess the dietary intake in Kuwait and included 146 food items and photographs of food portion sizes. The validated Food4Me FFQ, the European Prospective Investigation of Cancer (EPIC) Norfolk FFQ (version CAMB/PQ/6/1205), and a paper-based FFQ (PFFQ) for Kuwait were used as guides in the development of the EatWellQ8 FFQ food items and categories of food [13-16]. Good agreement between the web-based Food4Me FFQ and the EPIC-Norfolk FFQ for the estimation of energy-adjusted nutrient intake was shown earlier [15,16].

Objectives

The aim of this study is to develop and test the reproducibility of the EatWellQ8 FFQ for the assessment of food and nutrient intake in a Kuwaiti population for use in the EatWellQ8 study and to compare estimates of dietary intake using this tool with data obtained from a 4-day WFR and a validated paper, Kuwaiti FFQ (PFFQ) [14].

Methods

Study Sample

A sample size between 50 and 100 is recommended to accurately evaluate the Bland-Altman limits of agreement (LOA) between 2 methods [5]. Participants aged 18 to 65 years were recruited from Kuwait through email, poster advertisement, word of mouth, booths at colleges and health institutions, and social media (WhatsApp, Facebook, YouTube, and Instagram). Participants were then provided with an information sheet clarifying the study, a consent form, or an assent form (for participants aged 18-21 years) and asked to complete a web-based screening questionnaire. Participants were emailed a feedback response dependent on whether they met the inclusion criteria. A minimal set of exclusion criteria were applied (subjects aged below 18 years; pregnant or lactating; no or limited access to the internet; following a prescribed diet, including a weight-reducing diet in the previous 3 months; diabetes; celiac disease; Crohn disease; and previous chronic medical conditions requiring continuing therapeutic intervention apart from hypertension medication and statins). The study was approved by the Research Ethics Committee at the University of Reading (School of Chemistry, Food, and Pharmacy Research Ethics Committee, Ref. No. 13/17) and conformed with the Declaration of Helsinki. The study also received ethical approval from the Research Ethics Committee at the Dasman Diabetes Institute (DDI), Kuwait (RA-2015-018).

Study Design

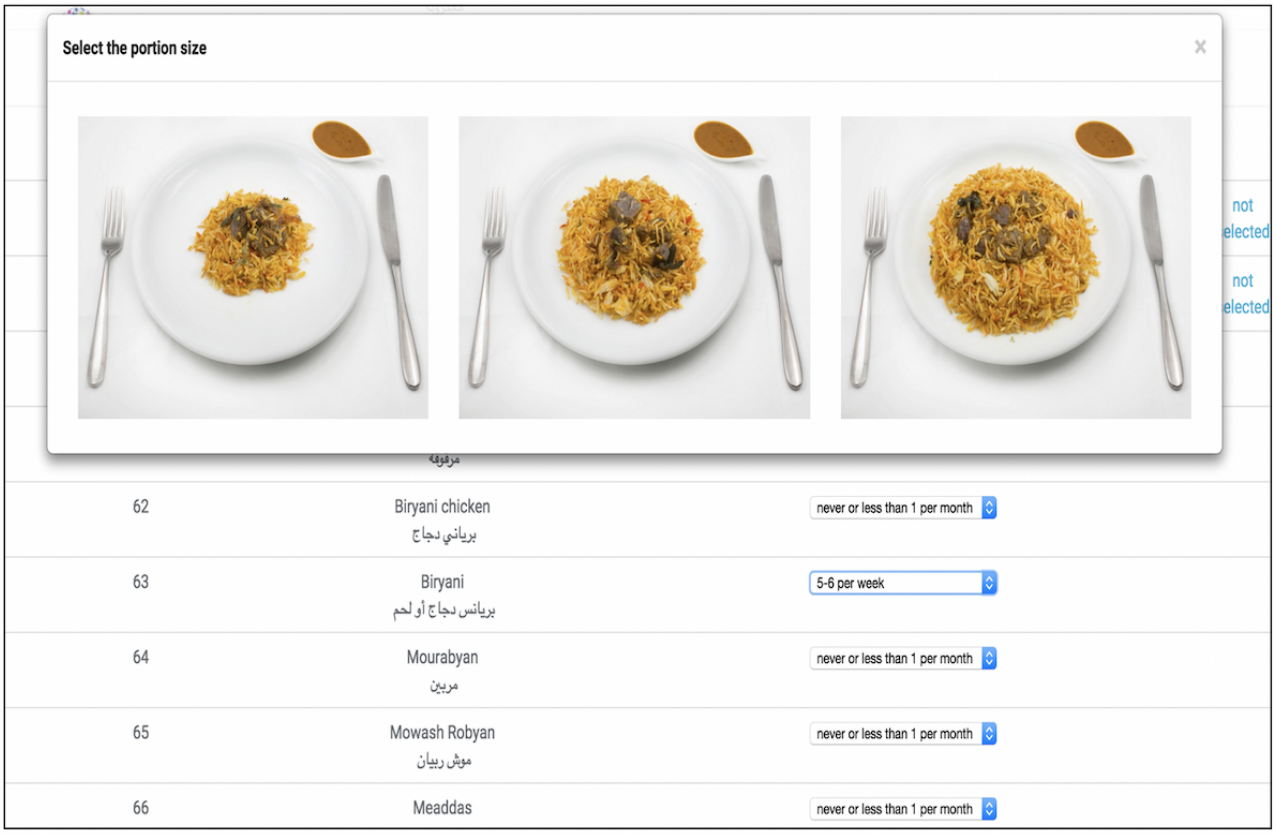
To assess the reproducibility of the EatWellQ8 FFQ, 100 participants were asked to complete the web-based FFQ twice, 4 weeks apart for intake over the past month, between the months of September and December. To assess the relative validity of the EatWellQ8 FFQ, participants were also asked to complete a 4-day WFR, a week after completing the web-based FFQ. An additional 50 participants were asked to take the EatWellQ8 FFQ at baseline and to complete a validated PFFQ for Kuwait a week after completing the web-based FFQ. The Kuwaiti PFFQ and the 4-day WFR were delivered to the participants in person or sent via email, depending on the participant's preference. Participants were asked to complete the forms and hand them in person or to scan and email them to the researcher. Participants were asked to complete a usability survey after completing the first EatWellQ8 FFQ [17]. Reminders were sent biweekly to participants in the form of email and text messages to encourage completion of the tools. All participants were requested to maintain their usual diet during the study.

The EatWellQ8 FFQ

The web-based EatWellQ8 semiquantitative FFQ was designed to measure the short-term nutritional and dietary intakes of adults in Kuwait. The design and development of the novel EatWellQ8 FFQ was led by researchers from the Hugh Sinclair

Unit of Human Nutrition and the Biomedical Engineering section at the University of Reading. The validated Food4Me FFQ, the well-validated EPIC-Norfolk FFQ (version CAMB/PQ/6/1205), and a valid semiquantitative FFQ for Kuwait were used as a guide in the development of the novel FFQ to identify food items and categorize food into different food groups [13,14]. To ensure that the EatWellQ8 FFQ was suitable for use among people in Kuwait, participants were able to choose between 2 languages—Arabic and English. The novel FFQ comprised 146 food items (Multimedia Appendix 1) that represented food items and composite dishes commonly consumed in Kuwait. Several new foods that are commonly consumed in Kuwait were added to the existing food categories, for example, pomegranate, guava, and mango were added to the fruit list and *Lebanese bread and Iranian bread* were added to the bread and savory biscuit list. A new food section titled *Kuwaiti composite dishes* was added, which included 23 food items such as *Machboos Laham, Biryani, and Harees* to ensure that commonly consumed foods were included in the FFQ. In addition, traditional Kuwaiti desserts such as *Konafa, maamoul, and luqaimat* were added to the sweets and snacks section to ensure the inclusion of most of the commonly consumed food items. Alcoholic drinks and pork were removed from the FFQ as they were not commonly consumed items and also to respect the religious culture in Kuwait. Food items on the web-based FFQ appeared as a list where all the food items are displayed on a single page, as compared with displaying foods in food groups that are presented over several consecutive pages (Figure 1).

Figure 1. Screenshot of the web-based EatWellQ8 food frequency questionnaire illustrating the 3 portion size photographs for the assessment of the portion size.



Photographs

Portion size photographs for 64 of the foods were derived from the Food4Me food portion size photograph list [15]. The remaining 83 food items were purchased from local supermarkets and local restaurants and bakeries in Kuwait. All foods were prepared and photographed at the DDI, Kuwait, over a period of 7 days or sessions by a professional photographer from DDI. Photographs were taken in the demo kitchen at DDI using the same lighting and a standard dining set of plates and cutlery that were positioned consistently for each session. All foods were weighed using calibrated portable food scales (Salter), and the calculated Food4Me portion sizes were used as a guide for all food items.

The Four-Day WFR

Participants were asked to record all food items and beverages consumed over a four-day period that included 3 weekdays (Sunday to Thursday) and 1 weekend day (Friday to Saturday). Before beginning the WFR, participants were asked to attend a preliminary training session at DDI given by a dietitian on how to describe food products and use the provided food scales (Salter Disc Electronic Kitchen Scales SKU1036 WHSSDR). Participants were given the flexibility of estimating portion sizes when they were unable to weigh the food items (eg, when dining out).

The Kuwaiti Validated FFQ

A total of 50 participants were asked to complete a PFFQ after completing the initial web-based FFQ. The Kuwait Validated FFQ is a self-administered semiquantitative FFQ that was developed in 2009. The FFQ was developed to target the frequency of consumption and portion sizes of food and beverages regularly consumed by the Kuwaiti population [14]. Standardized portions of the food items and beverages were used to estimate portion sizes, and 9 frequencies ranging from *never or once a month* to *more than 6 times/day* were used for frequency estimation [14]. The FFQ included questions on the average intake of 201 food items over the past 4 months. However, the time frame was reduced to 1 month for the purpose of the validation study. The food items were divided into the following 14 groups: *cereals*, *composite dishes*, *marag (stew)*, *soups*, *meat dishes*, *snacks*, *desserts*, *dairy products*, *beverages*, *fruits*, *vegetables*, *stuffed vegetables*, *salads*, and *miscellaneous*. The food intake (g/day) was calculated by multiplying the portion of each food listed in the FFQ by the frequency of consumption and by the nutrient composition of the food using the United States Department of Agriculture nutrient database [14].

Dietary Intake Analysis

Estimated dietary intake data from the EatWellQ8 FFQ were generated automatically by the web-based EatWellQ8 app, which was described previously by Franco et al [17]. Nutritional composition and portion sizes of the 146 food items were calculated using the Food4Me food list [18], fifth and sixth editions of *The Composition of Foods* by McCance and Widdowson [19,20], the Kingdom of Bahrain Food Composition Tables [21], and the National Kuwait Food Composition List [22]. From these lists, the most commonly consumed food items were selected and used to calculate the composition of the lists of foods in the EatWellQ8 FFQ. The nutritional compositions of all the Kuwaiti composite dishes were determined using the Kingdom of Bahrain food composition list and a Kuwaiti food composition list [21,22]. Portion sizes were primarily derived using the Food4Me food list [15,16]. To calculate the portion sizes, the food codes for each of the frequently consumed foods were identified from the Food4Me database and used to formulate the code for the food items in the FFQ. PASW Statistics version 24 (SPSS Inc) was used to calculate the 25th, 50th, and 75th percentile of daily food intake, which corresponds, respectively, to small, medium, and large portion of these foods when consumed by the general population [15]. Estimated nutrient intakes for the PFFQ were analyzed using a Microsoft Excel file that was based on the web-based EatWellQ8 programmed system. The four-day WFR intakes were analyzed using Nutritics software (version 1.8, database MW6, Nutritics Ltd, Co).

Over-Underreporting

Participants' results were excluded from the analysis if their daily energy intake was found to be less than 500 kcal or greater than 4500 kcal in any of the methods [23].

Statistical Analysis

Statistical analyses were performed using SPSS (version 24.0, PASW). Normality was assessed using the Shapiro-Wilk test, and log transformation was used for nonparametric data when necessary. A paired two- tailed *t* test was performed to assess differences in participants' energy intake (kcal) between the methods used. SDs and mean nutrient intakes were calculated for baseline, repeated EatWellQ8 FFQ, PFFQ, and four-day WFR. Comparisons between nutrient intakes were performed using a general linear model analysis, which was controlled further for energy and gender, where there was a significant interaction between nutrient intake and gender. To check for normality, data were analyzed using a Shapiro-Wilk test, and either the Pearson or the Spearman correlation coefficient (SCC) was used for normally or nonnormally distributed data, respectively. Correlations were considered statistically significant if the *P* value was <.05.

To test for agreement between the different dietary intake methods and repeated EatWellQ8 FFQ, cross-classification of nutrient intakes to assess the percentage of participants classified into the following quartiles: exact agreement (percentage of cases cross-classified into the same quartile), exact agreement plus adjacent (percentage of cases cross-classified into the same or adjacent quartile), disagreement (percentage of cases cross-classified 2 quartiles apart), and extreme disagreement (percentage of cases cross-classified into extreme quartiles). The Bland-Altman [24] method was used to further analyze the LOA for energy intakes and macronutrients between the repeated EatWellQ8 FFQ and between the 3 methods (EatWellQ8 FFQ, WFR, and PFFQ). On the basis of the Bland-Altman method, dietary intake methods were found to be repeatable or comparable if greater than 95% of the data plots fell within the 2 SD of the mean (LOA) and by calculating the bias calculated by the mean difference and SD of the differences.

Results

Overview

Of the 235 participants screened for the study, 218 were found to be eligible. Participants were excluded (n=17) because of incomplete FFQs or not fulfilling the screening requirements because of medication use, food allergies, or an existing illness. The mean BMI, weight, and height of participants included in the study were 25.6 kg/m^2 (SD 4.4), 70.3 kg (SD 14.0), and 165.5 cm (SD 8.6), respectively, and the mean BMI, weight, and height of dropouts were 25.7 kg/m^2 (SD 4.3), 70.7 kg (SD 13.8), and 166 cm (SD 8.2), respectively. A high dropout rate of 48.6% (106/218) was found, which was mainly because of participants' unwillingness to complete all aspects of the study. The mean completion time of the EWQ8 FFQ was 14.3 minutes (95% CI 12.9-15.3) [17]. A total of 110 participants completed the EatWellQ8 FFQ1, of which 60 completed EatWellQ8 FFQ2 and a four-day WFR and 50 were asked to complete a PFFQ. In total, 18 participants were excluded from the analysis because of reported energy intakes of <500 kcal or >4500 kcal [23]. Removal of under-overreporters did not have an impact on the outcomes of reproducibility and validity (data not shown). Of

these, 53 participants completed the second EatWellQ8 FFQ, 46 completed EatWellQ8 FFQ2 and the four-day WFR, and 46 participants completed the PFFQ. An illustration of the flow of the participants is shown in Figure 2. Demographic characteristics based on self-report are shown in Table 1. No significant differences were found between age and BMI for females and males. A higher percentage of females completed both studies (60/92, 65% in the validation study and 35/53, 66% in the reproducibility study).

Figure 2. Participant flow during the study.

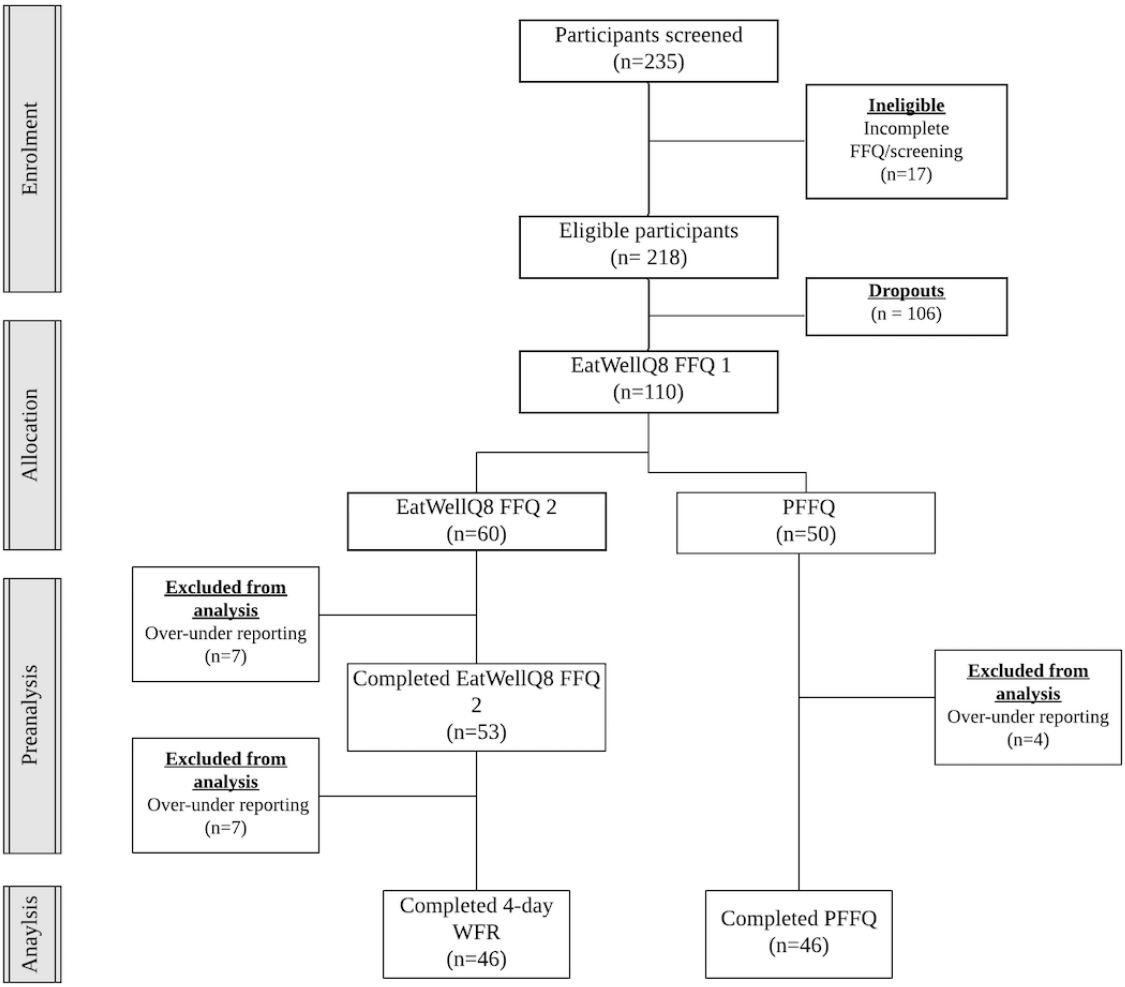


Table 1. Demographic characteristics of participants who completed the reproducibility and validation studies.

Study	Population, n (%)	Demographic characteristics, mean (SD)			
		Age (years)	Weight (kg)	Height (cm)	BMI (kg/m^2)
Reproducibility					
All	53 (100)	37 (9.9)	70.3 (13.9)	165.5 (8.6)	25.6 (4.4)
Female	35 (66)	36 (9.8)	64.9 (11.1)	161.6 (6.2)	24.9 (4.6)
Male	18 (34)	39 (9.7)	80.8 (12.8)	173.3 (7.3)	26.8 (3.5)
Validation					
All	92 (100)	36 (8.3)	71.9 (15.5)	167.4 (8.1)	25.2 (4.4)
Female	60 (65)	37 (9.2)	65.6 (11.9)	163.8 (5.6)	24.3 (4.4)
Male	32 (35)	34 (6.1)	82.6 (15.4)	173.6 (7.9)	27.0 (3.9)

Reproducibility of the EatWellQ8 FFQ

Comparison of Nutrient Intake Between Repeated EatWellQ8 FFQs

No significant differences were found between macronutrient and micronutrient intake evaluated in FFQ1 and FFQ2 (Table 2). Correlations were found to be significant for all nutrients (P=.01) and ranged from 0.37 (polyunsaturated fatty acids [FAs], percentage total energy [%TE]) to 0.82 (iron), with a mean value of r=0.67 (SD 0.14; 95% CI 0.11-0.89; Multimedia Appendix 2 Table S1). Adjustments for energy and gender did not modify these correlations. Results of the cross-classifications for percentage of participants classified into quartiles of exact agreement ranged from 40% (polyunsaturated FAs, %TE) to 62% (total folate). Classifications of exact agreement plus adjacent regions ranged from 77% (monounsaturated FAs, %TE) to 100% (energy, kcal). Disagreement was relatively low, the mean percentage of participants classified into quartiles of disagreement was 8%, and the mean of participants classified as having extreme disagreement was 1.40%.

The Bland-Altman plots for estimates of energy (kcal), protein (%TE), total fat (%TE), and carbohydrate (%TE) intakes are shown in Figure 3. Good agreement was found in the Bland-Altman plots, as the majority of the cases fell within the 95% LOA. The EatWellQ8 FFQ presented good reproducibility for the evaluation of daily fat intake, with less than 4% of cases outside the LOA. For energy and carbohydrate, less than 6% fell outside the LOA and 7% for protein. On the basis of the LOA values, greater agreement was found for protein (%TE) compared with energy, total fat, and total carbohydrate (%TE). No significant bias was identified for any nutrient. Variation between estimates of energy and energy-adjusted macronutrient intakes increased with higher mean intakes (Figure 3).

Table 2. Mean daily energy and nutrient intakes estimated by repeated measures of the web-based EatWellQ8 food frequency questionnaire (N=53).

Nutrient	EatWellQ8 FFQ1[a], mean (SD)	EatWellQ8 FFQ2, mean (SD)	*P* value[b]	*P* value[c]
Energy (kcal)	2724 (1355)	2524 (1232)	.09[d]	N/A[e]
Total fat (g)	104.7 (56.2)	96.1 (50.2)	.79	.79
Total fat (%TE[f])	34.2 (7.7)	34.2 (7.8)	.96	.96
SFA[g] (g)	43.5 (27.0)	38.5 (22.1)	.49	.49
SFA (%TE)	14.0 (4.4)	13.6 (4.8)	.73	.73
MUFA[h] (g)	45.4 (25.0)	40.0 (21.4)	.34	.34
MUFA (%TE)	14.9 (4.2)	14.3 (3.9)	.41	.41
PUFA[i] (g)	18.0 (9.1)	17.5 (8.9)	.46	.46
PUFA (%TE)	6.1 (1.7)	6.4 (1.7)	.58	.58
Omega 3 (g)	0.17 (0.2)	0.20 (0.4)	.47	.47
Protein (g)	117.8 (57.3)	111.4 (52.2)	.84	.84
Protein (%TE)	17.7 (4.1)	18.5 (5.1)	.53	.54
Carbohydrate (g)	348 (197)	323 (182)	.87	.88
Carbohydrate (%TE)	51.1 (9.8)	50.3 (10.7)	.76	.77
Total sugars (g)	149 (91)	134 (79)	.69	.69
Total sugars (%TE)	22.1 (9.2)	21.4 (7.3)	.67	.67
Calcium (mg)	1288 (682)	1192 (633)	.86	.86
Total folate (µg)	405 (204)	358 (167)	.29	.29
Iron (mg)	16.8 (9.5)	15.0 (7.8)	.41	.41
Total carotene (µg)	6581(6161)	5548 (4079)	.41	.41
Riboflavin (mg)	2.3 (1.2)	2.2 (1.1)	.90	.89
Thiamin (mg)	2.1 (1.1)	1.9 (0.9)	.78	.77
Vitamin B6 (mg)	2.8 (1.2)	2.7 (1.2)	.85	.85
Vitamin B12 (µg)	5.2 (3.3)	5.4 (3.4)	.33	.33
Vitamin C (mg)	200 (116)	183 (135)	.77	.76
Vitamin A RE[j] (µg)	1319 (1048)	1153 (718)	.48	.48
Retinol (µg)	241 (190)	258 (191)	.15	.15
Vitamin D (µg)	3.1 (2.3)	3.4 (3.2)	.36	.36
Vitamin E (mg)	15.4 (7.5)	14.3 (7.6)	.89	.89
Sodium (mg)	3159 (1570)	2948 (1485)	.99	.99

[a]FFQ: food frequency questionnaire.

[b]Controlled for energy.

[c]Controlled for energy and gender.

[d]Value derived from paired sample *t* test.

[e]N/A: not applicable.

[f]%TE: percentage total energy.

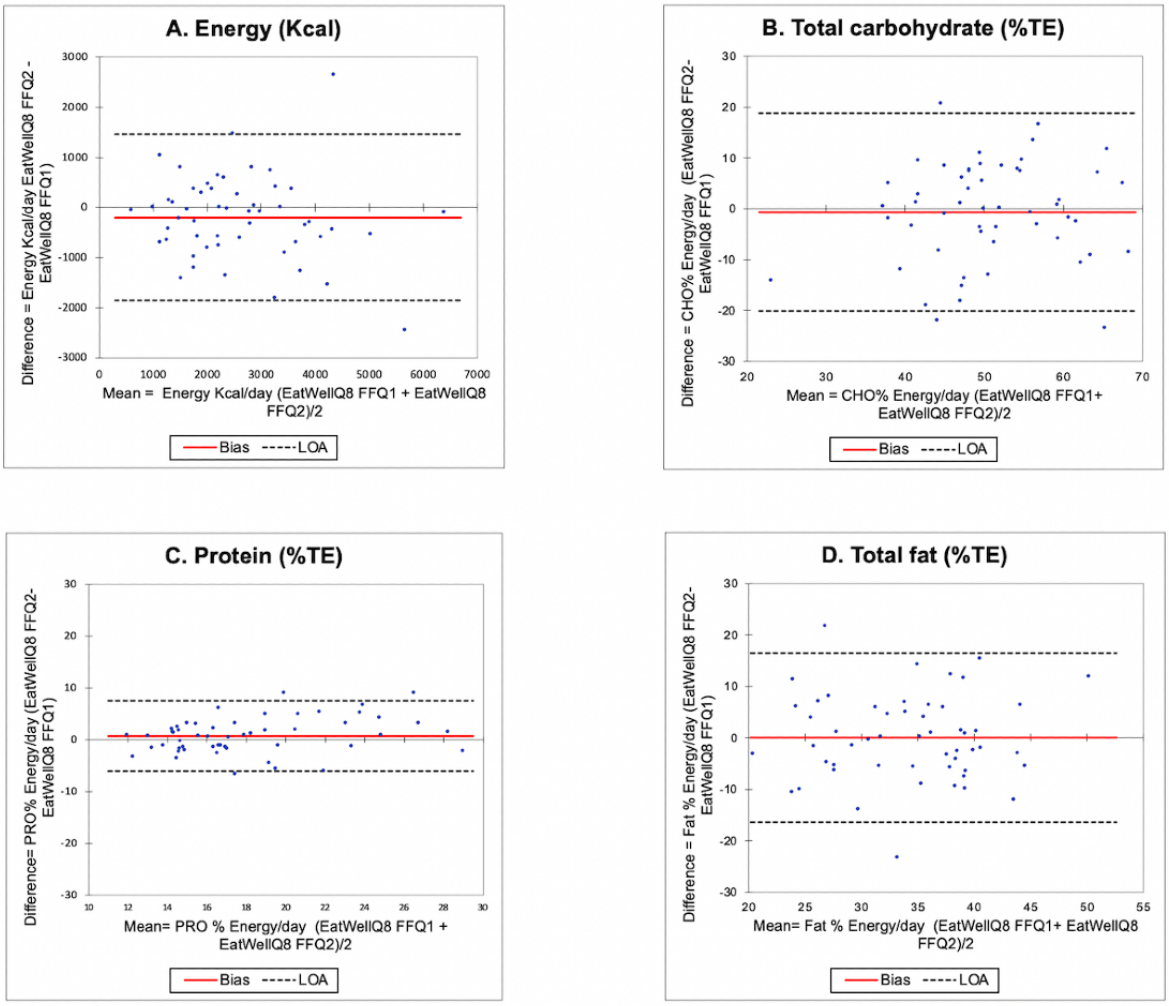
[g]SFA: saturated fatty acid.

[h]MUFA: monounsaturated fatty acid.

[i]PUFA: polyunsaturated fatty acid.

[j]RE: retinol equivalent.

Figure 3. Reproducibility study of Bland-Altman plots for (a) energy, (b) total carbohydrate, (c) protein, and (d) total fat with the bias (mean difference) and limits of agreement. The solid line represents the bias (mean difference), and the dotted lines represent the limits of agreement. %TE: percentage total energy; CHO: carbohydrates; PRO: protein.



Comparison of Food Group Intakes Between Repeated EatWellQ8 FFQ

Food items were categorized into 32 food groups to assess the differences between repeated administrations of the web-based EatWellQ8 FFQ. SCCs ranged from 0.40 (savories) to 0.93 (meat products) with a mean value of r=0.67 (SD 0.14; 95% CI 0.11-0.95; Multimedia Appendix 2 Table S2). Significant correlations were found for all food groups (P=.01). The cross-classification of participants classified into quartiles of exact agreement ranged from 45% (salad vegetables) to 76% (meat products). Moderately high classifications of exact agreement and adjacent were found, which ranged from 66% (confectionary and savory snacks) to 98% (meat products).

Validation of the EatWellQ8 FFQ

Comparison of Nutrient Intakes Between the EatWellQ8 FFQ and the Kuwaiti PFFQ

No significant differences were found between 70% of the macronutrients and micronutrients evaluated by the EatWellQ8 FFQ1 and PFFQ (Table 3). Estimated energy intakes were found to be significantly higher (difference 398 kcal/day) and 17% higher (P<.001) in the EatWellQ8 FFQ1 than in the PFFQ.

After controlling for energy, similar estimated intakes of macronutrients and micronutrients were observed for EatWellQ8 FFQ1 and the PFFQ except for saturated fatty acids (SFAs) and monounsaturated fatty acids (MUFAs; g, %TE), which were significantly higher for EatWellQ8 FFQ than for PFFQ (P<.001). Furthermore, the estimated intakes of total folate (P=.01), retinol (P<.001), and vitamin B12 (P<.001) were higher in the PFFQ than in the EatWellQ8 FFQ.

With the exception of omega 3 FAs and retinol, correlations were found to be significant for all nutrients (P=.01) and ranged from 0.42 (vitamin D) to 0.73 (energy), with a mean value of r=0.54 (SD 0.12; Multimedia Appendix 2 Table S3). However, large variations were found in 95% CI ranging from −0.02 to 0.84, and weak 95% CIs were found for retinol (−0.02) and omega 3 FAs (−0.11). The results of the cross-classifications for percentage of participants classified into quartiles of exact agreement ranged from 35% (total fat) to 57% sodium (Na), exact agreement plus adjacent, ranging from 76% (total fat, %TE) to 93% (energy) with low levels in disagreement (13.41%) and extreme disagreement (2.40%).

Table 3. Mean daily energy and nutrient intakes estimated by the web-based EatWellQ8 food frequency questionnaire and a paper-based food frequency questionnaire and general linear model results (N=46).

Nutrient	EatWellQ8 FFQ[a], mean (SD)	PFFQ[b], mean (SD)	*P* value[c]	*P* value[d]
Energy (kcal)	2297 (779)	1899 (505)	<.001[e]	N/A[f]
Total fat (g)	92.1 (40.8)	69.1 (23.4)	.12	.12
Total fat (%TE[g])	35.5 (7.8)	32.4 (5.1)	.13	.13
SFA[h] (g)	38.4 (17.9)	26.9 (9.1)	.01	.01
SFA (%TE)	14.9 (4.8)	12.6 (2.7)	.01	.01
MUFA[i] (g)	39.5 (19.5)	26.4 (9.7)	<.001	.001
MUFA (%TE)	15.1 (4.3)	12.3 (2.3)	<.001	.001
PUFA[j] (g)	16.1 (7.8)	13.1 (6.5)	.64	.58
PUFA (%TE)	6.2 (1.7)	6.1 (1.8)	.93	.96
Omega 3 (g)	0.1 (0.2)	0.3 (0.3)	<.001	<.001
Protein (g)	104 (41)	93 (41)	.35	.33
Protein (%TE)	18.2 (5.0)	19.4(6.1)	.39	.37
Carbohydrate (g)	280 (102)	241 (70)	.53	.53
Carbohydrate (%TE)	49.1 (10.5)	51.2 (7.9)	.53	.54
Total sugars (g)	125 (52)	105 (32)	.56	.53
Total sugars (%TE)	22.3 (8.8)	22.6 (6.2)	.53	.53
Calcium (mg)	1126 (542)	933 (358)	.86	.84
Total folate (µg)	323 (135)	328 (116)	.01	.01
Iron (mg)	13.5 (5.7)	11.5 (4.0)	.48	.49
Total carotene (µg)	5042 (3430)	4781 (4325)	.82	.82
Riboflavin (mg)	2.0 (0.9)	1.8 (0.7)	.30	.31
Thiamin (mg)	1.7 (0.7)	1.5 (0.4)	.18	.18
Vitamin B6 (mg)	2.4 (0.8)	2.1 (0.8)	.77	.77
Vitamin B12 (µg)	4.8 (2.8)	5.5 (3.2)	<.001	<.001
Vitamin C (mg)	163 (141)	156 (98)	.67	.68
Vitamin A RE[k] (µg)	1054 (590)	1110 (796)	.22	.22
Retinol (µg)	237 (138)	387 (452)	<.001	<.001
Vitamin D (µg)	2.6 (2.0)	2.3 (1.7)	.48	.45
Vitamin E (mg)	12.7 (7.0)	9.5 (3.9)	.44	.45
Sodium (mg)	2701 (1058)	2102 (771)	.21	.21

[a]FFQ: food frequency questionnaire.

[b]PFFQ: paper-based food frequency questionnaire.

[c]Controlled for energy.

[d]Controlled for energy and gender.

[e]Value derived from paired sample *t* test.

[f]N/A: not applicable.

[g]%TE: percentage total energy.

[h]SFA: saturated fatty acid.

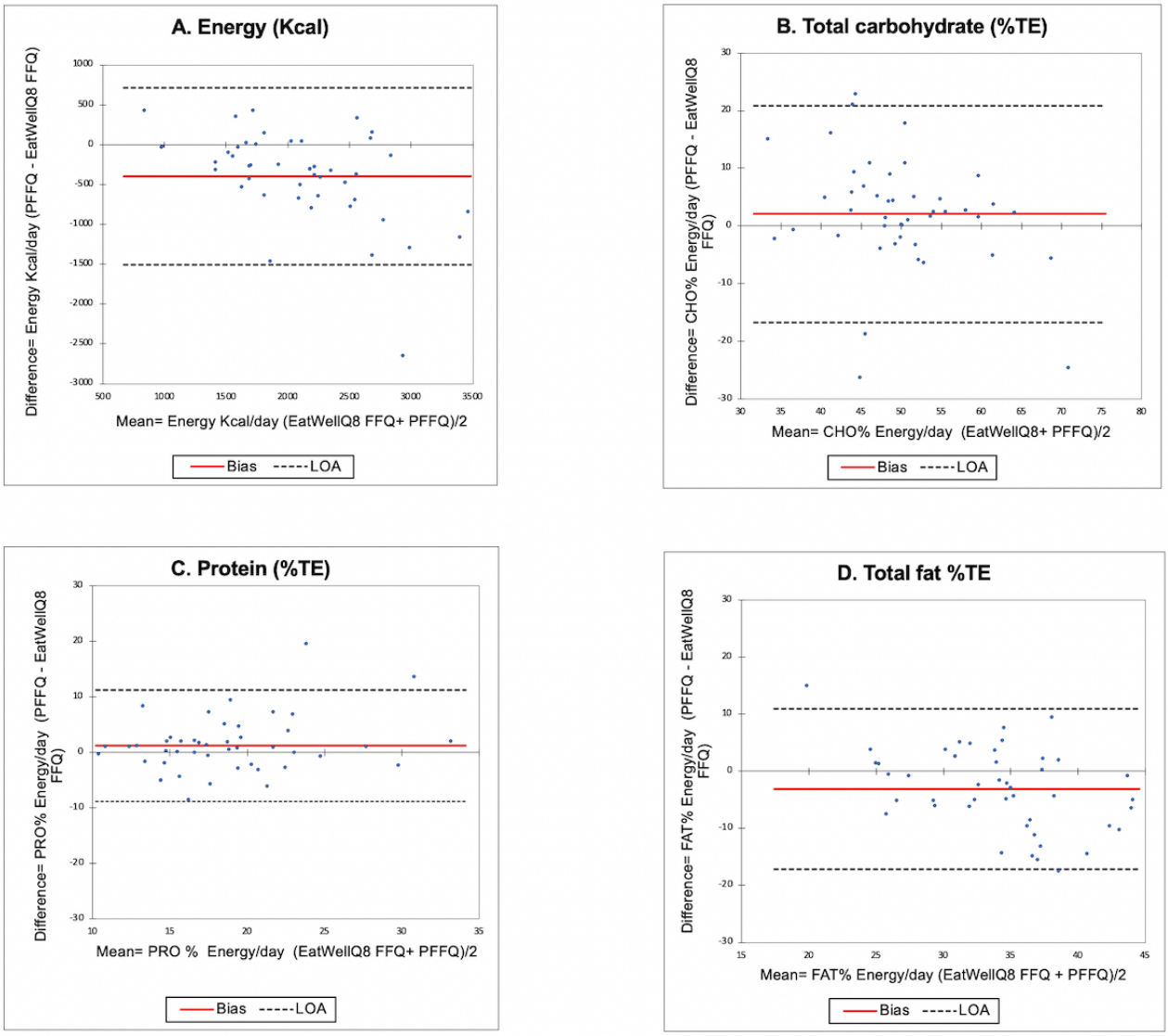
[i]MUFA: monounsaturated fatty acid.

[j]PUFA: polyunsaturated fatty acid.

[k]RE: retinol equivalent.

Overall, moderate agreement was found in the Bland-Altman plots between the EatWellQ8 and the paper form of PFFQ, with 87% of all cases falling within the 95% LOA (Figure 4). The EatWellQ8 FFQ presented good validation for the evaluation of energy and daily intake of fat, with approximately 4% of cases falling outside the LOA. For daily intake of carbohydrates, less than 6% fell outside the LOA and, for protein, 8% fell out of the LOA. Protein (%TE) had the narrowest LOA, which signifies better agreement compared with energy, total fat (%TE), and carbohydrate (%TE). The bias (mean difference) between energy intakes was significantly higher (398 kcal/day), with greater intakes reported in the EatWellQ8 FFQ. A high mean bias was found for total fat (−3.05%TE) compared with total carbohydrate (2.67%TE) and protein (1.20 %TE). No other significant differences were observed.

Figure 4. Validation study of Bland-Altman plots comparing the EatWellQ8 food frequency questionnaire (to a paper-based food frequency questionnaire for (a) energy, (b) total carbohydrate, (c) protein, and (d) total fat with the bias (mean difference) and limits of agreement. The solid line represents the bias (mean difference), and the dotted lines represent the limits of agreement. %TE: percentage total energy; CHO: carbohydrates; FFQ: food frequency questionnaire; LOA: limit of agreement; PRO: protein; PFFQ: paper-based food frequency questionnaire.



Comparison of Food Group Intakes Between the EatWellQ8 FFQ and the Kuwaiti PFFQ

SCCs ranged from 0.51 (bananas) to 0.22 (fish and fish product or dishes; 95% CI −0.07 to 0.71; Multimedia Appendix 2 Table S4). With the exception of fish and fish products or dishes, significant correlations were found for all food groups (P=.047). The cross-classification percentages of participants classified into quartiles of exact agreement ranged from 60% (soups, sauces, and miscellaneous foods) to 24% (white bread). Classifications of exact agreement plus adjacent ranged from 65% (ice cream, creams, and desserts) to 82% (teas and coffees). The mean percentage of participants classified into quartiles of disagreement was 15% and, for extreme disagreement, it was 9%.

Comparison of Nutrient Intakes Between the EatWellQ8 FFQ and a Four-Day WFR

Estimated macronutrient intakes were found to be similar between the EatWellQ8 FFQ and the 4-day WFR after controlling for energy (Table 4). However, estimated intakes of SFA (P=.001), total carbohydrates (P=.03), and total sugars

(g, %TE; *P*=.03) were significantly higher in the EatWellQ8 FFQ than in the four-day WFR. Significantly higher estimated intakes of omega 3 FAs, folate, total carotene, thiamin, vitamin B6, vitamin C, vitamin A retinol equivalent, and Sodium (*P*=.02) were found for the EatWellQ8 FFQ compared with the four-day WFR. Similar results were found after controlling for both energy and gender.

Significant correlation for all nutrients was found at the *P*=.01 level, correlation coefficients ranged from 0.40 (iron) to 0.88 (energy), with a mean value of *r*=0.61 (SD 0.13; 95% CI 0.13-0.93; Multimedia Appendix 2 Table S5). The percentage of volunteers classified into quartiles of exact agreement ranged from 28% (polyunsaturated fatty acid, g) to 67% (energy, kcal). Values were higher for classifications of exact agreement plus adjacent and ranged from 71% (MUFA, %TE) to 95% (protein, g). The mean percentage of volunteers classified into quartiles of disagreement was 11%, and less than 2% of volunteers were classified as having extreme disagreement.

In total, good agreement between the methods was found, with less than 5% of cases falling outside of the LOA for all of the plots (Figure 5). On the basis of the LOA values, the highest agreement was found for protein (%TE) compared with energy total carbohydrate (%TE) and total fat (%TE). Bias (mean difference) between energy intakes was small (81 kcal/day), with greater values estimated in the EatWellQ8 FFQ. Higher bias for energy-adjusted total carbohydrate (4.39%TE) and total fat (1.20%TE) intake was measured in the EatWellQ8 FFQ. However, a higher bias for energy-adjusted protein (1.65%TE) intakes was measured in the four-day WFR.

Table 4. Mean daily energy and nutrient intakes estimated by the web-based EatWellQ8 food frequency questionnaire and a four-day weighed food record and general linear model results (N=46).

Nutrient	EatWellQ8 FFQ[a], mean (SD)	WFR[b] four-day, mean (SD)	*P* value[c]	*P* value[d]
Energy (kcal)	2199 (862)	2119 (772)	.17[e]	N/A[f]
Total fat (g)	84.2 (39.0)	74.3 (27.6)	.08	.08
Total fat (%TE[g])	34.0 (7.4)	32.8 (8.5)	.47	.48
SFA[h] (g)	36.8 (18.2)	28.0 (12.0)	.001	.001
SFA (%TE)	14.8 (3.8)	11.9 (2.8)	.002	.002
MUFA[i] (g)	35.8 (17.9)	29.8 (13.9)	.04	.04
MUFA (%TE)	14.5 (4.1)	13.2 (5.1)	.17	.17
PUFA[j] (g)	14.5 (6.8)	11.7 (6.5)	.03	.03
PUFA (%TE)	5.9 (1.6)	5.2 (2.7)	.07	.07
Omega 3 (g)	0.1 (0.2)	0.3 (0.3)	.002	.002
Protein (g)	106 (50)	112 (55)	.11	.11
Protein (%TE)	19.3 (5.2)	21.00 (6.44)	.16	.16
Carbohydrate (g)	272 (117)	238 (102)	.03	.03
Carbohydrate (%TE)	49.8 (9.7)	45.4 (9.5)	.03	.03
Total sugars (g)	130 (74)	104 (63)	.03	.03
Total sugars (%TE)	23.1 (8.2)	19.3 (8.1)	.03	.02
Calcium (mg)	1191 (668)	1005 (508)	.07	.07
Total folate (µg)	345 (152)	288 (121)	.02	.02
Iron (mg)	13.4 (5.4)	11.9 (5.0)	.13	.13
Total carotene (µg)	5106 (4439)	3407 (3480)	.04	.04
Riboflavin (mg)	2.1 (1.2)	1.9 (1.1)	.40	.40
Thiamin (mg)	1.8 (0.8)	1.4 (0.5)	.001	.001
Vitamin B6 (mg)	2.5 (1.2)	2.0 (0.9)	.01	.01
Vitamin B12 (µg)	5.3 (3.5)	6.6 (11.3)	.37	.38
Vitamin C (mg)	178 (153)	126 (80)	.04	.04
Vitamin A RE[k] (µg)	1057 (766)	739 (497)	.02	.02
Retinol (µg)	223 (147)	202 (112)	.56	.57
Vitamin D (µg)	2.6 (1.9)	2.9 (1.8)	.26	.26
Vitamin E (mg)	11.8 (6.7)	10.6 (5.3)	.39	.39
Sodium (mg)	2552 (898)	2010 (815)	.001	.001

[a]FFQ: food frequency questionnaire.

[b]WFR: weighed food record.

[c]Controlled for energy.

[d]Controlled for energy and gender.

[e]Value derived from paired sample *t* test.

[f]N/A: not applicable.

[g]%TE: percentage total energy.

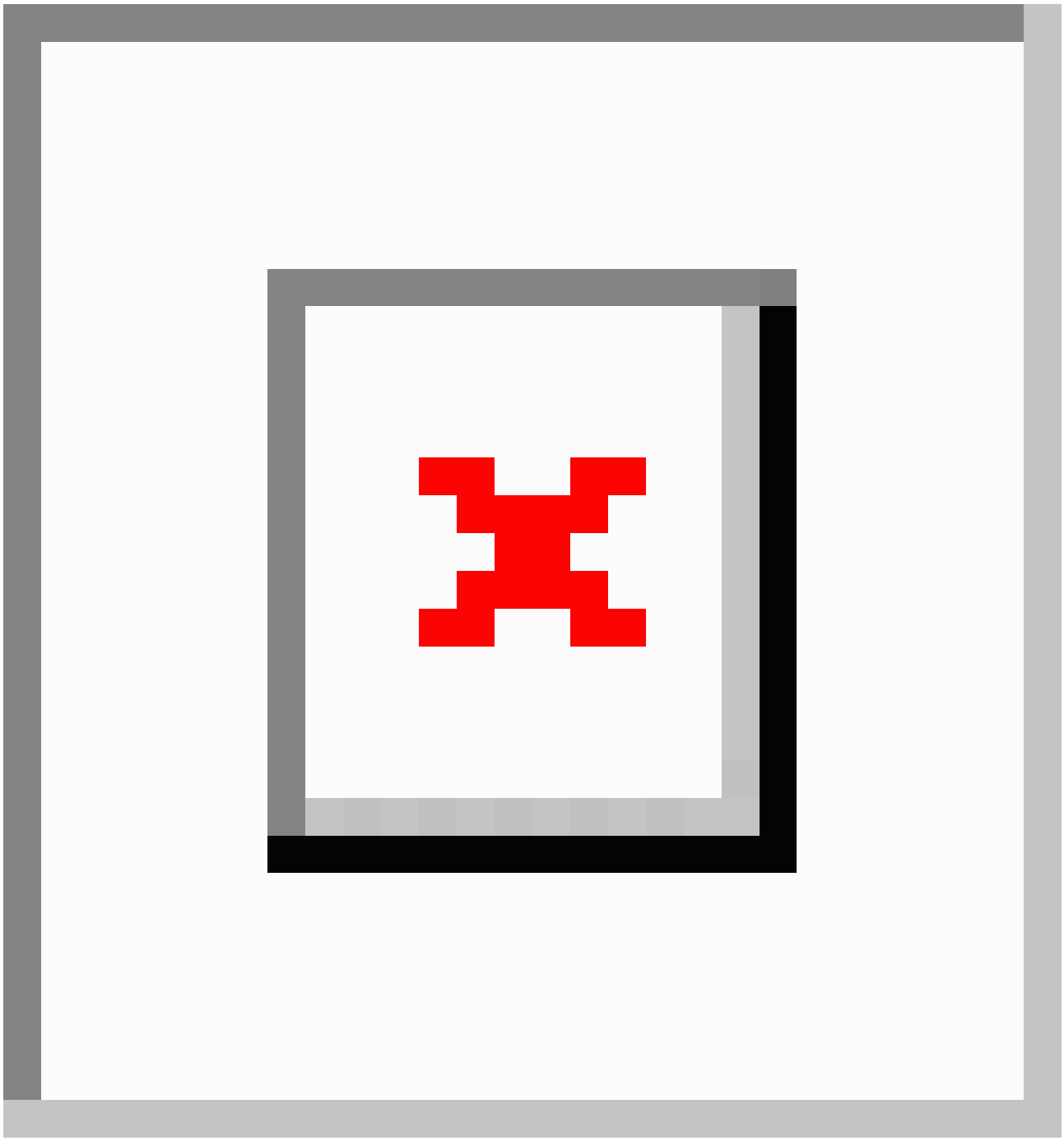
[h]SFA: saturated fatty acid.

[i]MUFA: monounsaturated fatty acid.

[j]PUFA: polyunsaturated fatty acid.

[l]RE: retinol equivalent.

Figure 5. Validation study of Bland-Altman plots comparing the EatWellQ8 food frequency questionnaire with a four-day weighed food record for (a) energy, (b) total carbohydrate, (c) protein, and (d) total fat with the bias (mean difference) and limits of agreement. The solid line represents the bias (mean difference), and the dotted lines represent the limits of agreement. %TE: percentage total energy; CHO: carbohydrates; FFQ: food frequency questionnaire; PRO: protein; WFR: weighed food record.



Comparison of Food Group Intakes Between the EatWellQ8 FFQ and the 4-Day WFR

The correlation coefficients and cross-classifications of mean food group intakes between the EatWellQ8 FFQ and the four-day WFR. SCCs ranged from 0.30 (bananas) to 0.88 (red meat; Multimedia Appendix 2 Table S6). Significant correlations were found for all food groups (95% CI 0.01-0.91), and weak 95% CIs were found for breakfast cereals and porridge (0.08) and bananas (0.01; *P*=.046). The cross-classification percentages of participants classified into quartiles of exact agreement ranged from 28% (green vegetables) to 65% (wholegrain and brown breads and rolls). Relatively high classifications of exact agreement plus adjacent were found, ranging from 71% (green vegetables) to 97% (red meat). The mean percentage of participants classified into quartiles of disagreement was 11% and, for extreme disagreement, it was 5%.

Discussion

Principal Findings

This study aimed to evaluate the reproducibility of the EatWellQ8 FFQ and to test its relative validity against a semiquantitative Kuwaiti PFFQ and a four-day WFR. The

EatWellQ8 FFQ has been developed to assess dietary and nutrient intakes in the EatWellQ8 study to investigate the effectiveness of delivering personalized face-to-face dietary advice compared with web-based dietary advice in Kuwait. It included images of 3 different portion sizes for each food item to aid in portion size estimation and food recognition. The need to develop a culturally sensitive FFQ that reflected the diet of the Kuwaiti population was necessary to avoid misclassifications of dietary intakes. The results of this study indicated that the EatWellQ8 FFQ is a suitable tool with moderate validity for the assessment of nutrient and food intake in a sample of healthy adults living in Kuwait.

Reproducibility

Overall, the EatWellQ8 FFQ had good reproducibility for the estimation of nutrient intake and food groups over a period of 4 weeks. The correlation coefficients for all nutrients were significant, compared with previous studies, and nearly all fell within the acceptable range of 0.5 to 0.7 for reproducibility trials proposed by Cade et al [5,16,25-29]. Similarly, strong associations were found between food groups with a mean SCC value of 0.67, which was comparable with previous web-based FFQ reproducibility studies by Fallaize et al [16] and Vereecken et al [30] that reported mean correlations of 0.75 and 0.64, respectively. However, a limitation in the trial by Vereecken et al [30] was the short assessment time between repeatability of the FFQs of only 1 to 2 weeks, which may have impacted the power of the trial. The usage of correlation analysis to assess agreement has been questioned as it only measures the degree of association between 2 variables and does not assess agreement [5,24]. Cross-classifications into quartiles of agreements and Bland-Altman plots were used to measure agreement. Analysis of cross-classifications of exact plus adjacent agreement of energy, nutrients, and food group intakes (mean value of 88%) indicated a high level of agreement and a low level of misclassification (<10%), similar to the results of previous web-based FFQ studies [10,16]. The high level of reproducibility may be in part due to the short period (4 weeks) between FFQ administrations, as true changes in dietary intakes are less likely to occur within a short period [31]. These data were also supported by the level of reproducibility from the Bland-Altman analysis for energy-controlled total protein, fat, and carbohydrate, which compared with findings from Fallaize et al [16] and Papazian et al [32]. Limitations to the trial by Papazian et al [32] were the relatively small sample size of 38 and the short interval time between FFQ administrations of 3 weeks, which may have impacted trial outcomes.

Results from several previous reproducibility studies have shown greater intakes in energy and nutrient intakes in the first FFQ compared with the second FFQ [10,16,26,29,33]. No significant differences between intakes were observed in this study, except for SFA and MUFA; however, quantitatively higher estimated energy and nutrient intakes were found in the initial administration of the EatWellQ8 FFQ compared with the second administration, which may be because of questionnaire boredom as a result of the short period between FFQs [5,34]. However, it has been proposed that the good reproducibility of the EatWellQ8 FFQ may be influenced by the short interval between FFQ administrations. It has been proposed that if the interval time between FFQs is short (1-6 months), participants' memory may influence the outcome, leading to overestimation in the reproducibility [16,35]. In contrast, underestimation was found in FFQs with longer time intervals (>6 months) because of changes in dietary habits [36]. We were keen for participants not to change dietary habits and explicitly asked for no change, which could have contributed to good reported reproducibility in our study. An additional factor that may have contributed to the good reproducibility is the use of photographs as an aid to food portion size estimation. It has been proposed that reproducibility is enhanced in FFQs that take into account food portion sizes, especially when participants are allowed to specify their own portion size [5].

Relative Validity

Overall, the results of the validation study demonstrated moderate to weak agreement between the EatWellQ8 FFQ and 2 dietary collection tools for the estimation of energy and nutrient intakes: a PFFQ and a four-day WFR. This was reflected by the higher level of bias being estimated by the EatWellQ8 FFQ for macronutrients (except for protein) and the level of disagreement in the cross-classifications, particularly in relation to food groups. This was also reflected by the large variations in the 95% CI range for both nutrients and food groups. The mean absolute intakes for most of the nutrients did not differ significantly between the tools. However, significant differences were found for specific FA (eg, SFA), which could possibly be because of differences in the food items presented in the FFQs. Similar to previous findings by Forster et al [15] and Beasley et al [26], compared with a PFFQ, the EatWellQ8 FFQ estimates of energy intake were significantly higher (P<.001). It has been reported that underestimation of dietary intake is common in PFFQs, which has been proposed to be because of errors such as skipped questions and a broad or vague use of portion size description [28].

With the exception of 2 nutrients (omega 3 FA and retinol), SCC fell within the range considered acceptable for FFQ validation trials from 0.4 to 0.7 [37,38]. The mean SCC for nutrients attained in this trial (r=0.54) was higher than that reported for a web-based FFQ validated against a PFFQ (r=0.47) and the one reported in the validation of a web-based diet history questionnaire against a four-day WFR [26,39]. The weakest SCC was found for specific FA (eg, omega 3 FA), and this finding was supported further in the results of cross-classifications, which also showed the least agreement for FA. This may be explained by a higher within-subject variation in fat intake. In this study, correlation coefficients for food groups were found to be relatively lower than the correlations found in trials by Forster et al [15] and Boeckner et al [40] that ranged from 0.42 to 0.90, which may be because of differences in the length of the PFFQs and number of food groups analyzed. Wide variations were observed in SCC between the EatWellQ8 FFQ and the PFFQ for food groups, suggesting that participants were able to estimate certain food items (eg, bananas) more accurately [41]. The proposed reasons for these variations are answering fatigue as a result of the length of the FFQs and may be a result of an overestimation of items that are perceived as healthy, such as vegetables and fruits, which is also common in other web-based FFQs [16,39].

The results of cross-classification for energy and nutrient intakes indicated that most participants were classified into exact plus adjacent quartiles that ranged from 76% to 93% and extreme disagreement or misclassification was <5% for most nutrients. Comparable cross-classifications that ranged from 77% to 99% were found when the Food4Me web-based FFQ was validated against the well-validated EPIC-Norfolk FFQ [15]. However, disagreement was high for food groups, especially for the food groups that were located at the end of PFFQ (eg, ice cream, creams, and desserts), suggesting answering fatigue. The results of the Bland-Altman plots showed moderate agreement between the methods for estimates of energy and energy-adjusted macronutrient intakes, and the least agreement was for %TE of fat. A possible reason for the disagreement between the tools may be participants' inability to assess portion sizes accurately using the PFFQ because of the lack of food photographs of portion sizes.

The EatWellQ8 FFQ was found to estimate higher energy, nutrient, and food group intakes compared with the 4-day WFR. These results were expected as it has been found in previous studies that FFQs that contain >100 items tend to show an overestimation in energy, nutrient, and food intake compared with WFR and 24-hour recalls, which may be because of underestimation of the latter methods or overestimation of FFQ [35,42]. Comparable percentages of individuals classified into quartiles of exact agreement (mean 49%) and exact plus adjacent agreement (mean 84%) were found between the EatWellQ8 FFQ and the four-day WFR for energy, nutrient intakes, and food groups, and low levels of disagreement were found. Cross-classifications were within the range reported in previous trials, which were both validated against WFRs [16,43]. The results of the Bland-Altman plots established good agreement between the 2 methods for energy and energy-adjusted macronutrient intakes. In addition, 28 of 30 nutrients measured had a correlation of higher than the 0.40 threshold recommended by Cade et al [5]. The relatively short period between administrations of the 2 methods (7 to 10 days) could have contributed to the high correlations. Highly variable SCCs were found for food group intakes that ranged from 0.29 for bananas to 0.88 for red meat, with a mean value of 0.55. Results from previous studies found similarly high variations that ranged from 0.09 to 0.95 [16,23,44,45]. It may be difficult to compare our results with previous studies because of differences in the type of food record used, food items included in specific food groups, and differences in the time intervals in each of the studies. Variations between the EatWellQ8 FFQ and the 4-day WFR were greatest for bananas, green vegetables, meat products, and tinned fruit or vegetables. This may be because of overestimations by the FFQ of foods perceived as healthy and can be because of the relatively short WFR (4 days), which may not reflect the individuals' dietary habits compared with the EatWellQ8 FFQ, which conveys the diet over the previous month, especially for foods that are not consumed regularly [46]. The wide variations observed in correlations between the EatWellQ8 FFQ and the four-day WFR may indicate whether volunteers could accurately estimate the consumption of some food items compared with others [16,41]. Compared with previous studies that compared FFQs with WFR, our results showed strong agreement for red meat (r=0.88) intake and fish and fish products, which are often consumed less frequently than other groups. A possible reason could be the differences in diets consumed in the Gulf compared with Western countries [47] and may be because of the short interval between the administration of the FFQ and the four-day WFR.

Strengths and Limitations

This study had many strengths, which included the comparison of the EatWellQ8 FFQ with 2 frequently used methods for dietary collection, one of them being the gold standard (a WFR) and the other a PFFQ to assess the reproducibility and relative validity of the EatWellQ8 FFQ [16]. Moreover, the sample size in this validation study was found to be adequate and comparable with the sample size used in previous studies [16,28,43,48].

Limitations of the validation study include the short interval time between administrations of the comparison tools (WFR and PFFQ) and the EatWellQ8 FFQ administration (7-10 days), which may have resulted in the similarity of responses between the tool. The addition of composite Kuwaiti dishes to the EatWellQ8 FFQ may have led to double reporting of food items and overestimation of caloric intake. In addition, the relatively large number of items (n=146) in the EatWellQ8 FFQ may have led to increased confusion and questionnaire fatigue or boredom. However, the validated comparison tools used (WFR and PFFQ) may also have been considered time consuming and burdensome for participants as the PFFQ, which included more than 200 items, takes approximately 35 to 40 minutes to complete, and the WFR required the weighing of food several times per day. Questionnaire tiring or boredom is particularly concerning as it can lead to underreporting of food items and may have therefore compromised the results [5]. However, the PFFQ was the only validated paper FFQ available for comparison in Kuwait, and WFR is a recommended comparison in validation studies.

Willet et al [49] suggested that the preferred sample size for FFQ validation studies is between 100 and 200, especially if they also take into account nutrient intakes; thus, the smaller sample size achieved in this study may be a limitation. However, it should be noted that the current trial did face recruitment issues. In addition, there was a high dropout rate in the trial, which may have resulted from the lack of an incentive upon completion or to participants' unwillingness to complete all 3 aspects of the study because of fatigue or boredom. Owing to the limited data available on the participants (eg, lack of sociodemographic and habitual data), it was not possible to account for known issues in self-reporting or deduce whether the sample was representative of the Kuwaiti population.

An additional limitation is the lack of biomarker data or alternative objective reference measures to validate the subjective questionnaire. It should also be noted that the EatWellQ8 FFQ and nutrient assessments did not consider supplements in the calculation of nutrition intakes, which may have led to their inaccurate assessments. Although seasonality is a common limitation in validation studies, it was not a concern in the current trial as the period of assessment for both validation and reproducibility fell between fall and winter. Owing to the lack of recent food composition tables specific to Kuwait, an

additional limitation may be inaccurate assessments of nutrient content of some food items, which further highlights the need for dietary assessment software that is specific for Kuwait.

Conclusions

In conclusion, the web-based self-administered EatWellQ8 FFQ, developed to assess energy and nutrient intake in healthy adults living in Kuwait, was found to have good reproducibility and moderate relative validity compared with a PFFQ and a four-day WFR. The results indicate that the novel web-based FFQ could be used as a dietary intake tool for the assessment of dietary intake in healthy adults living in Kuwait.

Acknowledgments

This work was funded by the Public Authority of Applied Education and Training (PAAET), Kuwait. The authors would like to thank the DDI for their assistance and support with the running of the study. In addition, they would like to thank the Food4Me consortium for the provision of portion size photographs. R Franco was sponsored by the National Council of Technological and Scientific Development of the Brazilian government.

Authors' Contributions

BA was responsible for data collection, analysis, and drafting of the manuscript. All authors have approved the final version for publication and have contributed to the research design, edited, and reviewed the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Unadjusted correlation coefficients, cross-classification and spearman correlation coefficients of quartiles of mean energy, nutrients and food group intakes from repeat measures of the web-based EatWellQ8 FFQ, paper-based food frequency questionnaire, and weighed food record.
[DOCX File , 34 KB - formative_v5i3e13591_app1.docx]

Multimedia Appendix 2
EatWellQ8 food frequency questionnaire food list.
[DOCX File , 62 KB - formative_v5i3e13591_app2.docx]

References

1. Noncommunicable diseases country profiles 2018. In: World Health Organization. Geneva: WHO; 2018:223.
2. Margetts B. FAO/WHO launch expert report on diet, nutrition and prevention of chronic diseases. Public Health Nutr 2007 Jan 02;6(4):323-325. [doi: 10.1079/phn2003481] [Medline: 12795818]
3. Roberts CK, Barnard RJ. Effects of exercise and diet on chronic disease. J Appl Physiol (1985) 2005 Jan;98(1):3-30 [FREE Full text] [doi: 10.1152/japplphysiol.00852.2004] [Medline: 15591300]
4. Alkerwi A, Guillaume M, Zannad F, Laufs U, Lair M, NESCAV project group. Nutrition, environment and cardiovascular health (NESCAV): protocol of an inter-regional cross-sectional study. BMC Public Health 2010 Nov 15;10(1):698 [FREE Full text] [doi: 10.1186/1471-2458-10-698] [Medline: 21078172]
5. Cade J, Thompson R, Burley V, Warm D. Development, validation and utilisation of food-frequency questionnaires – a review. Public Health Nutr 2007 Jan 02;5(4):567-587. [doi: 10.1079/phn2001318] [Medline: 12186666]
6. Thompson FE, Subar AF, Loria CM, Reedy JL, Baranowski T. Need for technological innovation in dietary assessment. J Am Diet Assoc 2010 Jan;110(1):48-51 [FREE Full text] [doi: 10.1016/j.jada.2009.10.008] [Medline: 20102826]
7. Falomir Z, Arregui M, Madueño F, Corella D, Coltell O. Automation of Food Questionnaires in Medical Studies: a state-of-the-art review and future prospects. Comput Biol Med 2012 Oct;42(10):964-974. [doi: 10.1016/j.compbiomed.2012.07.008] [Medline: 22898338]
8. Ngo J, Engelen A, Molag M, Roesle J, García-Segovia P, Serra-Majem L. A review of the use of information and communication technologies for dietary assessment. Br J Nutr 2009 Jul 01;101(S2):102-112. [doi: 10.1017/s0007114509990638] [Medline: 19594959]
9. Matthys C, Pynaert I, De Keyzer W, De Henauw S. Validity and reproducibility of an adolescent web-based food frequency questionnaire. J Am Diet Assoc 2007 Apr;107(4):605-610. [doi: 10.1016/j.jada.2007.01.005] [Medline: 17383266]
10. Labonté ME, Cyr A, Baril-Gravel L, Royer M, Lamarche B. Validity and reproducibility of a web-based, self-administered food frequency questionnaire. Eur J Clin Nutr 2012 Feb 21;66(2):166-173. [doi: 10.1038/ejcn.2011.163] [Medline: 21934698]
11. Elkum N, Al-Arouj M, Sharifi M, Shaltout A, Bennakhi A. Prevalence of childhood obesity in the state of Kuwait. Pediatr Obes 2016 Dec 11;11(6):30-34. [doi: 10.1111/ijpo.12090] [Medline: 26663908]

12. Musaiger AO, Al-Kandari FI, Al-Mannai M, Al-Faraj AM, Bouriki FA, Shehab FS, et al. Perceived barriers to weight maintenance among university students in Kuwait: the role of gender and obesity. Environ Health Prev Med 2014 May 9;19(3):207-214 [FREE Full text] [doi: 10.1007/s12199-013-0377-z] [Medline: 24402439]
13. European Prospective Investigation of Cancer EPIC-Norfolk food frequency questionnaire (FFQ) entry and processing documentation. EPIC (2013). 2018 Feb 12. URL: https://www.epic-norfolk.org.uk/for-researchers/ffq/ [accessed 2021-01-27]
14. Dehghan M, Al-Hamad N, McMillan CR, Prakash P, Merchant AT. Comparison of a semi-quantitative food frequency questionnaire with 24-hour dietary recalls to assess dietary intake of adult Kuwaitis. Saudi Med J 2009 Jan;30(1):159-161. [Medline: 19139794]
15. Forster H, Fallaize R, Gallagher C, O'Donovan CB, Woolhead C, Walsh MC, et al. Online dietary intake estimation: the Food4Me food frequency questionnaire. J Med Internet Res 2014 Jun 09;16(6):150 [FREE Full text] [doi: 10.2196/jmir.3105] [Medline: 24911957]
16. Fallaize R, Forster H, Macready AL, Walsh MC, Mathers JC, Brennan L, et al. Online dietary intake estimation: reproducibility and validity of the Food4Me food frequency questionnaire against a 4-day weighed food record. J Med Internet Res 2014 Aug 11;16(8):190 [FREE Full text] [doi: 10.2196/jmir.3355] [Medline: 25113936]
17. Franco RZ, Alawadhi B, Fallaize R, Lovegrove JA, Hwang F. A web-based graphical food frequency assessment system: design, development and usability metrics. JMIR Hum Factors 2017 May 08;4(2):13 [FREE Full text] [doi: 10.2196/humanfactors.7287] [Medline: 28483746]
18. Ryan NM, O'Donovan CB, Forster H, Woolhead C, Walsh MC. New tools for personalised nutrition: the Food4Me project. Nutrition Bulletin 2015 May 19;40(2):134-139. [doi: 10.1111/nbu.12143]
19. McCance and Widdowson’s the composition of foods: seventh summary edition: edition 7. London: Royal Society of Chemistry; 2014.
20. Chan W, Brown J, Lee SM, Buss DH. Meat, poultry and game: fifth supplement to the fifth edition of McCance and Widdowson's the composition of foods 1995. In: Department for Environment, Food and Rural Affairs. London: Royal Society of Chemistry; 1996.
21. Musaiger AO. Food composition tables for Kingdom of Bahrain. In: Semantic Scholar. Manama, Bahrain: Arab Center for Nutrition; 2011:150.
22. Dashti B, Al-Awadi F, Khalafawi M, Al-Zenki S, Sawaya W. Nutrient contents of some traditional Kuwaiti dishes. Food Chemistry 2001 Aug;74(2):169-175. [doi: 10.1016/s0308-8146(01)00111-x]
23. Hu F, Rimm E, Smith-Warner S, Feskanich D, Stampfer MJ, Ascherio A, et al. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. Am J Clin Nutr 1999 Feb;69(2):243-249. [doi: 10.1093/ajcn/69.2.243] [Medline: 9989687]
24. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet 1986 Feb;327(8476):307-310. [doi: 10.1016/s0140-6736(86)90837-8]
25. Chiu TH, Huang H, Chen K, Wu Y, Chiu JP, Li Y, et al. Relative validity and reproducibility of a quantitative FFQ for assessing nutrient intakes of vegetarians in Taiwan. Public Health Nutr 2013 Jun 20;17(7):1459-1466. [doi: 10.1017/s1368980013001560]
26. Beasley JM, Davis A, Riley WT. Evaluation of a web-based, pictorial diet history questionnaire. Public Health Nutr 2009 May 01;12(5):651-659. [doi: 10.1017/s1368980008002668]
27. Timon CM, Blain RJ, McNulty B, Kehoe L, Evans K, Walton J, et al. The development, validation, and user evaluation of Foodbook24: a web-based dietary assessment tool developed for the Irish adult population. J Med Internet Res 2017 May 11;19(5):158 [FREE Full text] [doi: 10.2196/jmir.6407] [Medline: 28495662]
28. Kristal AR, Kolar AS, Fisher JL, Plascak JJ, Stumbo PJ, Weiss R, et al. Evaluation of web-based, self-administered, graphical food frequency questionnaire. J Acad Nutr Diet 2014 Apr;114(4):613-621 [FREE Full text] [doi: 10.1016/j.jand.2013.11.017] [Medline: 24462267]
29. Cantin J, Latour E, Ferland-Verry R, Salgado SM, Lambert J, Faraj M, et al. Validity and reproducibility of a food frequency questionnaire focused on the Mediterranean diet for the Quebec population. Nutr Metab Cardiovasc Dis 2016 Feb;26(2):154-161. [doi: 10.1016/j.numecd.2015.11.003] [Medline: 26719221]
30. Vereecken CA, Bourdeaudhuij ID, Maes L. The HELENA online food frequency questionnaire: reproducibility and comparison with four 24-h recalls in Belgian-Flemish adolescents. Eur J Clin Nutr 2010;The HELENA online food frequency questionnaire:541-548. [doi: 10.1038/ejcn.2010.24] [Medline: 20216566]
31. Cade JE, Burley VJ, Warm DL, Thompson RL, Margetts BM. Food-frequency questionnaires: a review of their design, validation and utilisation. Nutr Res Rev 2007 Dec 14;17(1):5-22. [doi: 10.1079/nrr200370] [Medline: 19079912]
32. Papazian T, Hout H, Sibai D, Helou N, Younes H, Osta NE, et al. Development, reproducibility and validity of a food frequency questionnaire among pregnant women adherent to the Mediterranean dietary pattern. Clin Nutr 2016 Dec;35(6):1550-1556. [doi: 10.1016/j.clnu.2016.04.015] [Medline: 27161893]
33. Collins CE, Boggess MM, Watson JF, Guest M, Duncanson K, Pezdirc K, et al. Reproducibility and comparative validity of a food frequency questionnaire for Australian adults. Clin Nutr 2014 Oct;33(5):906-914. [doi: 10.1016/j.clnu.2013.09.015] [Medline: 24144913]

34. Wong SS, Boushey CJ, Novotny R, Gustafson DR. Evaluation of a computerized food frequency questionnaire to estimate calcium intake of Asian, Hispanic, and non-Hispanic White Youth. J Am Diet Assoc 2008 Mar;108(3):539-543. [doi: 10.1016/j.jada.2007.12.006] [Medline: 18313437]
35. Fernández-Ballart JD, Piñol JL, Zazpe I, Corella D, Carrasco P, Toledo E, et al. Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. Br J Nutr 2010 Jan 27;103(12):1808-1816. [doi: 10.1017/s0007114509993837] [Medline: 20102675]
36. Tsubono Y, Nishino Y, Fukao A, Hisamichi S, Tsugane S. Temporal change in the reproducibility of a self-administered food frequency questionnaire. Am J Epidemiol 1995 Dec 01;142(11):1231-1235. [doi: 10.1093/oxfordjournals.aje.a117582] [Medline: 7485070]
37. Willett WC. Future directions in the development of food-frequency questionnaires. Am J Clin Nutr 1994;59(1):171-174. [doi: 10.1093/ajcn/59.1.171s] [Medline: 8279418]
38. Willett WC. Nutritional epidemiology. In: Oxford Scholarship Online. New York: Oxford University Press; 1998.
39. González Carrascosa R, García Segovia P, Martínez Monzó J. Paper and pencil vs online self-administered food frequency questionnaire (FFQ) applied to university population: a pilot study. Nutr Hosp 2011;26(6):1378-1384. [doi: 10.1590/S0212-16112011000600027] [Medline: 22411386]
40. Boeckner LS, Pullen CH, Walker SN, Abbott GW, Block T. Use and reliability of the world wide web version of the block health habits and history questionnaire with older rural women. J Nutri Edu and Behavior 2002 Mar;34:20-24. [doi: 10.1016/s1499-4046(06)60307-2] [Medline: 12047825]
41. Fidanza F, Gentile MG, Porrini M. A self-administered semiquantitative food-frequency questionnaire with optical reading and its concurrent validation. Eur J Epidemiol 1995 Apr;11(2):163-170. [doi: 10.1007/bf01719482] [Medline: 7672070]
42. Verger EO, Armstrong P, Nielsen T, Chakaroun R, Aron-Wisnewsky J, Gøbel RJ, MetaCardis Consortium. Dietary assessment in the metacardis study: development and relative validity of an online food frequency questionnaire. J Acad Nutr Diet 2017 Jun;117(6):878-888. [doi: 10.1016/j.jand.2016.10.030] [Medline: 28024800]
43. Sanjeevi N, Freeland-Graves J, George GC. Relative validity and reliability of a 1-week, semiquantitative food frequency questionnaire for women participating in the supplemental nutrition assistance program. J Acad Nutr Diet 2017 Dec;117(12):1972-1982. [doi: 10.1016/j.jand.2017.05.013] [Medline: 28669810]
44. Feskanich D, Rimm EB, Giovannucci EL, Colditz GA, Stampfer MJ, Litin LB, et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. J Am Dietetic Asso 1993 Jul;93(7):790-796. [doi: 10.1016/0002-8223(93)91754-e] [Medline: 8320406]
45. Mazzeo T, Roncoroni L, Lombardo V, Tomba C, Elli L, Sieri S, et al. Evaluation of a modified Italian European prospective investigation into cancer and nutrition food frequency questionnaire for individuals with celiac disease. J Acad Nutr Diet 2016 Nov;116(11):1810-1816. [doi: 10.1016/j.jand.2016.04.013] [Medline: 27245762]
46. Apovian CM, Murphy MC, Cullum-Dugan D, Lin P, Gilbert KM, Coffman G, et al. Validation of a web-based dietary questionnaire designed for the DASH (Dietary Approaches to Stop Hypertension) diet: the DASH online questionnaire. Public Health Nutr 2009 Nov 16;13(05):615. [doi: 10.1017/s1368980009991996] [Medline: 19912673]
47. Musaiger AO. Nutritional disorders associated with affluence in Bahrain. Fam Pract 1990 Mar;7(1):9-13. [doi: 10.1093/fampra/7.1.9] [Medline: 2318377]
48. Lee Y, Park K. Reproducibility and validity of a semi-quantitative FFQ for trace elements. Br J Nutr 2016 Jul 05;116(5):864-873. [doi: 10.1017/s0007114516002622] [Medline: 27378401]
49. Willett WC. Nutritional epidemiology. In: Oxford Scholarship Online. New York: Oxford University Press; 2012.

Abbreviations

%TE: percentage total energy
DDI: Dasman Diabetes Institute
EPIC: European Prospective Investigation of Cancer
FA: fatty acid
FFQ: food frequency questionnaire
LOA: limits of agreement
MUFA: monounsaturated fatty acid
Na: sodium
NCD: noncommunicable disease
PFFQ: paper-based food frequency questionnaire
PN: personalized nutrition
SCC: Spearman correlation coefficient
SFA: saturated fatty acid
WFR: weighed food record

Edited by G Eysenbach; submitted 07.02.19; peer-reviewed by J Shin, P Wark; comments to author 02.10.19; revised version received 08.06.20; accepted 16.11.20; published 02.03.21.

Please cite as:
Alawadhi B, Fallaize R, Franco RZ, Hwang F, Lovegrove J
Web-Based Dietary Intake Estimation to Assess the Reproducibility and Relative Validity of the EatWellQ8 Food Frequency Questionnaire: Validation Study
JMIR Form Res 2021;5(3):e13591
URL: https://formative.jmir.org/2021/3/e13591
doi:10.2196/13591
PMID:33650974

Original Paper

A Rest Quality Metric Using a Cluster-Based Analysis of Accelerometer Data and Correlation With Digital Medicine Ingestion Data: Algorithm Development

Zahra Heidary[1], PhD; Jeffrey M Cochran[1], PhD; Timothy Peters-Strickland[1], MD; Jonathan Knights[1], PhD

Otsuka Pharmaceutical Development & Commercialization, Inc, Princeton, NJ, United States

Corresponding Author:
Jeffrey M Cochran, PhD
Otsuka Pharmaceutical Development & Commercialization, Inc
508 Carnegie Center Dr
Princeton, NJ, 08540
United States
Phone: 1 609 524 6788
Email: jeffrey.cochran@otsuka-us.com

Abstract

Background: Adherence to medication regimens and patient rest are two important factors in the well-being of patients with serious mental illness. Both of these behaviors are traditionally difficult to record objectively in unsupervised populations.

Objective: A digital medicine system that provides objective time-stamped medication ingestion records was used by patients with serious mental illness. Accelerometer data from the digital medicine system was used to assess rest quality and thus allow for investigation into correlations between rest and medication ingestion.

Methods: Longest daily rest periods were identified and then evaluated using a k-means clustering algorithm and distance metric to quantify the relative quality of patient rest during these periods. This accelerometer-derived quality-of-rest metric, along with other accepted metrics of rest quality, such as duration and start time of the longest rest periods, was compared to the objective medication ingestion records. Overall medication adherence classification based on rest features was not performed due to a lack of patients with poor adherence in the sample population.

Results: Explorations of the relationship between these rest metrics and ingestion did seem to indicate that patients with poor adherence experienced relatively low quality of rest; however, patients with better adherence did not necessarily exhibit consistent rest quality. This sample did not contain sufficient patients with poor adherence to draw more robust correlations between rest quality and ingestion behavior. The correlation of temporal outliers in these rest metrics with daily outliers in ingestion time was also explored.

Conclusions: This result demonstrates the ability of digital medicine systems to quantify patient rest quality, providing a framework for further work to expand the participant population, compare these rest metrics to gold-standard sleep measurements, and correlate these digital medicine biomarkers with objective medication ingestion data.

(JMIR Form Res 2021;5(3):e17993) doi:10.2196/17993

KEYWORDS

serious mental illness; rest quality; actimetry; behavioral health; digital medicine; accelerometer; medication adherence

Introduction

Lack of adherence to medication regimens is a significant public health issue that contributes to increased health care utilization [1,2]. Adherence is of particular concern in patients with serious mental illness (SMI), including schizophrenia, bipolar disorder, and major depressive disorder, with estimates of nonadherence as high as 60% [1,3]. Within this population, effective pharmacotherapy is critical for mitigating the risk of serious adverse events, such as psychosis, symptom recurrence, poor social functions, hospitalizations, and suicide attempts [4,5]. Conventional methods of inferring medication adherence to pharmacotherapy have limited utility, as they are typically either subjective or acquire only surrogate markers of medication ingestion [6]. Thus, there is a clear, unmet clinical need for adherence monitoring that digital medicine is ideally suited to address. In this context, digital medicine refers to a system that combines an active pharmaceutical and an ingestible sensor that

communicates to a mobile app or web-based application to record that patients have taken their medication at a specific time [7], providing an objective measure of ingestion.

Another potentially useful biomarker for patients with SMI is disruption in sleep [8,9] and activity patterns [10,11], both of which have been significantly linked to mental health. The digital medicine system (DMS) used here noninvasively records activity-related parameters such as step count, physical patient orientation, and heart rate. Due to battery optimization for ingestion monitoring, the DMS does not record these metrics at a sufficiently high temporal resolution to perform common sleep identification techniques, such as heart rate variability analysis [12,13]. However, accelerometer-based actigraphy has been extensively validated against gold-standard sleep measurements like polysomnography [10,14-17]; thus, the available data from the DMS can noninvasively provide valuable insight into patterns of rest and activity in patients with SMI at the currently available temporal resolution and battery consumption while simultaneously recording previously unavailable objective medication ingestion data with a single device in a natural care setting.

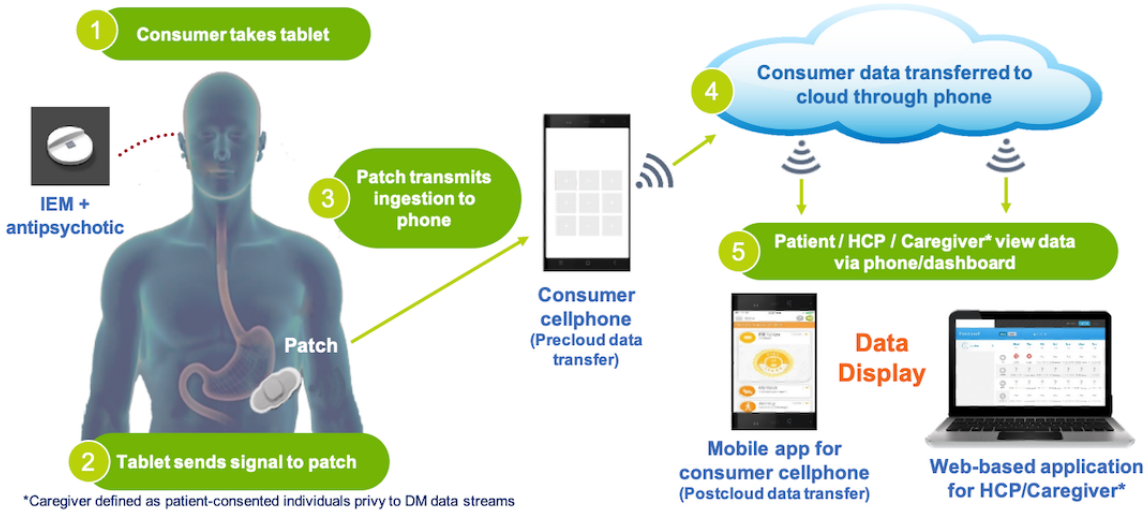
Thus, the goals of this study were twofold. First, 3-axis accelerometer data from the DMS and a novel analysis algorithm were used to identify and analyze patients' longest period of rest during each day they were on the system. The relative quality of these rest periods was quantified using a modified k-means clustering algorithm of accelerometer-derived features; previous actimetry research has correlated these types of activity features with wakefulness in sleep studies [14-17]. These activity features should thus also be correlated with measurements of rest quality as defined by the National Sleep Foundation, such as sleep efficiency and wake after sleep onset [18]. The stability of the rest period duration and consistency of the rest period starting time were also assessed across all days for each patient, as measurements of rest duration [18-20] and starting time [9,10,18] have also been correlated with quality of rest. Finally, preliminary correlations between rest and medication ingestion records were explored.

Methods

DMS

The DMS used here has been described previously [21]. Briefly, the DMS is a combination of an ingestible sensor embedded in an active pharmaceutical and a sensor patch that is attached to the torso, which records ingestion events and measures activity via a 3-axis accelerometer and an estimated sample heart rate with a single-lead electrocardiogram (ECG). The patch, which the patient must apply, also contains a temperature sensor and an impedance sensor and is designed to be worn for 7 days. The collected data is uploaded and stored in the cloud via a mobile phone app that also displays a patient dashboard (see Figure 1).

Figure 1. Schematic of the DMS. The DMS consists of an ingestible sensor embedded within an aripiprazole tablet. Upon ingestion, this sensor communicates with a patient-worn patch, which records the ingestion event and provides measures of activity and heart rate using an accelerometer and ECG, respectively, as well as temperature and impedance across the patch. The patch then communicates all data to a smartphone app, which stores the data on a secure cloud server that can be accessed by patients or designated caregivers via a mobile or web-based dashboard. DM: digital medicine; DMS: digital medicine system; ECG: electrocardiogram; HCP: health care provider; IEM: ingestible event marker.



This system is designed for monitoring medication ingestion in patients with SMI. Each measured ingestion event is recorded with a timestamp. The accelerometer data are measured across a 14-second sample every minute, and a built-in algorithm converts the raw data to a step count, mean acceleration along all 3 axes (a_x, a_y, a_z), and mean body orientation angle (θ), which are read out at 1-minute intervals. When the patch initially observes activity, the accelerometer is calibrated such that the x-axis is approximately oriented along the longitudinal axis of the body regardless of the orientation of the patch. The ECG provides the mean heart rate every 5 minutes.

Study Population

The population used in this analysis comprised 102 participants enrolled across 2 clinical studies (NCT02219009, NCT02722967) [22,23]. Table 1 contains the demographic and clinical characteristics for these patients. All participants were previously diagnosed with schizophrenia, bipolar disorder, or major depressive disorder and were regularly receiving

once-daily doses of the atypical antipsychotic aripiprazole. For the duration of the study, patients received the digital version of aripiprazole. Both studies were reviewed and approved by the appropriate institutional review boards, and patients were all deemed capable of using the DMS and provided signed informed consent. To ensure the reliability of individual rest metrics, the data set presented here excludes 14 additional patients enrolled in the studies that did not have at least 7 days of recorded data from the DMS patch.

Table 1. Demographic and clinical characteristics of the sample population.

Characteristic	Values (N=102)
Gender, n (%)	
Female	40 (39)
Male	62 (61)
Age (years), mean (SD)	45.9 (11.3)
Race, n (%)	
Black or African American	56 (55)
White	39 (38)
Asian	5 (5)
Other	2 (2)
Ethnicity, n (%)	
Hispanic or Latino	5 (5)
Not Hispanic or Latino	97 (95)
Diagnosis, n (%)	
Schizophrenia	71 (70)
Bipolar 1 disorder	21 (21)
Major depressive disorder	10 (10)
Observed ingestion (%)	
Median (IQR)	71.3 (28.9)
Range	14.8-96.6
Observed ingestion duration (days)	
Median (IQR)	54 (14)
Range	7-64

Definition of Longest Rest Periods

The DMS patch provides step count and posture angle measurements from the accelerometer every minute and the mean heart rate from the ECG every 5 minutes. These measurements were partitioned into 15-minute nonoverlapping time intervals, which were assessed for data quality. A 15-minute interval was considered analyzable if all 3 of the following conditions were met: (1) there was no point in the interval at which the patch was pairing with the mobile app, (2) there were at least 10 collected accelerometer measurements within the interval, (3) there were at least 2 ECG records within the window.

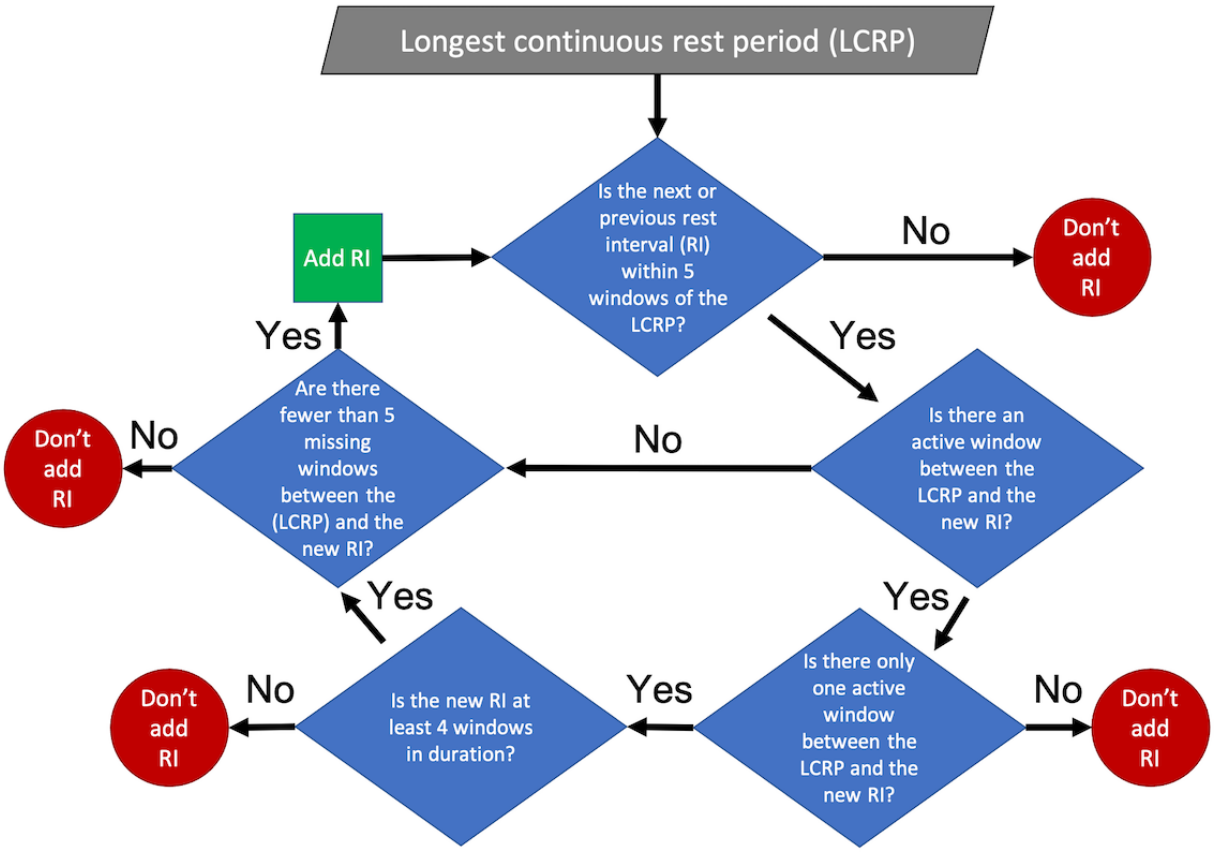
Any intervals that failed 1 or more of these criteria were labelled as missing, ensuring that intervals were only used if they contained sufficient reliable patch data.

For analyzable intervals, each 1-minute accelerometer record was then examined, and if the posture angle of that record was less than 30° away from horizontal, it was classified as rest.

Based on these accelerometer measurement labels, the analyzable 15-minute intervals were then classified as either rest (if greater than 70% of the records in that interval were classified as rest) or active (if the rest criterion was not met).

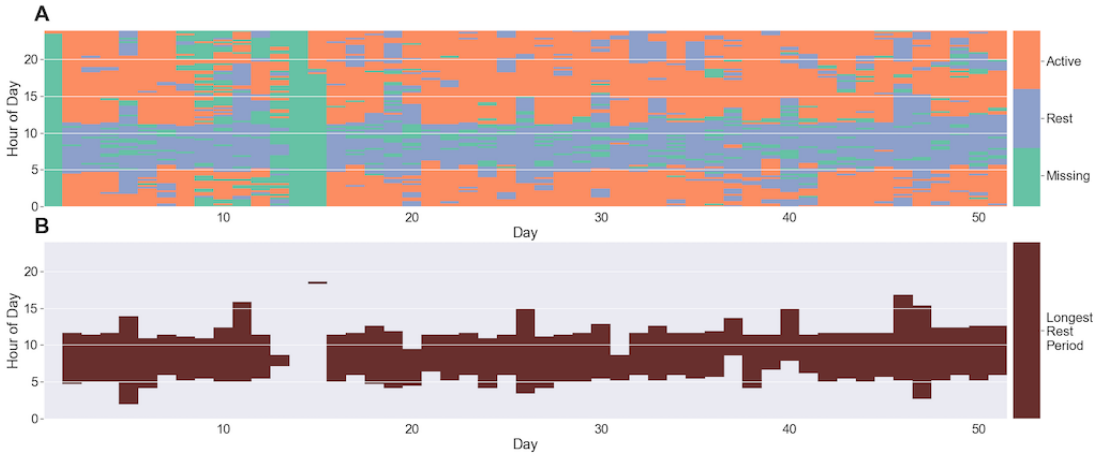
These 15-minute windows were then used to define each patient's longest period of rest for a given day. The first step was to identify the longest period of consecutive rest intervals that did not contain any active or missing intervals; however, this simple longest continuous rest period (LCRP) was likely an insufficient representation of the patient's true longest rest period (LRP). For one, intervals labeled as missing could artificially shorten these periods. Additionally, a single active interval is not necessarily indicative of the end of that rest period; for example, sleep is often interrupted by a short period of activity. Thus, an algorithm was used, as illustrated in Figure 2, to extend these longest periods of rest to more accurately capture the duration of each rest period.

Figure 2. Flowchart for finding the longest rest period. The first step in defining the longest rest period for a given day is to locate the LCRP (ie, the longest duration of consecutive 15-minute rest windows without any active or missing windows). The algorithm described in the flowchart is then used to determine if nearby windows are added to the LCRP.



Briefly, sequences of continuous rest windows that were within 5 intervals of either the beginning or end of the initial longest LCRP were added to the LRP if the intervening intervals were all missing (ie, none of them were labeled as active). If exactly 1 active interval was present between the longest interval and another rest interval in any of the 5-interval windows, only sequences that contained 4 or more consecutive rest intervals were added to the LRP. If more than 2 intervals between the original period and the additional rest period were labeled as active, the new rest period was not added to the LRP. Figure 3 provides the 15-minute interval rest designations and identified LRPs for a single patient across all days.

Figure 3. Rest state interval designations and longest rest periods for a single subject. (A) All 15-minute windows were classified as either rest, active, or missing, according to the procedure discussed in the Methods section. (B) The longest rest period for each day was identified using the procedure shown in Figure 2.

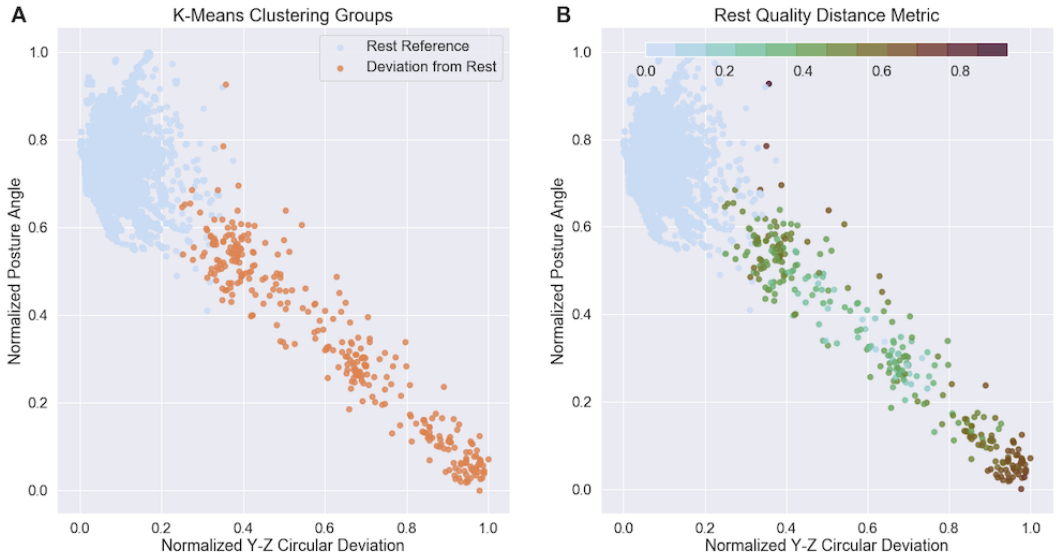


Rest Features and Metrics

With the LRP identified for each patient on each day, features were then developed to quantify the relative quality of rest within these intervals. For this purpose, 3-minute rolling windows of the accelerometer data, consisting of 3D acceleration (a_x, a_y, a_z), posture angle (θ), and step count, were used, resulting in 13 individual data points for each 15-minute window. In all, 35 features (see Multimedia Appendix 1 Table S1) were tested, and 4 features were chosen via a feature agglomeration technique which is similar to agglomeration clustering but uses recursive merging of features instead of samples [24]. The feature agglomeration clustering algorithm was run independently for each patient in the data set, and 4 features that were consistently grouped into separate clusters across all patients were chosen.

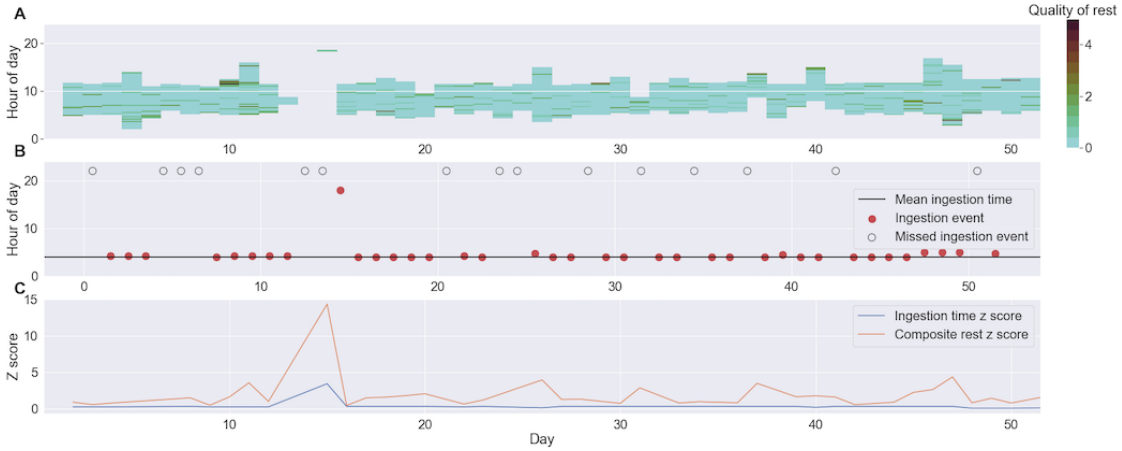
These 4 features were then normalized to their respective ranges and used in a k-means clustering algorithm with 2 groups (k=2); 2 groups were purposely chosen to identify both a rest-reference (RR) cluster, in which the mean posture angle was closer to a horizontal position, and a deviation-from-rest (DFR) cluster. A Euclidean distance across the 4D feature space quantified the deviation from the rest cluster for each data point; because the 4 individual features used in this model are associated with rest quality, this distance can be interpreted as a metric for the quality of rest. For all points in the DFR cluster, the distance was calculated from that point to the center of the RR cluster in the 4D feature space. All points that were designated as part of the RR cluster were given a distance measure of 0; thus, low values indicated higher rest quality, while higher values indicated poorer rest quality. Figure 4 shows a sample patient cluster diagram with the Euclidean distance calculation.

Figure 4. K-means clustering and relative rest quality in a single patient for 2 of the 4 features. (A) K-means clustering diagram plotted across 2 dimensions representing 2 of the 4 features in the clustering algorithm. All data points for each patient were classified using 2 clusters defined as the rest reference cluster (in blue) and the deviation from rest cluster (in orange). (B) Relative rest quality distance metric calculation. A quality of rest metric, represented here by variations in color, was also determined for each data point by calculating the Euclidean distance of the point from the center of the rest reference cluster. All points within the rest reference cluster were assigned a distance of 0. Note that the rest quality color does not have a perfect correlation with apparent distance on this 2D graph because the Euclidean distance is calculated across all 4 feature dimensions.



Within each 15-minute window, the distance measurements of all points were then summed to create a single rest quality metric. Figure 5A displays these LRPs and the calculated rest quality in each window for a single patient. The relative rest quality for each LRP was defined as the mean of the relative rest quality in all 15-minute windows within that LRP. Finally, for a given patient, the mean and SD of the relative rest quality of the LRP across all days were used to transform the rest quality metric for each LRP into a *z* score.

Figure 5. Daily relative rest quality and ingestion for a single patient. (A) The longest rest period for each day. Each bar represents the time windows that constitute the longest rest period for a given day. The color of each 15-minute interval within each bar represents the rest quality metric for that window. (B) Ingestion times for each day. Recorded ingestion events are marked with a red circle, missed ingestions are marked with an empty black circle, and the mean ingestion time across all days is represented by the horizontal line. (C) Ingestion and composite rest z scores. The composite rest z score (in orange), which is the sum of the rest quality, starting time, and duration z score, and the ingestion time z score (in blue) are plotted across all days for the same patient. We searched for data points where outliers in the composite rest z score occurred near the same day as a corresponding outlier in the ingestion time z score (eg, on day 15). Note that days without recorded patch data were excluded from the analysis.

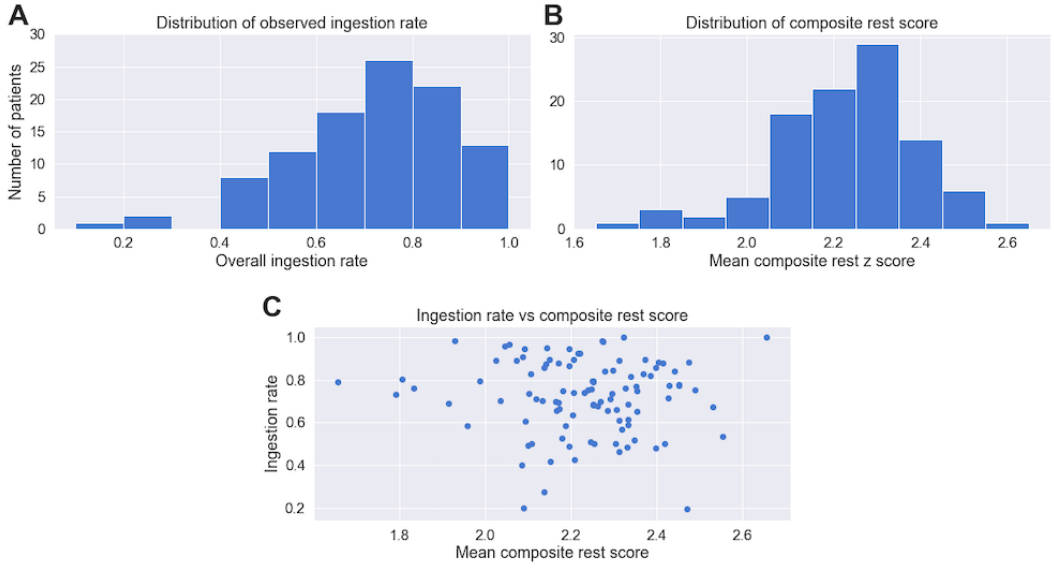


The *z* score metrics of the stability of the LRP duration and of the consistency of the LRP starting time were similarly calculated across all days for each participant. A composite rest *z* score for each LRP could then be calculated by summing the absolute values of the quality, stability, and consistency *z* score measurements.

Ingestion Metrics

Patient medication ingestion was quantified with 2 metrics. The first was the patient's overall observed ingestion rate, calculated as the number of days during which an ingested dose was recorded divided by the expected number of ingested doses across the entire regimen, which was defined as the time from the first day with recorded patch data to the last day with an ingestion record (see Table 1 and Figure 6A). The second metric was a *z* score of daily ingestion time using the mean and SD of ingestion time across all days for a given patient. These ingestion time markers were compared to the previously described rest metrics by searching for single-day outliers in the rest *z* score within 1 day of similar outliers in the ingestion time *z* score (see Figure 5B).

Figure 6. Ingestion rate and composite rest score. (A) Distribution of ingestion rate by patient. Ingestion rate was defined as the fraction of days that contained an ingestion event. Notice that there are relatively few patients with a poor ingestion rate, which made developing robust classification algorithms difficult. (B) Distribution of mean composite rest score across all days for each patient. (C) Ingestion rate versus mean composite rest score. There is no clear correlation between the ingestion rate and the composite rest score. Note, however, that there are no patients with very poor adherence (ie, an ingestion rate less than 0.5) that have a low composite rest score, which would be indicative of better rest.



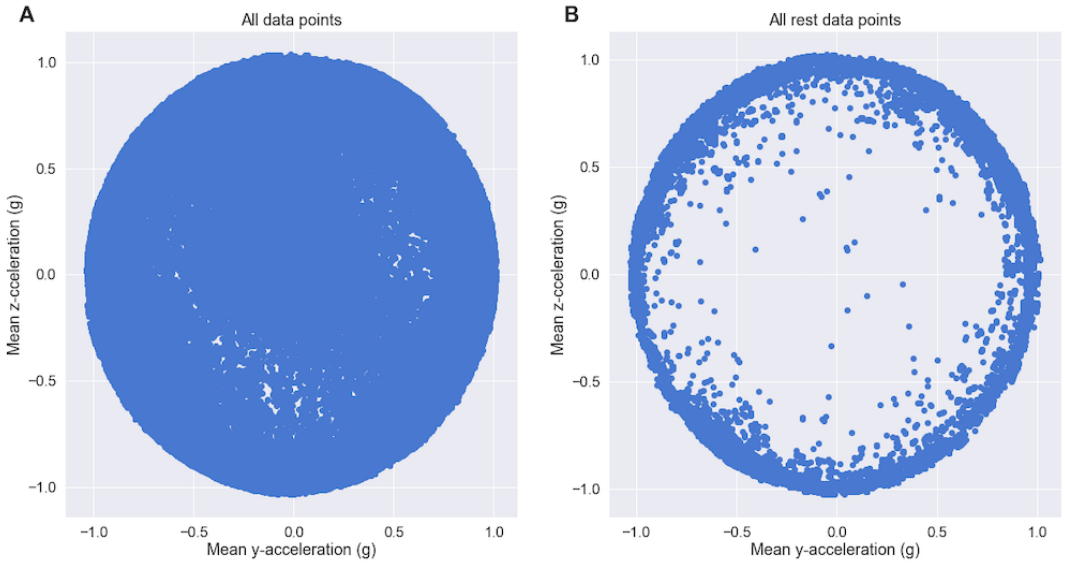
Results

A total of 35 features, calculated across 3-minute rolling windows, were explored to characterize rest in this population of 102 patients with SMI. There were no significant observed differences across the 3 indications in this population (ie, schizophrenia, bipolar disorder, or major depressive disorder) for any of the forthcoming metrics. The 4 features chosen by the aforementioned agglomeration technique were the mean of the circular deviation in the y–z acceleration plane, the mean of the posture angle, the mean of the 3D acceleration norm, and the SD of the x acceleration (see Table 2), which were all calculated across 3-minute rolling windows. The circular deviation in the y–z acceleration, which represents the degree to which the combination of the y and z components of the acceleration differ from the full 1-g resting acceleration, was particularly successful at differentiating rest windows from those designated as active (see Figure 7).

Table 2. Rest features.

Name	Definition
y–z acceleration circular deviation (mean)	
Posture angle (mean)	
Acceleration norm (mean)	
x acceleration (SD)	$\sigma(\alpha_x)$

Figure 7. The y–z acceleration circular deviation. (A) Mean z-axis acceleration versus mean y-axis acceleration for all data points in units of g. (B) Mean z-axis acceleration versus mean y-axis acceleration for all rest data points in units of g. Note that, for the remaining data points, the y-axis and z-axis accelerations tend to cluster around a circle with radius 1. Thus, the circular deviation in the y–z plane is an effective feature for differentiating rest from nonrest.



These 4 features were used in the k-means clustering algorithm to create the RR and DFR groups as well as the relative rest quality distance metric (see Figure 4). The distance metrics for each 3-minute window were summed over each 15-minute interval, and the mean across the LRP was calculated for each patient on each day, providing a metric of rest quality throughout each LRP (see Figure 5A). Note that the value of this metric is inversely related to the quality of rest. The start time and duration of each LRP was also calculated. The *z* scores for this rest quality, along with the start time stability and duration of the LRPs, were calculated with respect to the mean and SD of all LRPs within a given patient's data set. A composite rest score was then defined by summing the absolute value of the 3 *z* score metrics for each LRP. This composite score was thus a measure of the combined magnitude of deviation from a patient's typical pattern for these 3 rest metrics. The distribution of the mean value of this composite rest score for all patients can be seen in Figure 6B.

These rest metrics were then compared to the overall ingestion rate for each patient, which was defined as the fraction of days a patient was using the DMS that contained an ingestion event. This study population had a notable lack of patients with poor adherence; for example, only 15% of patients had an observed ingestion percentage of 50% or less (see Figure 6A). Patients with lower adherence did appear to have a relatively low quality of rest as quantified by the composite rest metric; there were no patients with an ingestion rate equal to or less than 50% that

had a low composite rest *z* score (see Figure 6C), which signifies better rest. However, the asymmetry in high versus low adherence and the small sample size precluded statistically significant conclusions about the overall correlation between low adherence and rest quality.

We were, however, able to explore correlations between daily outliers in the composite rest score with outliers in ingestion time that occurred within 1 day of the rest outlier. One such example can be seen in Figure 5C. Although this sample does not contain sufficient events of this type to develop quantitative conclusions, this could be a potentially useful marker for identifying rest-related adherence risk factors in the time domain.

Discussion

Principal Results

This DMS combines objective, time-stamped records of medication ingestion with noninvasive physiological markers of activity and heart rate in patients with SMI. It thus provides a unique opportunity to measure the quality of rest and explore correlations between rest and objective medication ingestion data. Here, algorithms were developed to both identify rest periods and quantify the relative quality of the rest within these periods using only the data available from the DMS without increasing battery consumption.

The LRP identification algorithm used accelerometer data, as provided by the DMS, to first find 15-minute windows that could be classified as rest and then to combine these windows to find a daily period of longest rest. Thirty-five potential accelerometer-measured features were explored, and the four best were selected via a feature agglomeration methodology (see Table 2). These features included the acceleration norm and SD of the x-axis acceleration, with larger values associated with greater activity. Another feature was the posture angle, with 0 corresponding to a patient lying down and deviations representing a greater degree of uprightness. The most novel feature was the circular deviation of the y–z acceleration, which is also correlated with rest. If a patient were lying still and perfectly horizontally, the x-axis (ie, head-to-toe) acceleration would be small, and thus the total y–z magnitude of acceleration would be near 1. This would be an example of high-quality rest and result in a y–z circular deviation near 0 (see Figure 4). Larger y–z circular deviations can thus be interpreted as an indicator of decreased rest quality.

These 4 features served as the input for an unsupervised k-means clustering algorithm (k=2) that included a 4D Euclidean distance metric for each data point to the center of the RR cluster, which was then applied to all windows in each LRP. This distance metric thus could quantify the degree to which a patient deviates from the RR cluster. Because each of the 4 features that constitute this distance can individually be correlated with rest quality, this distance metric can thus be interpreted as a deviation from quality rest. All patients used in this clustering algorithm were examined to ensure that their data points were not well represented by a single rest cluster. It should be noted that other populations could contain such a patient, which could minimize the contrast between the RR and DFR clusters. This algorithm demonstrates the ability of the DMS to quantify patient-specific rest quality despite the relatively low temporal resolution of the available accelerometer data due to the need to prioritize power consumption of the ingestion detection module. Thus, rest quality can be quantified for patients already engaged with the DMS without the use of an additional activity tracker that would require further patient engagement.

Exploratory correlations between the calculated rest metrics and ingestion data were hindered by the sample's asymmetry in medication adherence, in which few patients exhibited poor ingestion rates. Thus, no statistically significant conclusions about rest and overall adherence were drawn. Patients with poor adherence did tend to have larger composite rest *z* scores, which is indicative of inconsistent rest quality; however, patients that were largely adherent to the regimen did not necessarily have more consistent rest as measured by the composite rest *z* score. A *z* score–based comparison of daily outliers in rest and ingestion outliers within 1 day could also serve as a useful metric for exploring the time-domain prediction of variability in medication ingestion time in a study with a larger sample.

Limitations

Thus, this analysis is an important first step in leveraging the available data from this DMS system to quantify personalized rest metrics for eventual correlation with objective ingestion data, providing insight into the behavioral context of medication adherence for patients with SMI. However, the study does have several limitations that can be addressed with future research. The most notable is the lack of patients with very poor adherence, which prevented the use of rest data to truly classify patients by their adherence. This shortcoming can be easily addressed by accruing more participants in future studies. It would also be interesting to apply these rest metrics to patients in a controlled sleep study, which would enable an assessment of the degree to which these markers for rest are an accurate surrogate of sleep. Finally, an inherent limitation of the DMS is that it provides accelerometer data at only 1-minute intervals. This study nonetheless demonstrates that this relatively sparse data can still be used to effectively monitor patient rest.

There are many other future analysis directions that could be pursued with this data. Two of the most direct extensions of this study would be to more fully incorporate the system's measurement of heart rate and to begin exploring markers that quantify a patient's activity level throughout the day.

Conclusions

Data from the accelerometer in a DMS that provides objective time-stamped medication ingestion records were collected and used to develop novel algorithms for identifying and assessing the quality of daily rest periods for individual patients. A lack of patients with poor adherence in the sample population prevented the use of a quantitative classifier; however, pilot explorations of the relationship between these rest metrics and ingestion provided several insights. Patients with poor adherence experienced lower quality of rest, while patients with high adherence did not exhibit a consistent rest pattern. Additionally, the correlation of outliers in composite rest score with outliers

in ingestion time is an interesting potential application of this algorithm. This study is an important first demonstration of the ability of the DMS to track patient rest, which provides a framework for future correlation of DMS-based biomarkers with medication adherence in patients with SMI.

Acknowledgments

Funding for this study was provided by Otsuka Pharmaceutical Development & Commercialization, Inc.

Authors' Contributions

ZH and JK made substantial contributions to the design and conception of this study, the data collection, and the analysis and interpretation of the data. JMC participated in the analysis and interpretation of the data and created the first draft of the manuscript. TP contributed to the design and conception of this study and the data collection. All authors provided intellectual contributions to the manuscript reviews and approved the final version for submission.

Conflicts of Interest

ZH, JMC, JK, and TPS are employees of Otsuka Pharmaceutical Development & Commercialization, Inc.

Multimedia Appendix 1
List of all accelerometer features used.
[DOCX File , 15 KB - formative_v5i3e17993_app1.docx]

References

1. Kane JM, Kishimoto T, Correll CU. Non-adherence to medication in patients with psychotic disorders: epidemiology, contributing factors and management strategies. World Psychiatry 2013 Oct;12(3):216-226 [FREE Full text] [doi: 10.1002/wps.20060] [Medline: 24096780]
2. Cutler DM, Everett W. Thinking outside the pillbox--medication adherence as a priority for health care reform. N Engl J Med 2010 Apr 29;362(17):1553-1555. [doi: 10.1056/NEJMp1002305] [Medline: 20375400]
3. Valenstein M, Ganoczy D, McCarthy JF, Myra Kim H, Lee TA, Blow FC. Antipsychotic adherence over time among patients receiving treatment for schizophrenia: a retrospective review. J Clin Psychiatry 2006 Oct;67(10):1542-1550. [doi: 10.4088/jcp.v67n1008] [Medline: 17107245]
4. Velligan DI, Sajatovic M, Hatch A, Kramata P, Docherty JP. Why do psychiatric patients stop antipsychotic medication? A systematic review of reasons for nonadherence to medication in patients with serious mental illness. Patient Prefer Adherence 2017;11:449-468 [FREE Full text] [doi: 10.2147/PPA.S124658] [Medline: 28424542]
5. El-Mallakh P, Findlay J. Strategies to improve medication adherence in patients with schizophrenia: the role of support services. Neuropsychiatr Dis Treat 2015;11:1077-1090 [FREE Full text] [doi: 10.2147/NDT.S56107] [Medline: 25931823]
6. Andrade SE, Kahler KH, Frech F, Chan KA. Methods for evaluation of medication adherence and persistence using automated databases. Pharmacoepidemiol Drug Saf 2006 Aug;15(8):565-74; discussion 575 [FREE Full text] [doi: 10.1002/pds.1230] [Medline: 16514590]
7. FDA approves pill with sensor that digitally tracks if patients have ingested their medication. The U.S. Food and Drug Administration. URL: https://tinyurl.com/y3zbavse [accessed 2018-02-22]
8. Pandi-Perumal SR, Kramer M, editors. Sleep and Mental Illness. Cambridge: Cambridge University Press; 2010.
9. Jagannath A, Peirson SN, Foster RG. Sleep and circadian rhythm disruption in neuropsychiatric illness. Curr Opin Neurobiol 2013 Oct;23(5):888-894. [doi: 10.1016/j.conb.2013.03.008] [Medline: 23618559]
10. Jones SH, Hare DJ, Evershed K. Actigraphic assessment of circadian activity and sleep patterns in bipolar disorder. Bipolar Disord 2005 Apr;7(2):176-186. [doi: 10.1111/j.1399-5618.2005.00187.x] [Medline: 15762859]
11. Soundy A, Roskell C, Stubbs B, Vancampfort D. Selection, use and psychometric properties of physical activity measures to assess individuals with severe mental illness: a narrative synthesis. Arch Psychiatr Nurs 2014 Apr;28(2):135-151. [doi: 10.1016/j.apnu.2013.12.002] [Medline: 24673789]
12. Elsenbruch S, Harnish MJ, Orr WC. Heart rate variability during waking and sleep in healthy males and females. Sleep 1999 Dec 15;22(8):1067-1071. [doi: 10.1093/sleep/22.8.1067] [Medline: 10617167]
13. Tobaldini E, Nobili L, Strada S, Casali KR, Braghiroli A, Montano N. Heart rate variability in normal and pathological sleep. Front Physiol 2013 Oct 16;4:294 [FREE Full text] [doi: 10.3389/fphys.2013.00294] [Medline: 24137133]
14. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. Sleep 2003 May 01;26(3):342-392. [doi: 10.1093/sleep/26.3.342] [Medline: 12749557]
15. Sadeh A, Sharkey KM, Carskadon MA. Activity-based sleep-wake identification: an empirical test of methodological issues. Sleep 1994 Apr;17(3):201-207. [doi: 10.1093/sleep/17.3.201] [Medline: 7939118]

16. Reid K, Dawson D. Correlation between wrist activity monitor and electrophysiological measures of sleep in a simulated shiftwork environment for younger and older subjects. Sleep 1999 May 01;22(3):378-385. [doi: 10.1093/sleep/22.3.378] [Medline: 10341389]
17. Hedner J, Pillar G, Pittman SD, Zou D, Grote L, White DP. A novel adaptive wrist actigraphy algorithm for sleep-wake assessment in sleep apnea patients. Sleep 2004 Dec 15;27(8):1560-1566. [doi: 10.1093/sleep/27.8.1560] [Medline: 15683148]
18. Ohayon M, Wickwire EM, Hirshkowitz M, Albert SM, Avidan A, Daly FJ, et al. National Sleep Foundation's sleep quality recommendations: first report. Sleep Health 2017 Feb;3(1):6-19. [doi: 10.1016/j.sleh.2016.11.006] [Medline: 28346153]
19. Strine TW, Chapman DP. Associations of frequent sleep insufficiency with health-related quality of life and health behaviors. Sleep Med 2005 Jan;6(1):23-27. [doi: 10.1016/j.sleep.2004.06.003] [Medline: 15680291]
20. Lemola S, Ledermann T, Friedman EM. Variability of sleep duration is related to subjective sleep quality and subjective well-being: an actigraphy study. PLoS One 2013;8(8):e71292 [FREE Full text] [doi: 10.1371/journal.pone.0071292] [Medline: 23967186]
21. Knights J, Heidary Z, Peters-Strickland T, Ramanathan M. Evaluating digital medicine ingestion data from seriously mentally ill patients with a Bayesian Hybrid Model. NPJ Digit Med 2019;2:20 [FREE Full text] [doi: 10.1038/s41746-019-0095-z] [Medline: 31304367]
22. Open-label study to assess usability of the medical information device #1 (MIND1) system in adults with schizophrenia on oral aripiprazole. ClinicalTrials.gov. URL: https://clinicaltrials.gov/ct2/show/NCT02219009 [accessed 2021-02-11]
23. Exploratory trial to assess the functionality of an integrated call center for the digital medicine system. ClinicalTrials.gov. URL: https://clinicaltrials.gov/ct2/show/NCT02722967 [accessed 2021-02-11]
24. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Machine Learning 2011;12(85):2825-2830.

Abbreviations

DFR: deviation from rest
DMS: digital medicine system
ECG: electrocardiogram
LCRP: longest continuous rest period
LRP: longest rest period
RR: rest reference
SMI: serious mental illness

Edited by G Eysenbach; submitted 28.01.20; peer-reviewed by E Vanegas, S Patel; comments to author 31.08.20; revised version received 14.10.20; accepted 17.01.21; published 02.03.21.

Please cite as:
Heidary Z, Cochran JM, Peters-Strickland T, Knights J
A Rest Quality Metric Using a Cluster-Based Analysis of Accelerometer Data and Correlation With Digital Medicine Ingestion Data: Algorithm Development
JMIR Form Res 2021;5(3):e17993
URL: https://formative.jmir.org/2021/3/e17993
doi:10.2196/17993
PMID:33650981

Original Paper

Older Adult Peer Support Specialists' Age-Related Contributions to an Integrated Medical and Psychiatric Self-Management Intervention: Qualitative Study of Text Message Exchanges

Mbita Mbao[1*], LICSW; Caroline Collins-Pisano[2*]; Karen Fortuna[2*], LICSW, PhD

[1]Simmons University, School of Social Work, Boston, MA, United States

[2]Dartmouth College, Lebanon, NH, United States

[*]all authors contributed equally

Corresponding Author:
Karen Fortuna, LICSW, PhD
Dartmouth College
Department of Psychiatry, Geisel School of Medicine, Dartmouth College
Lebanon, NH
United States
Phone: 1 603 722 5727
Email: karen.L.Fortuna@dartmouth.edu

Abstract

Background: Middle-aged and older adults with mental health conditions have a high likelihood of experiencing comorbid physical health conditions, premature nursing home admissions, and early death compared with the general population of adults aged 50 years or above. An emerging workforce of peer support specialists aged 50 years or above or "older adult peer support specialists" is increasingly using technology to deliver peer support services to address both the mental health and physical health needs of middle-aged and older adults with a diagnosis of a serious mental illness.

Objective: This exploratory qualitative study examined older adult peer support specialists' text message exchanges with middle-aged and older adults with a diagnosis of a serious mental illness and their nonmanualized age-related contributions to a standardized integrated medical and psychiatric self-management intervention.

Methods: Older adult peer support specialists exchanged text messages with middle-aged and older adults with a diagnosis of a serious mental illness as part of a 12-week standardized integrated medical and psychiatric self-management smartphone intervention. Text message exchanges between older adult peer support specialists (n=3) and people with serious mental illnesses (n=8) were examined (mean age 68.8 years, SD 4.9 years). A total of 356 text messages were sent between older adult peer support specialists and service users with a diagnosis of a serious mental illness. Older adult peer support specialists sent text messages to older participants' smartphones between 8 AM and 10 PM on weekdays and weekends.

Results: Five themes emerged from text message exchanges related to older adult peer support specialists' age-related contributions to integrated self-management, including (1) using technology to simultaneously manage mental health and physical health issues; (2) realizing new coping skills in late life; (3) sharing roles as parents and grandparents; (4) wisdom; and (5) sharing lived experience of difficulties with normal age-related changes (emerging).

Conclusions: Older adult peer support specialists' lived experience of aging successfully with a mental health challenge may offer an age-related form of peer support that may have implications for promoting successful aging in older adults with a serious mental illness.

(JMIR Form Res 2021;5(3):e22950) doi:10.2196/22950

KEYWORDS

older adults; peer support; self-management; mobile technology

Introduction

Middle-aged and older adults with mental health conditions have a high likelihood of experiencing comorbid physical health conditions, premature nursing home admissions, and early death compared with the general population of middle-aged and older adults [1]. Despite challenges associated with mental health and comorbid physical health conditions in late life, there is a shortage of trained professionals to address the medical, psychiatric [2-4], and psychosocial age-related needs of this vulnerable population. An emerging workforce of peer support specialists aged 50 years or above is one of the fastest growing mental health workforces and may be a suitable community-based workforce to simultaneously support the mental health, physical health, and aging needs of middle-aged and older adults with a serious mental illness. A serious mental illness is defined as a diagnosable mental, behavioral, or emotional disorder that an adult has experienced in the past year that causes them serious functional impairment that substantially interferes with or limits at least one major life activity (ie, schizophrenia, bipolar disorder, and treatment refractory major depressive disorder) [5].

Older adult peer support specialists are people with a lived experience of aging into middle age and older adulthood with a mental health condition. As people with serious mental illnesses die up to 32 years earlier than the general population [1], older age is commonly defined as 50 years or above in this population of peer support specialists and their service user counterparts. For Medicaid reimbursement, older adult peer support specialists are trained and accredited by their respective state to provide support services that augment the traditional mental health system. Although accreditation and certification of older adult peer specialists varies by state, as of 2017, 41 states were billing Medicaid for peer support services [6]. Promising evidence indicates that older adult peer support specialist services reduce hospitalizations; improve engagement and treatment adherence; improve feelings of loneliness among older adults [7,8]; and promote hope, empowerment, and quality of life [9,10].

Similar to peer support specialists of any age, older adult peer support specialists employ a collaborative approach, in which caring for others creates an upward spiral of positivity to both professionals and service users [11]. This approach may be central to recovery and support bidirectional successful aging in people with serious mental illnesses and older adult peer support specialists themselves. Successful aging is defined as preventing late-life disease and disability; maintaining high cognitive, mental, and physical function; and being actively engaged in late life [12]. As such, successful aging cannot exist without the absence of disease, disability, and impairment, and thus, may not apply to middle-aged and older adults with serious mental illnesses, as the majority of this population is also diagnosed with one or more chronic health conditions [2]. As such, subjective factors of successful aging for people with serious mental illnesses may include resilience, optimism, adaptability, life satisfaction, and physical and mental health–related quality of life [12,13].

Older adult peer support specialists are increasingly using technology to deliver peer support services related to addressing *both* the mental health and physical health needs of older adults [9,10], including text messaging, videoconferencing, social media, and virtual reality. Text messages may be a low-cost high-reach intervention to support people in the community between clinical encounters with psychiatrists, social workers, etc. Clinician-supported text message exchanges have shown promising evidence of positive outcomes for improving mental health disorders and cardiac outcomes [14]. Other supportive clinician-based text message interventions have reported potential improvements in users with a comorbid diagnosis of depression and alcohol use [15]. A similar randomized controlled trial of supportive text messages for users with depression also reported positive outcomes for depression [16]. Despite evidence of clinician-based text message effectiveness for people with mental health challenges, limited literature exists on older adult peer support specialists, and unlike clinicians, their role is to intentionally disclose their lived experiences of aging with mental health challenges to support older adults. The purpose of this study was to explore older adult peer support specialists' text message exchanges with middle-aged and older adults with a diagnosis of a serious mental illness and their nonmanualized age-related contributions to a standardized integrated medical and psychiatric self-management intervention.

Methods

Study Design

The study design and recruitment procedures have been described in a previously published article [17]. Briefly, a medical and psychiatric self-management intervention enhanced with the smartphone app "PeerTECH" was offered to 10 older adults with serious mental illness (ie, bipolar disorder, schizophrenia spectrum disorder, and persistent major depressive disorder) and one medical comorbidity (ie, cardiovascular disease, obesity, diabetes, chronic obstructive pulmonary disease, hypertension, and/or high cholesterol). All the participants were over the age of 60 years, and the PeerTECH intervention was delivered in the participant's home using eModules augmented with the smartphone app (ie, it includes text messaging between older adult peer support specialists and older service users). Eight of the 10 older service users completed the intervention and were included in the current analysis. Three trained older adult peer support specialists delivered PeerTECH. All procedures were conducted in accordance with the ethical standards of the Dartmouth College Institutional Review Board and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Written informed consent was obtained from all participants.

An older adult peer support specialist and a research staff member met with potential participants in their homes. As part of the informed consent process, research staff members provided an overview of the study and discussed the voluntary nature of study participation and confidentiality issues. If interested, the potential participant completed written informed

consent and could ask any questions as needed. This study was approved by Dartmouth College Institutional Review Board. After the informed consent form was completed by the participant, an older adult peer support specialist met with the participant in his/her home for 1 hour each week over 12 weeks.

A study of PeerTECH text messaging exchanges has been conducted with the data. The related findings have been described in detail in a previously published report [17]. Previous research explored how older adult peer support specialists used text messages to support integrated illness self-management (ie, via health behavior change, self-management therapeutic techniques, engagement in health technology, and peer support) [17]. This study seeks to expand on prior work by exploring unique older adult peer support specialists' age-related contributions to a medical and psychiatric self-management intervention enhanced with a smartphone app.

Older Adult Peer Support Training

Older adult peer support specialists completed their state-certified peer specialist support training. In addition to their state credentialing, as part of this study, they were provided a 4-day training by the third author (20 hours), which included the following content: (1) information about the interconnected relationship between aging, physical health, and mental health; (2) training older adults to use technology and mitigating normal age-related challenges using technology; (3) older adult mental health (eg, suicidality in late life, anxiety and depression in late life, and role of life events in mental health in older adults); (4) techniques used in PeerTECH (ie, motivational interviewing, psychoeducation, coping skills training, and behavioral tailoring for medication adherence); (5) setting personalized goals and action steps to achieve goals; (6) delivering PeerTECH sessions in person using eModules on a tablet; (7) structure of the weekly sessions; (8) using lived experience of successful aging, physical health, and mental health challenges and role play in teaching self-management skills; and (9) orientation to the smartphone app. Each older adult peer support specialist had a two to three-person caseload and worked a total of 10 hours per week, including direct care, text messaging participants, and supervision. Older adult peer support specialists were supervised by a trained older adult peer support supervisor for 1 hour each week in person or over the telephone. Verbal informed consent was obtained from older adult peer support specialists during the older adult peer support specialist training.

Text Messaging Requirements

The smartphone intervention portion of PeerTECH was designed to reinforce in-person sessions and to provide support to people in real-world environments. Older adult peer support specialists were instructed to message participants a minimum of three times a week. There was no maximum number of text messages required. Text message exchanges were unstructured and were to focus on nonmanualized peer support, follow-up of service users' goals, and discussions during in-person sessions facilitated by eModules (eg, "hope you are doing well on your goals- journaling and walking"). Older adult peer support specialists were encouraged to text message at times consistent with the preferences of the service users they were working with. All text message content was logged, and the time/date was recorded. Requirements for text messaging were purposely left unstructured in an effort to examine naturalistic interactions between older adult peer support specialists and service users with a diagnosis of a serious mental illness.

Smartphones and data plans were provided free of charge or service user participants could use their own smartphone. Participants were not provided incentives to send text messages; however, they were provided US $20 compensation to complete baseline, 1-month, and 3-month assessments (total of US $60 over the entire study duration). For this study, incoming and outgoing text messages were securely stored within the smartphone app database. Text message transcript data were extracted into an Excel worksheet and analyzed.

Data Analysis

Transcripts were analyzed for eight participants and three older adult peer support specialists. The codebook consisted of a priori older adult peer support specialist and nonpeer researcher–driven codes, which were derived from text messages, and inductively derived codes from qualitative data [18]. The first and third authors read data and incorporated new codes and operational definitions from transcript coding, which is a validated approach that allows for multiple perspectives [18]. Codes were assigned to text, grouped, and checked for themes. Thematic analysis was used to summarize themes identified in the text message data [19]. Analyses assessed within-group consensus or disagreement. Member checking was employed to validate results and resolve any incongruent findings. As such, the third author contacted the participating older adult peer support specialists to discuss the key themes that emerged from the text message data. This approach helped ensure these findings are consistent with how older adult peer support specialists intend to use these text message exchanges. Quantitative data comprised the frequency of text messages by either older adult peer support specialists or service users. Frequency data were captured directly from the PeerTECH app. Frequency data were integrated at the conclusion of the study.

Results

Study Sample

The sample consisted of eight service user participants and three older adult peer support specialists. Service user participants had a mean age of 68.8 years (SD 4.9 years; range 62-77 years) and were primarily women (7/8, 88%), White (8/8, 100%), and married (6/8, 75%). The sample included people diagnosed with major depressive disorder (5/8, 63%), schizophrenia spectrum disorder (2/8, 25%), and bipolar disorder (1/8, 13%). Older adult peer support specialists were all aged 55 years or above. Additionally, 100% (3/3) were female, 66% (2/3) identified as White, and 33% (1/3) identified as African American.

Text Message Exchanges

Older adult peer support specialists sent text messages to participants' smartphones based on participants' preferences from 8 AM to 10 PM EST on weekdays and weekends. Over the course of the 12-week intervention, a total of 356 text messages were sent. For this study, only the text message exchanges were analyzed. Five themes emerged including (1)

simultaneously managing mental health and physical health issues through empowerment and technology; (2) realizing new coping skills in late life; (3) sharing roles as parents and grandparents; (4) wisdom; and (5) sharing lived experience of difficulties with normal age-related changes.

Using Technology to Simultaneously Manage Mental Health and Physical Health

The first and most predominant theme was using technology to simultaneously manage mental health and physical health (29/53, 55%). Older adult peer support specialists encouraged participants to take control of their mental health and physical health needs using technology. For example, an older adult peer support specialist texted as follows:

> *You were able to relax enough to get the exercise. It was amazing what 10 minutes of focused quiet can do. Take a timer and just do a mental relaxation exercise for 5 minutes 2x a day. It will be great for your heart. You can find guided imagery and meditations videos on YouTube also.*

Another older adult peer support specialist texted as follows:

> *I can feel the stress you are under. In the app, there are several short videos on stress reduction like deep breathing, mindfulness, meditation that you can try to have some peace.*

Another older adult peer support specialist provided tools for reducing anxiety as follows:

> *Doing slow deep breathing and short meditations can relax you and dissipate anxiety more effectively than cigarettes. Think about it. I will look for some YouTube videos on relaxation and/or smoking cessation.*

Realizing New Coping Skills in Late Life

The second most predominant theme was realizing new coping skills in late life (8/53, 15%). Older adult peer support specialists purported that they are learning new skills in late life related to coping skill development. For example, an older adult peer support specialist texted as follows:

> *Since I have been in recovery, I am learning new skills to cope with stressful days, weeks, and months.*

Another older adult peer support specialist texted as follows:

> *I do a lot of journaling (writing). It helps me in many ways- when I am upset; when I don't have someone to talk to and I need to get it out of my head; when I have to make a big decision...Writing has been a blessing for me especially in tough times.*

Sharing Roles as Parents and Grandparents

The third most predominant theme was sharing roles as parents and grandparents in late life (7/53, 13%). For example, an older adult peer support specialist texted as follows:

> *Being (Grandma...) means so much to me. I know you love your grandchildren and in time we can work through all of this with grace and ease.*

Another older adult peer support specialist texted as follows:

> *We as parents and women tend to put everybody first. Now it is our time to take care of ourselves.*

Older adult peer support specialists used their shared experiences to help participants solve the challenges they were having with their families, which were impacting service users' mental health. Older adult peer support specialists and older adults with mental health conditions shared similar experiences, for example, an older adult peer support specialist texted as follows:

> *We grandmothers can't hold back our sheer joy and love we have for our grandchildren. I wish I had a video of your "grandmother moment." Your smile lit up the room as you shared about your excitement from hearing from your granddaughter (that's is tears of joy smile).*

Wisdom

The fourth theme was wisdom (6/53, 11%). Wisdom, for the purpose of this manuscript, was defined as advanced-level knowledge that leads to good judgment [20]. Older adult peer support specialists and older adults with mental health conditions both offered each other wisdom regarding aging successfully. For example, one older adult peer support specialist texted as follows:

> *We form habits without knowing it and once we identify the "bad" habits, we can turn those unwanted habits around in 14-21 days. We just have to be persistent.*

Another older adult peer support specialist texted as follows:

> *What you resist, persist what we focus on, we also get... as a man thinketh, so is he... What the mind can conceive and believe, it can achieve... In other words, instead of focusing on what you don't want, put your focus on what you do want. … What can you do to transform your thinking? What actions can you take to feel better today? What do you want to be focused on? WORDS HAVE POWER and so do we. We have a saying "Be, Do, Have" ex: Be Happy, Do things that make you feel happy and you will have Happiness in your life.*

Sharing Lived Experience of Difficulties With Normal Age-Related Changes

The fifth emerging theme was sharing lived experience of difficulties with normal aging (3/53, 6%). An older adult peer support specialist texted as follows:

> *Sleep is a challenge for me too. It is a process; we will work on it. I added the Sleep module. I hope you can get some good tips on getting a good night sleep. I get home so late at night it's hard for me to wind down. Let me know what works best for you, so I can try it too.*

Discussion

Principal Findings

This exploratory qualitative study examined older adult peer support specialists' text message exchanges with middle-aged and older adults with a diagnosis of a serious mental illness and their nonmanualized age-related contributions to a standardized integrated medical and psychiatric self-management intervention. The following themes emerged: (1) simultaneously managing mental health and physical health issues through empowerment and technology; (2) realizing new coping skills in late life; (3) sharing roles as parents and grandparents; (4) wisdom; and (5) sharing lived experience of difficulties with normal age-related changes. Older adult peer support specialists' lived experience of aging successfully with a mental health challenge may offer an age-related form of peer support that may have implications for promoting successful aging in older adults with serious mental illnesses.

The age of older adult peer support specialists and their experiences with multimorbidity led to text message exchanges that focused on the management of mental health and physical health challenges using technologies outside of the PeerTECH platform (eg, smartphone apps, YouTube, and social media platforms). Older adults in the general population with multimorbidity are often faced with burdens of managing treatment, such as increased health care visits, refilling prescriptions, managing diet concerns, and self-managing care [14], and this is compounded by functional and structural challenges (eg, lack of transportation) associated with serious mental illnesses in late life [2]. As increasing numbers of older adults with serious mental illnesses are using technologies [21], technologies may be viable tools to support late-life self-management of mental health and physical health challenges. Further, older peer support specialists' personal experience of coping with multimorbidity using technologies may influence service users to engage with these new technologies to support their own medical and psychiatric self-management skill development. Prior research has found that clinician [22] and peer support interactions [23] within digital mental health services facilitate engagement with technologies. As such, older adult peer support specialists' shared experience may have a role in influencing engagement in digital interventions among older adults with serious mental illnesses.

The shared lived experience of parenting and grandparenting between older adult peer support specialists and service users may have facilitated the development of a supportive alliance. Older adult peer support specialists' practice principles, unlike those of clinicians, encourage sharing of their lived experiences (or self-disclose) to support the recovery of individuals. The use of intentional self-disclosure may have facilitated the development of a supported alliance between older adult peer support specialists and older adult service users.

Through sharing of experiences, older adult peer support specialists offered wisdom related to navigating some of the challenges resulting from age, illness, and life experiences. Wisdom is a prototype of successful aging [24] and has been found to enhance mental health and promote well-being in older adults (without serious mental illnesses) [25]. Possibly, through bidirectional sharing of knowledge of aging, illness, and life experiences, both older adult peer support specialists and older service users may assist one another in navigating some of the challenges of aging successfully with a serious mental illness, which could support older adults between clinical encounters. This is also an important area for future inquiry. As peer support is a nonmanualized form of support [26], peer support specialists aged 50 years or above may offer different lived experience expertise than their younger adult peer support specialist counterparts, based on their level of expert knowledge (ie, wisdom).

This study is not without limitations. First, it is not known whether we met saturation. Qualitative interviews are conducted until there is saturation of data (ie, saturation means that researchers reach a point in their analysis where sampling more data will not lead to more information related to their research questions) [27]. By drawing from grounded theory design [18], saturation would generally occur with 20 to 30 participants in total [28]; however, the sample size was small because the primary study was conducted to assess feasibility [9]. It is important to note that findings cannot be generalized; however, the themes identified can be used to guide the development of peer support text-messaging services as an adjunct to evidence-based interventions [9]. Further, we were unable to stratify our data by demographic characteristics owing to the sample size. For example, one peer had a master's degree in social work, and this advanced educational background likely influenced the person's delivery of services. Second, peers met in person with participants over a 12-week time frame, and in-person follow-up discussions from text messages are not reported. Third, the sample involved a heterogeneous group of people with psychotic disorders and mood disorders that predominately included those with major depressive disorder. Fourth, the participants in this study were all receiving mental health services, and therefore, our findings cannot be generalized to individuals with serious mental illnesses not enrolled in care or without access to mental health services. Finally, the results elucidate text message themes between older adult peer support specialists and older adults with serious mental illnesses and chronic health conditions; however, it is not known whether the peer-to-participant text message exchanges can improve self-management and other clinical outcomes.

Conclusion

Older adult peer support specialists are an emerging part of the service delivery system for older adults. Older adult peer support specialists offered text message–based age-related experiential contributions to support aging successfully with a mental health and physical health condition. Through older adult peer support specialists' wisdom, sharing of new late-life coping skills and similar age-specific roles in life and encouragement to use technology to support medical and psychiatric self-management may promote engagement in nontraditional support services (ie, YouTube) and support for older adults with serious mental illnesses in the community between clinical encounters.

Acknowledgments

KF was funded by a K01 award from the National Institute of Mental Health (K01MH117496)

Conflicts of Interest

KF provides consulting services through Social Wellness and discloses interest with Trust and InquisitHealth. The other authors have no conflicts to disclose.

References

1. Walker E, McGee RE, Druss BG. Mortality in mental disorders and global disease burden implications: a systematic review and meta-analysis. JAMA Psychiatry 2015 Apr;72(4):334-341 [FREE Full text] [doi: 10.1001/jamapsychiatry.2014.2502] [Medline: 25671328]
2. Bartels S, DiMilia PR, Fortuna KL, Naslund JA. Integrated Care for Older Adults with Serious Mental Illness and Medical Comorbidity: Evidence-Based Models and Future Research Directions. Psychiatr Clin North Am 2018 Mar;41(1):153-164 [FREE Full text] [doi: 10.1016/j.psc.2017.10.012] [Medline: 29412843]
3. Everett A. Bringing Awareness to the Mental Health of Older Adults. SAMHSA. 2019. URL: https://blog.samhsa.gov/2019/05/20/bringing-awareness-to-the-mental-health-of-older-adults [accessed 2020-09-01]
4. Eden J, Maslow K, Le M, Blazer D. The Mental Health and Substance Use Workforce for Older Adults. In Whose Hands? Assessing the Service Needs of Older Adults with Mental Health and Substance Use Conditions. The Mental Health and Substance Use Workforce for Older Adults. In Whose Hands? Assessing the Service Needs of Older Adults with Mental Health and Substance Use Conditions 2012. [doi: 10.17226/13400] [Medline: 24851291]
5. Adults With SMI and Children/Youth With SED. SAMHSA. 2020. URL: https://www.samhsa.gov/dbhis-collections/smi [accessed 2020-10-10]
6. Videka L, Neale J, Page C, Buche J, Beck AJ, Wayment C, et al. National analysis of peer support providers: Practice settings, requirements, roles, and reimbursement. Behavioral Health Workforce. 2019. URL: https://behavioralhealthworkforce.org/wp-content/uploads/2019/10/BHWRC-Peer-Workforce-Full-Report.pdf [accessed 2020-08-16]
7. Conner K, Cadet T, Brown M, Barnett J. The Impact of Peer Support on the Risk of Future Hospital Readmissions among Older Adults with a Medical Illness and Co-Occurring Depression. Social Sciences 2018 Sep 12;7(9):156 [FREE Full text] [doi: 10.3390/socsci7090156]
8. Gagne C, Finch WL, Myrick KJ, Davis LM. Peer Workers in the Behavioral and Integrated Health Workforce: Opportunities and Future Directions. Am J Prev Med 2018 Jun;54(6 Suppl 3):S258-S266 [FREE Full text] [doi: 10.1016/j.amepre.2018.03.010] [Medline: 29779550]
9. Fortuna K, Lohman MC, Gill LE, Bruce ML, Bartels SJ. Adapting a Psychosocial Intervention for Smartphone Delivery to Middle-Aged and Older Adults with Serious Mental Illness. Am J Geriatr Psychiatry 2017 Aug;25(8):819-828 [FREE Full text] [doi: 10.1016/j.jagp.2016.12.007] [Medline: 28169129]
10. Fortuna KL, Naslund JA, LaCroix JM, Bianco CL, Brooks JM, Zisman-Ilani Y, et al. Digital Peer Support Mental Health Interventions for People With a Lived Experience of a Serious Mental Illness: Systematic Review. JMIR Ment Health 2020 Apr 03;7(4):e16460 [FREE Full text] [doi: 10.2196/16460] [Medline: 32243256]
11. Feeney B, Collins NL. A new look at social support: a theoretical perspective on thriving through relationships. Pers Soc Psychol Rev 2015 May;19(2):113-147 [FREE Full text] [doi: 10.1177/1088868314544222] [Medline: 25125368]
12. Jeste DV, Ardelt M, Blazer D, Kraemer HC, Vaillant G, Meeks TW. Expert consensus on characteristics of wisdom: a Delphi method study. Gerontologist 2010 Oct 15;50(5):668-680 [FREE Full text] [doi: 10.1093/geront/gnq022] [Medline: 20233730]
13. Ibrahim F, Cohen CI, Ramirez PM. Successful Aging in Older Adults With Schizophrenia: Prevalence and Associated Factors. The American Journal of Geriatric Psychiatry 2010 Oct;18(10):879-886 [FREE Full text] [doi: 10.1097/jgp.0b013e3181d57441]
14. Hoffmann T, Jansen J, Glasziou P. The importance and challenges of shared decision making in older people with multimorbidity. PLoS Med 2018 Mar;15(3):e1002530 [FREE Full text] [doi: 10.1371/journal.pmed.1002530] [Medline: 29534067]
15. Agyapong V, Ahern S, McLoughlin DM, Farren CK. Supportive text messaging for depression and comorbid alcohol use disorder: single-blind randomised trial. J Affect Disord 2012 Dec 10;141(2-3):168-176 [FREE Full text] [doi: 10.1016/j.jad.2012.02.040] [Medline: 22464008]
16. Agyapong V, Juhás M, Ohinmaa A, Omeje J, Mrklas K, Suen VYM, et al. Randomized controlled pilot trial of supportive text messages for patients with depression. BMC Psychiatry 2017 Aug 02;17(1):286 [FREE Full text] [doi: 10.1186/s12888-017-1448-2] [Medline: 28768493]

17. Fortuna K, Naslund JA, Aschbrenner KA, Lohman MC, Storm M, Batsis JA, et al. Text message exchanges between older adults with serious mental illness and older certified peer specialists in a smartphone-supported self-management intervention. Psychiatr Rehabil J 2019 Mar;42(1):57-63 [FREE Full text] [doi: 10.1037/prj0000305] [Medline: 30010355]
18. Martin P, Turner BA. Grounded Theory and Organizational Research. The Journal of Applied Behavioral Science 2016 Jul 26;22(2):141-157 [FREE Full text] [doi: 10.1177/002188638602200207]
19. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
20. Baltes PB, Kunzmann U. Wisdom. The Psychologist. 2003. URL: http://www.margret-baltes-stiftung.de/PBB-Website/Baltes&Kunzmann.pdf [accessed 2020-07-28]
21. Ben-Zeev D, Davis KE, Kaiser S, Krzsos I, Drake RE. Mobile technologies among people with serious mental illness: opportunities for future services. Adm Policy Ment Health 2013 Jul;40(4):340-343 [FREE Full text] [doi: 10.1007/s10488-012-0424-x] [Medline: 22648635]
22. Mohr D, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011 Mar 10;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
23. Fortuna KL, Brooks JM, Umucu E, Walker R, Chow PI. Peer Support: a Human Factor to Enhance Engagement in Digital Health Behavior Change Interventions. J. technol. behav. sci 2019 May 29;4(2):152-161. [doi: 10.1007/s41347-019-00105-x]
24. Parisi J, Rebok GW, Carlson MC, Fried LP, Seeman TE, Tan EJ, et al. Can the Wisdom of Aging be Activated and Make a Difference Societally? Educ Gerontol 2009;35(10):867-879 [FREE Full text] [doi: 10.1080/03601270902782453] [Medline: 21998491]
25. Jeste D, Lee EE. The Emerging Empirical Science of Wisdom: Definition, Measurement, Neurobiology, Longevity, and Interventions. Harv Rev Psychiatry 2019;27(3):127-140 [FREE Full text] [doi: 10.1097/hrp.0000000000000205]
26. Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. Psychiatr Rehabil J 2004;27(4):392-401. [doi: 10.2975/27.2004.392.401] [Medline: 15222150]
27. Seale C, Gobo G, Gubrium JF, Silverman D. Qualitative Research Practice. Thousand Oaks, CA, USA: SAGE Publications; 2004.
28. Creswell JW. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Thousand Oaks, CA, USA: SAGE Publications; 2007.

Edited by G Eysenbach; submitted 27.07.20; peer-reviewed by J Urff, M Wasilewski; comments to author 17.11.20; revised version received 26.12.20; accepted 17.01.21; published 02.03.21.

Please cite as:
Mbao M, Collins-Pisano C, Fortuna K
Older Adult Peer Support Specialists' Age-Related Contributions to an Integrated Medical and Psychiatric Self-Management Intervention: Qualitative Study of Text Message Exchanges
JMIR Form Res 2021;5(3):e22950
URL: https://formative.jmir.org/2021/3/e22950
doi:10.2196/22950
PMID:33650979

Original Paper

Feasibility and Acceptability of a Digital Intervention to Support Shared Decision-making in Children's and Young People's Mental Health: Mixed Methods Pilot Randomized Controlled Trial

Shaun Liverpool[1,2], PhD; Julian Edbrooke-Childs[1], PhD

[1]Evidence-Based Practice Unit, Anna Freud National Centre for Children and Families, University College London, London, United Kingdom
[2]Faculty of Health, Social Care & Medicine, Edge Hill University, Ormskirk, United Kingdom

Corresponding Author:
Shaun Liverpool, PhD
Evidence-Based Practice Unit
Anna Freud National Centre for Children and Families
University College London
4-8 Rodney Street
The Kantor Centre of Excellence
London, N1 9JH
United Kingdom
Phone: 44 7539468630
Email: shaun.liverpool.14@ucl.ac.uk

Abstract

Background: Interventions to involve parents in decisions regarding children's and young people's mental health are associated with positive outcomes. However, appropriately planning effectiveness studies is critical to ensure that meaningful evidence is collected. It is important to conduct pilot studies to evaluate the feasibility and acceptability of the intervention itself and the feasibility of the protocol to test effectiveness.

Objective: This paper reports the findings from a feasibility and acceptability study of Power Up for Parents, an intervention to promote shared decision-making (SDM) and support parents and caregivers making decisions regarding children's and young people's mental health.

Methods: A mixed method study design was adopted. In stage 1, health care professionals and parents provided feedback on acceptability, usefulness, and suggestions for further development. Stage 2 was a multicenter, 3-arm, individual, and cluster randomized controlled pilot feasibility trial with parents accessing services related to children's and young people's mental health. Outcome measures collected data on demographics, participation rates, SDM, satisfaction, and parents' anxiety. Qualitative data were analyzed using thematic analysis. Google Analytics estimates were used to report engagement with the prototype. Outcomes from both stages were tested against a published set of criteria for proceeding to a randomized controlled trial.

Results: Despite evidence suggesting the acceptability of Power Up for Parents, the findings suggest that recruitment modifications are needed to enhance the feasibility of collecting follow-up data before scaling up to a fully powered randomized controlled trial. On the basis of the Go or No-Go criteria, only 50% (6/12) of the sites successfully recruited participants, and only 38% (16/42) of parents completed follow-up measures. Nonetheless, health care practitioners and parents generally accessed and used the intervention. Themes describing *appearance and functionality, perceived need and general helpfulness, accessibility and appropriateness*, and *a wish list for improvement* emerged, providing valuable information to inform future development and refinement of the intervention.

Conclusions: Owing to the high attrition observed in the trial, proceeding directly to a full randomized controlled trial may not be feasible with this recruitment strategy. Nonetheless, with some minor adjustments and upgrades to the intervention, this pilot study provides a platform for future evaluations of Power Up for Parents.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN) 39238984; http://www.isrctn.com/ISRCTN39238984.

International Registered Report Identifier (IRRID): RR2-10.2196/14571

(JMIR Form Res 2021;5(3):e25235) doi:10.2196/25235

KEYWORDS

mental health; pilot projects; child; adolescent; parents; shared decision making

Introduction

Background

Shared decision-making (SDM) is an ethical imperative whereby health care professionals collaborate with service users to derive care and treatment decisions. The process involves an exploration of health care options, service users' values and preferences, and achieving treatment consensus [1]. There is a wealth of knowledge suggesting that adopting SDM practices is associated with better outcomes across health care settings [2,3]. However, implementing SDM in children's and young people's mental health (CYPMH) services remains a challenge. Barriers to implementation include professional, relational, service user or parent, and service-level and context-level factors [4]. Researchers agree that a primary reason could be the unique triad relationship involving multiple decision makers [5,6]. As a result, parents may navigate between feeling excluded from services, advocating or assuming the role of surrogate decision makers depending on the age and capacity of the child or young person [7,8]. Such feelings and roles sometimes result in added stressors for the parents involved [7]; low service engagement [9,10]; and treatment disagreements between parents, health care practitioners, and young people [11].

Parents and carers are recognized by the literature and by the law as key members of the CYPMH decision-making process [12,13], reporting significant benefits to involvement [14]. However, research in general pediatric care highlights parents' emotional states as a commonly reported barrier to adopting SDM measures [15]. Despite a range of interventions and service delivery models to support parents' involvement, time, accessibility, and appropriateness of the interventions also appear to influence use and successful implementation [16]. That study also revealed that research exploring available interventions is limited in that it targets specific populations (eg, attention-deficit and hyperactivity disorders or autism spectrum disorders), uses less innovative modalities (eg, face-to-face, paper-based, or static digital tools), or evaluates interventions using nonrandomized study designs (eg, pre or post, qualitative, or pilot trials). In addition, there appears to be a large number of interventions being developed and implemented without being tested for effectiveness or using small or unrepresentative samples [16].

In response to this, an evidence-based, theoretically informed interactive web-based app was designed and developed to support SDM in universal CYPMH services [17]. The core content and aims of the intervention are based on an earlier intervention named Power Up, an intervention designed to promote SDM among young people accessing mental health services [18,19]. This intervention builds on previous versions by using an affective-appraisal approach that takes into account the emotional states of parents and caregivers.

With the growing interest in applying digital technology in CYPMH services, several modes of delivery, including website and text messages, have been adopted [20]. In line with the Technology Acceptance Model (TAM), it is recommended to test the acceptance of new digital interventions to ensure successful implementation [21]. The TAM is an extension of the Theory of Reasoned Action that focuses on behavioral intention and attitude. The theory proposes that the assessment of perceived usefulness and perceived ease of use can determine whether users engage with the new digital intervention. However, researchers and clinicians agree that poorly designed studies to evaluate these interventions can result in false-positive findings and loss of research investments [22,23].

The evidence-based approach to evaluating effectiveness recognizes randomized controlled trials (RCTs) as the *gold standard* for generating the highest level of evidence. RCTs are viewed as the most rigorous when it comes to determining cause-effect relationships between treatment and outcomes [24]. However, to ensure a successful RCT, it is highly recommended that pilot and feasibility studies are first conducted [25,26]. The Medical Research Council's guidelines highlight that assessing the feasibility allows researchers to examine important components of the research, such as testing the procedures, estimating rates of recruitment and retention of participants, and determining the sample sizes for future trials [27]. Therefore, acknowledging the relevance of pilot and feasibility studies and in keeping with the Medical Research Council's framework for developing, evaluating, and implementing a complex intervention, this study was considered an important step.

Aims and Research Questions

This pilot feasibility study focused on obtaining end users' views of the intervention and exploring justifiable administration procedures to inform a full RCT. The primary aim is to investigate whether it is feasible to conduct a prospective RCT of an evidence- and web-based app (Power Up for Parents [PUfP]) to promote SDM in families accessing CYPMH services. In addition, this study assessed the perceived usefulness and acceptance of the intervention to determine whether end users would engage.

The following research questions (RQs) were addressed:

Quantitative RQs:

- RQ1: What are the eligibility, consenting, adherence, and engagement rates of participants using PUfP?
- RQ2: Are the outcome measures appropriate and acceptable for a prospective RCT?
- RQ3: What are the potential barriers and enablers to conducting a prospective RCT?
- RQ4: Which data collection procedures are appropriate and acceptable?

- RQ5: What is the scope of the pilot data collected from users and nonusers of PUfP?

Qualitative RQs:

- RQ6: Is PUfP acceptable and useful for parents and health care practitioners?
- RQ7: Can the feedback from PUfP users be used to further refine the prototype for prospective RCTs?

Methods

Changes to Protocol

During the initial stages of the study, it was discovered that the intervention may be applicable to settings beyond the CYPMH services offered by the United Kingdom's National Health Service (NHS). Parent experts in patient and public involvement (PPI) sessions confirmed this by expressing that the intervention was something they could use with limited guidance. In addition, typical service users accessing CYPMH support via the NHS were below the age of 18 years. In line with the United Nation's definition of young people [28], this research interest extended to parents of young people up to the age of 24 years. Therefore, to obtain more feedback and usage data during the feasibility and pilot testing phases, we added a second recruitment strand (community sampling). It also became clear at later stages of the study that recruitment from NHS services was slower than anticipated, and therefore, the second recruitment strand assisted in increasing the study's sample size. This change also strengthened the study by allowing further exploration of different recruitment strategies to partly address the aims of the feasibility study.

Study Design

A mixed methods study involving qualitative data collection and feasibility testing was adopted. Interviews and focus group discussions (FGDs) required user testing of the intervention by health care practitioners and parents to obtain feedback on acceptability and usefulness, and suggestions for further development and upgrading of the prototype. The second stage of the study included (1) a multicenter, 3-arm, randomized controlled, pilot feasibility trial with parents accessing NHS CYPMH services to explore efficiency and eliminate possible study contamination and (2) a web-based individually randomized trial with a community sample of parents to inform recruitment strategies for future trials.

Study Setting

A total of 18 NHS Trusts in England offering CYPMH services were identified as potential study sites. Community samples were recruited on the web through social media advertising. Parents in the community sample accessed the study via a link to the recruitment software Gorilla [29].

Intervention: PUfP

The development and evidence base for the PUfP prototype are described and outlined in more detail in the study protocol [17]. PUfP is a decision support intervention with 5 key features (ie, decisions, goals, journey, support, and resources). The intervention aims to encourage discussion, allow parents to ask questions during sessions or seek further information between sessions, and allow health care practitioners to tailor the SDM process to accommodate the needs of the parent and child or young person.

Participants

Health Care Practitioners

A contact person (site collaborator) circulated information about the study to all health care practitioners at the CYPMH services. Then, the health care practitioners attended an information session where a brief introduction and further details of the study were provided. Any health care practitioner who identified as being in contact with the families accessing care when making care and treatment decisions was eligible to participate in the study.

Parents

On the basis of the eligibility criteria, the health care practitioners identified the eligible participants. Posters and flyers were posted at the participating NHS sites. To obtain a community sample, the study was advertised on the Anna Freud National Centre for Children and Families' website between June and August 2019 and promoted through social media platforms (ie, Facebook and Twitter). In addition, a blog post was written on the Association of Child and Adolescent Mental Health's website to further advertise the study [30]. The recruitment process was guided and informed by the PPI activities and the study's steering committee.

All parents were screened against the eligibility criteria developed before the start of the trial. Parents were included on the following criteria:

1. Over the age of 18 years
2. No known mental health diagnosis
3. Ability to speak and understand English
4. Parent of at least one child or young person attending CYPMH services.

Parents were excluded on the following criteria:

1. Concurrent and/or involvement in other research that was likely to interfere with the intervention
2. Parents or guardians in cases where the child or young person was being treated under a section of the Mental Health Act.

Procedure and Materials

Qualitative Data Collection

Semistructured interviews and FGDs were conducted. The interview guide aimed to capture the perceived usefulness and acceptability of the intervention, including suggestions for content and prototype upgrade. At the end of the interview sessions, participants were debriefed and encouraged to contact researchers with any further concerns or suggestions.

Quantitative Data Collection

Study sites were randomly assigned to either the control group or one of the 2 intervention groups. Intervention group 1 (IG1) received the prospective version 1 of PUfP which included

decisions, *goals*, *journey*, *support*, and *resources* features. Intervention group 2 (IG2) received version 2 of PUfP without the *support* and *resources* features. The control group included participants who were not exposed to either version of the intervention. The cluster randomization for the NHS sample was completed independent of the research team. For the community sample, participant-level randomization was conducted using Gorilla recruitment software. Therefore, any parent accessing any form of CYPMH service (eg, school mental health support or private therapeutic services) coming in contact with the study information had a chance to participate.

Participants met with a researcher at a convenient time and completed a battery of baseline and follow-up questionnaires. These consisted of demographic details (gender, ethnicity, first language, relationship to child, and child's age), participation rates (completion of consent, pretest and posttest measures, and intervention use), SDM measures (the Control Preferences Scale for Pediatrics [31], the Pediatric Shared Decision-Making Questionnaire (modified) [32], and the Decisional Conflict Scale [DCS] [33]), experience of service (the Experience of Service Questionnaire [ESQ] [34]), usability and acceptance (the Poststudy Usability Questionnaire [35]), and an anxiety measure (the Spielberger State Anxiety Inventory Form for Adults [36]). Further details on the outcome measures are presented in the study protocol [17]. Depending on which group the participants were recruited into (ie, IG1, IG2, or control), they received help to access the app and were given a guided tour. Participants were then encouraged to use the app as much as they needed to. Participants completed follow-up measures at 3 months after or at dropout or discharge (whichever came first).

The health care practitioners completed an adapted version of the Control Preferences Scale to highlight observed changes in the amount of parental involvement in the child's care and treatment decisions. Clinicians were asked to select 1 of 5 statements on whether "the parent left all mental health care and treatment decisions about the child to the practitioner" or "the parent shared responsibility for care and treatment decisions with the practitioner."

At the end of the pilot testing phase, participants were encouraged to share their opinions on the study before being debriefed and thanked for their participation. The Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) diagram [37] as reported in the study protocol illustrates the participants' pathway through the trial.

Data Management and Analysis

Qualitative Data

All interviews and FGDs were audiorecorded and transcribed. The data were analyzed using thematic analysis [38]. Data were coded using a combination of a priori themes as categories and emergent themes [39]. The first step generated initial codes from open coding, in which units of meanings were derived from line-by-line analysis followed by axial coding (ie, locating linkages between data) to integrate and differentiate among subcategories [39]. A priori themes or categories were important in framing the emergent themes and assisted in reporting the findings. NVivo was used as the qualitative data management software [40]. An independent investigator reviewed 3 random transcripts and generated codes. The codes were compared and discussed to reach a consensus before inclusion.

Quantitative Data

Descriptive statistics were calculated for participant characteristics at baseline, and Google Analytics estimates [41] were used to report engagement with the prototype. To address the aims of the feasibility study, the main focus was on descriptive data. However, some exploratory significance testing using means and CIs were conducted on within- and between-group mean differences at the 2 time points (ie, baseline and follow-up) on the SDM measure using the *as-per-protocol* approach. The intraclass correlation was also calculated to prepare information for sample size calculation within a clustered randomized trial. Analyses were conducted using the Statistical Package for Social Sciences software [42]. Outcomes from both study designs were then tested against 8 Go or No-Go criteria (Multimedia Appendix 1). The criteria were informed by the key areas of focus for evaluating a feasibility study [26]. Upon completion of the study, the following decisions were possible:

- Ready to proceed to full RCT
- Ready to proceed with some action to be taken
- Not ready to proceed to full RCT.

Recording Adverse Events

Adverse events were identified as any untoward medical occurrence in a parent, child or young person, or HCP, which did not necessarily have a causal relationship with the intervention. Any adverse events arising during the study period were assessed for severity, causality, seriousness, and expectedness (ie, relating to the information provided by PUfP).

Ethical Approvals and Research Governance

The study was ethically reviewed by the London Surrey Research Ethics Committee (REC) and approved by the Health Research Authority (IRAS 236277) for recruitment at the CYPMH services provided by the NHS. Recruitment for the community sample was approved by the University College London REC.

Results

Overview

Recruitment was expected to begin in October 2018 and was scheduled to last for 1 year. However, NHS REC approval was received in December 2018; therefore, recruitment from NHS began in January 2019. We approached 18 NHS sites, of which 12 (67%) CYPMH sites expressed interest and were recruited to participate in both stages of the study. Web-based community advertising resulted in 387 unique visitors on the study webpage. The data collection was terminated on October 1, 2019. The results section is structured according to the study's RQs.

RQ1: What Are the Eligibility, Consenting, Adherence, and Engagement Rates of Participants in the Trial?

Through consultation with site collaborators and HCPs, the eligibility criteria were considered to be clear and

straightforward. However, one site expressed difficulties in recruiting parents due to the high percentage of parents at that site with a mental health diagnosis, which met the exclusion criteria. Consequently, this site withdrew from the study within 3 months of confirming its capacity and capability.

Qualitative Study

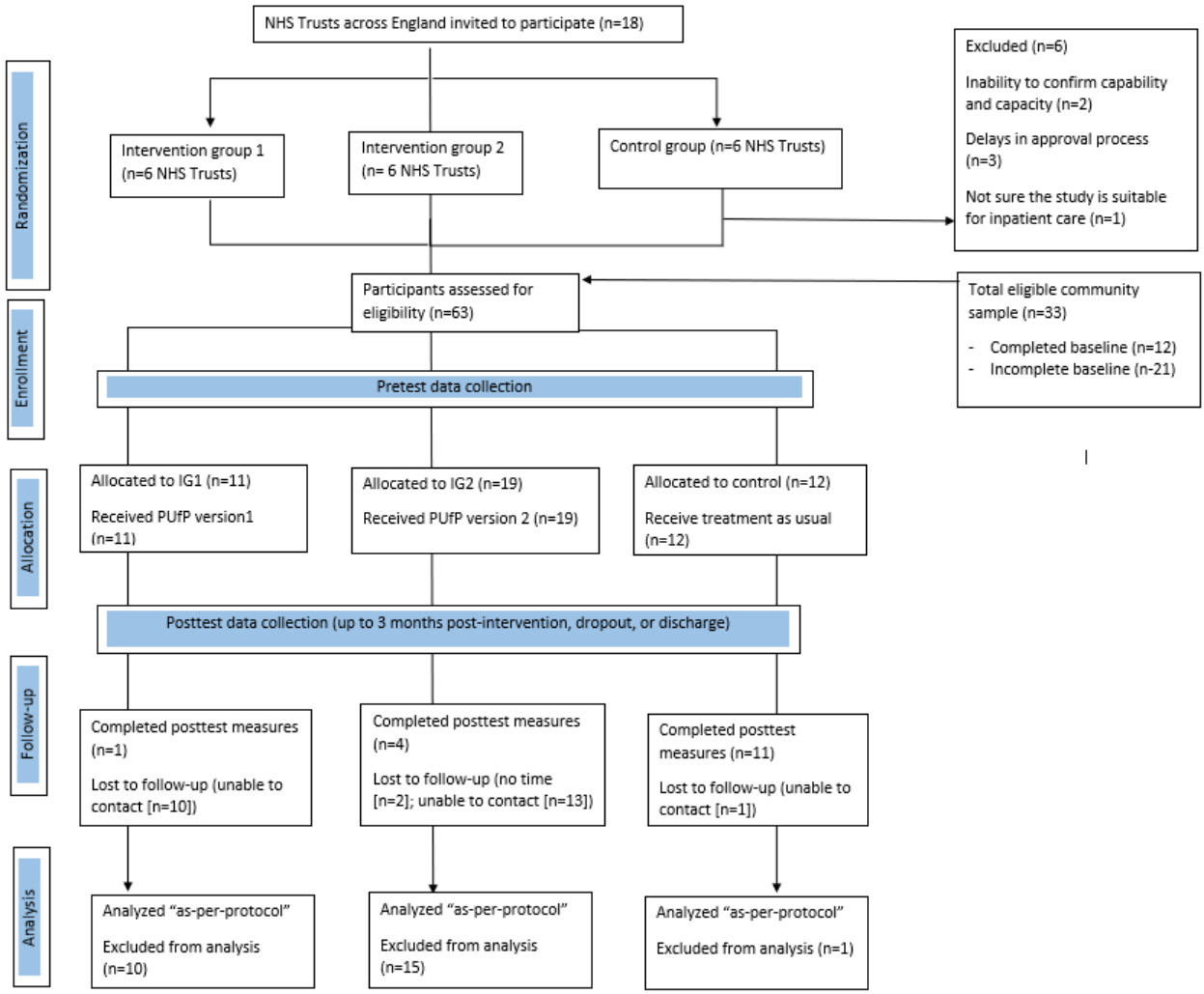
A total of 40 parents consented to participate in interviews or FGDs (ie, 36 from the NHS and 4 from community recruitment). In total, 36 parents from the NHS were recruited from 58% (7/12) of the participating sites. The remaining 5 sites not recruiting participants included the site that withdrew from the study, 1 site that wished to take part in the quantitative study only, and 3 sites that stated that the parents were too busy to commit to an interview or FGD. Consequently, a total of 24 parents participated (24/40, 60%): 14 parents were interviewed, and 10 participated in FGDs. For the remaining participants who consented but did not attend an FGD or interview, it was not possible to contact them on the email or phone contact provided by the site collaborator or to arrange a convenient time for an interview. The sample included 22 mothers and 2 fathers with a mean age of 44.9 (SD 6.76) years. The majority of the sample (23/24, 96%) was of White or White British ethnicity. The mean age of their children was 13.88 (SD 2.8) years, and the children were experiencing a range of mental health problems. Of the children, 29% (7/24) were boys, 66% (16/24) were girls, and 4% (1/24) identified as other (Multimedia Appendix 2).

A total of 33 HCPs from 8 NHS sites were included in the study. In total, 19 of the 33 participants were interviewed, and 12 participants participated in the FGDs. For the remaining 2 HCPs (6%), it was not possible to arrange a time that was convenient during the recruitment period. HCPs represented a broad range of clinical expertise, worked with children and young people aged from 0 to 25 years in an outpatient capacity and had an average of 7.54 (SD 6.24) years of working experience in CYPMH services (Multimedia Appendix 3).

Quantitative Study

A total of 63 parents met the eligibility criteria and consented to be part of the pilot RCT (ie, 30 from the NHS and 33 from the community sample; Figure 1). There were no significant demographic differences in the parents accessing the trial through community recruitment and those accessing through the NHS (χ^2_8=8.272; *P*=.41). Of the 63 parents, 42 (67%) parents completed baseline measures (30 from the NHS and 12 from the community sample) and were randomly assigned to control (n=12), IG1 (n=11), or IG2 (n=19). Of the 42 parents, 16 (40%) completed follow-up measures (ie, 12 from the NHS and 4 from the community sample). A total of 2 parents expressed not having time to complete the follow-up measures, and the remaining parents could not be reached. There were no significant differences between the parents who consented and completed baseline measures and those who consented but did not complete baseline measures (χ^2_8=8.766; *P*=.36). Similarly, there were no significant differences between the parents who completed follow-up measures and those who did not (χ^2_8=8.015; *P*=.43).

Figure 1. Consolidated Standards of Reporting Trials flowchart of participants in the quantitative study. IG1: intervention group 1; IG2: intervention group 2; NHS: National Health Service; PUfP: Power Up for Parents.



Only 50% (6/12) of the NHS sites were able to recruit parents to stage 2 of the study with an intraclass correlation of 0.042 on the Pediatric Shared Decision-Making Questionnaire (modified). Of the remaining 6 sites, one withdrew from the study and another site reported insufficient clinical staff to assist in identifying potential parents. The other 4 sites entered the study within the last 3 months of recruitment and reported insufficient time to recruit participants for both stages of the study. The randomized sample (n=42) was predominantly White British, English-speaking mothers, with a mean age of 45.98 (SD 6.45) years. The majority of the participants were primary caregivers of teenage girls with a mean age of 14.31 (SD 2.14) years (Multimedia Appendix 4).

Engagement With the Intervention

Google Analytics use data from January 7, 2019 and October 1, 2019, were used, as the data coincided with the recruitment of the first participant to the pilot study and the last day of data collection. App use data were made anonymous to comply with the General Data Protection Regulation and research ethical guidelines. Overall, 117 users cumulatively accessed versions 1 and 2 of the app and 72 registered an account. In total, users visited the app 288 times for an average duration of 5 min and 59 s. Less than 33% of the users visited the app and left immediately without viewing any of the features (bounce rate=32.99). An average of 3 active users were recorded for each 28 day-period during the study. The *decisions* feature was accessed 330 times, followed by *journey* 163 times, *goals* 160 times, *resources* 146 times, and *support* 103 times. All parents recruited via the NHS were guided through the setting up of the app, and web-based participants had to download the app before clicking next to indicate completion of baseline measures. Therefore, it was estimated that the majority of the intervention arm participants (n=30) accessed the intervention contributing to these statistics.

RQ2: Are the Outcome Measures Appropriate and Acceptable for a Prospective RCT?

For parents who completed baseline measures (n=42), the majority (40/42, 95%) had no missing data at baseline. The 2 cases with missing data failed to complete the Pediatric Shared Decision-Making Questionnaire (modified) and the DCS. For parents completing follow-up measures (n=16), all measures were completed by all parents, except the Poststudy System Usability Questionnaire (PSSUQ). Only parents belonging to the intervention groups were required to complete the PSSUQ, and all 5 completed it. At baseline, 53% (16/30) of the NHS cases had completed the HCP observed Control Preference Scale

(CPS). At follow-up, 58% (7/12) of the NHS cases had completed the HCP observed CPS. The HCP observed CPS was required only from parents recruited via NHS.

The outcome measures provided valuable information on parents' anxiety levels, decision-making preferences, and experiences of SDM. Data from the outcome measures were summarized and descriptively presented (Multimedia Appendix 5). Overall, the majority of parents (n=26) who participated in the study preferred to be involved in SDM. However, clinicians reported that, based on observations, parents either left the final decision to the HCP after sharing their views (n=6), got involved in SDM (n=5), or preferred to make the final decision themselves (n=4). The average Pediatric Shared Decision-Making Questionnaire (modified) score reported at baseline was 26.54, which increased to an average of 28.8 at the end of the study (higher values indicate greater levels of SDM). The average DCS increased from 35.44 to 38.18 during the study (higher values indicate greater levels of decisional conflict). In addition, the average overall satisfaction with care score increased from 20.62 to 26.18 by the end of the study (higher values indicate greater experience of service). However, the SDM construct of the ESQ highlighted that many parents did not experience SDM at baseline (32/42, 76%) and again at follow-up. The average overall anxiety scores for the sample showed scores that were above the cut-off (38) at both time points, indicating that the parents in the sample were moderate to highly anxious. The PSSUQ had a mean score ranging from 3 to 3.4 overall and on all the subscales (lower values indicate better performance and satisfaction).

RQ3: What is the Scope of the Pilot Data Collected From Users and Nonusers of PUfP?

As the Pediatric Shared Decision-Making Questionnaire scores for the overall sample shifted in a positive direction by the end of the study, this measure was investigated further to gain insight into the SDM outcome. The CIs around the estimated differences in mean scores were too wide to indicate any potential significant differences between the groups [43]. However, based on observed data, at baseline, there was a small observed difference between the control (mean 28.12, SD 9.17) and intervention groups (mean 25.86, SD 11.46) in the Pediatric Shared Decision-Making Questionnaire (2.26 points, 95% CI −5.31 to 9.92). At the end of the intervention period, a small difference was observed between the control (mean 29.36, SD 3.12) and intervention groups (mean 27.6, SD 11.89; 1.76 points, 95% CI −10.75 to 14.28). On the basis of the observations, both the control and intervention groups may have increased the behaviors of SDM over time.

For participants completing both baseline and follow-up, it was observed that the control group at baseline (mean 28.91, SD 29.36) showed very little change at follow-up (mean 29.36, SD 10.36) on the Pediatric Shared Decision-Making Questionnaire scores (−0.45 points, 95% CI −4.75 to 3.84). The intervention group showed a small difference from baseline (mean 22.2, SD 10.62) to follow-up (mean 27.6, SD 11.89; −5.4 points, 95% CI −26.56 to 15.76). Again, CIs around the estimated differences in mean scores were too wide to indicate any potential significant differences over time. These preliminary findings suggest that if the change over time was ignored, parents in the control and intervention groups may have had similar scores on the Pediatric Shared Decision-Making Questionnaire [43].

RQ4: What Are the Potential Barriers and Enablers to Conducting a Prospective RCT?

Potential barriers observed or reported by site collaborators included insufficient time for recruitment and site setup, as indicated by the challenge sites faced to recruit participants within the final 3 months of the study. Second, including a criterion that excluded parents with a mental health diagnosis decreased the number of potential participants. This was confirmed by the withdrawal of 1 site that expressed challenges with recruitment, as most parents reported having a diagnosis.

Although this pilot feasibility study highlighted potential barriers that can affect recruitment in a full RCT, the study highlighted no reports of adverse effects in either stage of the study. It was also possible to recruit a satisfactory sample (n=31) across 6 NHS CYPMH sites within 9 months and 33 participants within 3 months of community sampling. However, a high attrition was observed among the participants. No other barriers to upgrading to a full RCT were observed or identified. Input from PPI sessions and guidance from the study's steering committee were highlighted as beneficial to the intervention development and recruitment strategies.

RQ5: Which Data Collection Procedures Are Appropriate and Acceptable?

For the qualitative study, the majority of parents (12/14, 86%) opted for phone interviews. In addition, those participating in FGDs preferred this to be held at the CYPMH site and attached to an existing meeting, instead of the university. For quantitative data collection, the majority of parents preferred to complete the baseline (30/42, 71%) and follow-up (10/16, 63%) measures on the web.

Similarly, for clinicians participating in the qualitative study, all clinicians opted for phone interviews. Those participating in FGDs preferred this to take place at the CYPMH sites and attached to an existing staff meeting. Although there was no web-based option for HCPs completing the observed Control Preferences Scale, many HCPs requested to have the measure emailed or to receive a reminder email to prompt completion of the measure. In addition, both forms of randomization worked smoothly, with an unpredictable assignment to the comparison groups.

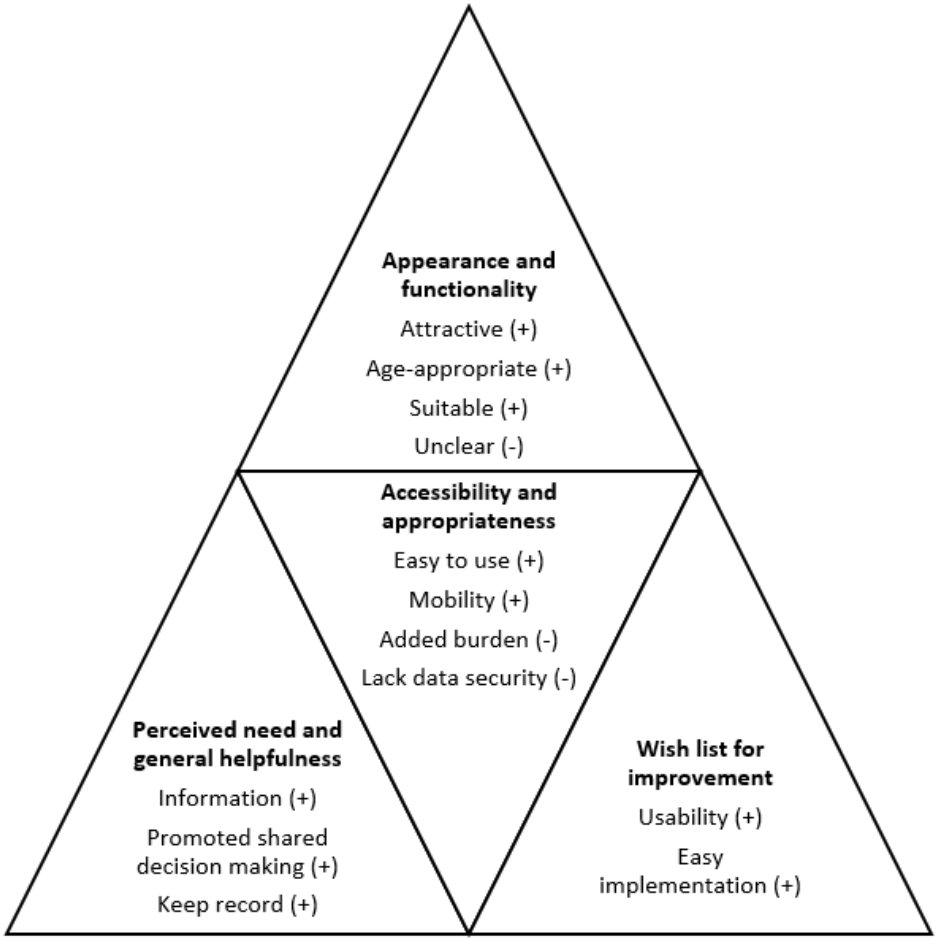
Qualitative Results

RQ6: Is PUfP Acceptable and Useful for Parents and HCPs?

Feedback revealed feasibility categories that represented acceptability, (perceived) usefulness, and scope for improvement. Participants described the *appearance and functionality* of the intervention as essential to the acceptability of PUfP. *Perceived need and general helpfulness of the intervention* and *accessibility and appropriateness* of the intervention emerged as 2 further themes describing the perceived usefulness of PUfP. Figure 2 provides a brief

overview of the themes emerging from the qualitative data, highlighting the important influencing factors.

Figure 2. An overview of the themes emerging from qualitative data.



Acceptability

Theme 1: Appearance and Functionality of the Interface

Parents' feedback on the intervention was mostly positive, generally expressing satisfaction with the intervention (11/14, 79%). Most parents described the appearance of the intervention as attractive. Parents also appreciated the layout and functionality of the intervention and described it as age-appropriate and suitable for their busy lifestyles. There was a general sentiment that images and graphics were ideal for parents:

> *I find it much easier on the eye. It gives a soothing vibe kinda thing.* [Parent, age 47 years]
>
> *Yeah, it looks good, colourful.* [Parent, age 40 years]
>
> *It's not overly childlike. Yeah, I think it looks very user friendly.* [Parent, age 53 years]

HCPs also expressed satisfaction with the appearance and provided favorable comments on the presentation of the intervention (15/19, 79%). HCPs were also positive in endorsing specific components of PUfP and its suitability for parents. HCPs highlighted that the layout and colors drew attention to the relevant features within the app:

> *I like the layout, in terms of the different sections. I think that's really good.* [HCP, 2.5 years of experience]
>
> *It's nice and clear in terms of the graphics. It tells you what it is, and the tabs are really nice.* [HCP, 7 months of experience]

Although parents and HCPs were generally satisfied with the intervention, some expressed dislike with some of the features. In addition, not all participants understood all features. Dislikes centered on having a preference for specific colors and wording. Although parents were able to find their way around the app after *clicking around* or browsing the user manual, participants expressed that clarity or further instructions are needed to guide users:

> *I would say, I don't like question mark boxes, because I think the text should maybe be in the main box itself, because it's just another thing to click on.* [HCP, 6 years of experience]
>
> *So it's not altogether clear what that [Support section] does...here I've got a plus and a minus...* [Parent, age 51 years]

Usefulness

Theme 2: Perceived Need and General Helpfulness

Parents generally provided positive feedback, highlighting that the intervention was useful (13/14, 93%). The intervention was well received by the parents, and they generally indicated that the intervention was or would be useful for them and may also help with various aspects of accessing CYPMH care. Parents echoed the potential value of the intervention to keep records, promote involvement in SDM, and signpost useful resources:

> *...and if it worked and it worked well, I'd be using it. It's really good to have all your appointments in one place as well. And the notes section, things that you think, "Oh, I need to talk to the doctor about that." Yeah, I think it sounds really good.* [Parent, age 39 years]

> *This definitely looks like something I would use.* [Parent, age 47 years]

Similarly, many HCPs expressed that the intervention was useful and would be relevant to their practice (16/19, 84%). The HCPs provided insight into the potential application and benefits of PUfP, with the majority expressing that it should make it easier to signpost families to useful resources that can support their practice:

> *It might also be helpful in terms of just understanding CAMHS. That's often one of the first hurdles that I have to get over with parents and young people, is they don't really understand our service and they don't really understand CAMHS. I think that could be quite helpful in this.* [HCP, 13 years of experience]

> *I think this can be used with any diagnosis. This is kinda helpful. With any kinda parents, this is helpful.* [HCP, 15 years of experience]

Theme 3: Accessibility and Appropriateness of the Intervention

The concept of an app received mixed views from parents and HCPs, mainly regarding usability. However, participants highlighted positive reasons for using an app and expressed that an *easy to use* and *easily accessible* app may motivate parents to at least try the intervention. Participants generally thought a digital resource provided that *instant* support and because of its dynamic nature may also help the parents by providing feedback and signposting. Participants also expressed appreciation that the intervention had the potential for use *on the go*:

> *I think even if there were parents with learning difficulties or struggled with using a bit of technology, I think, as long as they obviously had a phone, you know, that they brought with them and we were able to help guide them through it, I think that could still work as well.* [HCP, 2.5 years of experience]

> *I think I'd probably use it more on my phone because that's constantly with me. So, if something happened, like panic attack in McDonalds, like we've had before, something like this will be quite handy.* [Parent, age 36 years]

Although many participants highlighted that the intervention presented limited potential for harm, there were genuine concerns around specific groups of parents, suggesting that the intervention may be an additional burden to parents. Generally, a sense of excluding some users based on their comfort with technology or level of literacy was expressed. Similarly, data security and privacy were also highlighted as concerns. Participants expressed that sensitive data would be entered into the app, and therefore, reassurance of trustworthiness and safety would be needed:

> *just thinking about culture and ethnicity and language, and whether or not this would be available in different languages, for those that don't read English, basically.* [HCP, 2.5 years of experience]

> *as long as I'm assuming, it's obviously all secure with the data that you put on there and everything. As long as I was confident that what I was putting on there was all secure.* [Parent, age 39 years]

> *Well, I have a few illiterate parents so they may struggle with this.* [HCP, 16 years of experience]

RQ7: Can the Feedback From PUfP Users be Used to Further Refine the Prototype for the Prospective RCT?

The following theme emerged addressing the final RQ.

Theme 4: A Wish List for Improvement

Parents and HCPs appreciated that their input could potentially help further develop and improve the PUfP prototype for future research and before implementation. They suggested improvements that could enhance usability and facilitate easy implementation into practice. Feedback was either in line with refining what already existed (eg, attaching the user manual to the home screen) or adding new features that were seen as vital (eg, emergency help) or features that could promote use of the app (eg, options for emotional support such as mindfulness). The overarching theme emerged as *a wish list of improvements* for informing the development and refinement of PUfP:

> *A section on mindfulness, for themselves...* [HCP, 2.5 years of experience]

> *Sometimes a brief video of how to use the app can be useful, or testimony of another parent or carer talking about themselves can be helpful.* [HCP, 5 years of experience]

> *I think if there was under resource, if there was things like, "If this happens, do this." Maybe that would help.* [Parent, age 47 years]

> *Maybe having the manual where it is fine, but maybe there could be a smaller, I don't know, more compact, sorry, more compact version within the app itself to just remind people what each of the particular areas are for.* [Parent, age 39 years]

Overall Feasibility and Acceptability

The findings suggest that although the components of the study work well together, adjustments to the study protocol to improve recruitment are needed to proceed to an appropriately powered prospective RCT. In addition, the intervention appears to be acceptable and usable, with findings further suggesting upgrades and improvements that may benefit future trials. On the basis

of the 8 Go or No-Go criteria (Multimedia Appendix 1), this study achieved 15 points out of a possible 16. Although many points were accumulated throughout the study, on further reflection, we acknowledge the importance of an adequate sample size to facilitate a fully powered RCT. This is an important issue; therefore, we were particularly cautious in our interpretation of the points-based system.

Discussion

Summary

This study was a preliminary investigation to pilot PUfP, a novel digital evidence- and web-based app to promote SDM among parents of children and young people with mental health difficulties. This study aimed to assess the acceptability of the intervention and examine the feasibility of proceeding to a full RCT. To the best of our knowledge, this is the first RCT to pilot test an interactive parent-targeted digital SDM tool in CYPMH settings. Qualitative data revealed that PUfP may be acceptable and useful for parents and HCPs. The findings also indicate that there is scope for improvement of PUfP with suggestions for refinement and upgrade. A total of 63 parents met the eligibility criteria and consented to participate in the pilot RCT. However, 42 completed baseline measures and only 16 completed follow-up. Although there is some evidence indicating that general administrative procedures such as overall study design and selection of outcome measures are appropriate for a future RCT, the high attrition of participants suggests that some modifications should be applied to increase recruitment target numbers before upgrading to a fully powered RCT.

Results in Context With Other Research

The 15 points accumulated from the Go- or No-Go criteria indicated the successful completion of the study [44]. The main potential barrier identified for the future trial centered on recruitment. This is not uncommon among researchers recruiting in medical settings [45]. On the one hand, we obtained data from 24 parents and 31 clinicians, which were acceptable for the qualitative study and permitted data saturation [46]. On the other hand, approximately 50% of the participants were lost to follow-up when assessing feasibility measures. Although high attrition is consistent with research in web-based interventions [47], the small sample size was comparable with other studies exploring decision aids in CYPMH settings [48-52]. In addition, our sample size was higher than the recommended sample size for feasibility and pilot studies [53]. Nonetheless, a larger, more appropriate sample will render the research more efficient.

The number of eligible participants identified through social media (n=33) is consistent with other studies reporting social media as beneficial to recruitment rates [47]. In contrast to previous studies indicating challenges in recruiting HCPs [45,52], this study identified a fair sample (n=24) of interested HCPs. A possible explanation could be that the topic resonated with the clinical care agenda or is in an area of special interest to HCPs at CYPMH sites [54]. Use data, however, demonstrated the feasibility of parents accessing and using the intervention. Similar findings have been reported in the original Power Up for young people tested in schools and CYPMH services [18]. Taken together, these findings highlight that future trials of PUfP should clearly define and discriminate adherence to intervention and adherence to study protocol. The relationship between the 2 types of attrition could have implications for how the findings are interpreted [55]. For future trials of PUfP, adherence to PUfP may promote SDM; however, a meaningful sample size is needed to make the necessary comparisons.

In general, parents reported a preference to be involved in SDM. However, HCPs reported that some parents in their care displayed behaviors in line with SDM, left the decision up to HCP, or made the final decision themselves after sharing their views or listening to the HCP's recommendations. These preliminary findings are in agreement with other academics, suggesting that although SDM is preferred, not everyone engages [56,57], or it may be too challenging to implement [4,15]. This finding also highlights that, within triad relationships, varying levels of *shared* decision-making may exist [6]. This may be, in part, due to the age and capacity of children and young people.

The majority of the qualitative sample (>80%), including parents and HCPs, provided feedback consistent with the acceptance and (perceived) usefulness of the intervention. These findings demonstrate that additional support is generally well received in CYPMH settings, as indicated in other studies [48,49,52]. The 3 emerging themes highlighted the importance of the intervention for end users and may promote use. These themes also fit within previous research on the broader TAM [21]. Qualitative findings further highlighted *a wish list* of features and improvements to PUfP, which may potentially increase acceptability and usefulness. Incorporating the participants' views would be in line with human and computer interaction approaches to designing technological interventions and reinforces an opportunity to involve end users in the development of interventions. Researchers generally agree that this approach to co-designing improves usability and subsequent outcomes [58].

Investigations of parent responses to the Pediatric Shared Decision-Making Questionnaire using CIs resulted in no significant findings within and between groups. This is not surprising because of the small sample size obtained and the *as-per-protocol* analytic approach chosen [59]. However, parents' average anxiety levels were mostly above the cut-off for this study's sample. This is in line with other research suggesting that parents of children with mental health difficulties generally report higher stress levels [7].

Strengths and Limitations

The primary strength of this study was the ability to obtain data that could be useful for scaling up to a full RCT. Recruitment figures were improved by including web-based community sampling and social media advertising. Future trials could explore these forms of recruitment further, balancing the possible high proportion of incomplete data via the web-based platform versus the slow recruitment process via the NHS sites. Another strength of this trial was the consideration for respondent burden by providing the participants with options for phone or face-to-face interviews and options for completing outcome measures on the web or paper based. Notably, having limited contact details made it difficult to reach some of the

parents and resulted in a small number of parents completing follow-up measures.

Although emerging themes suggested acceptability and perceived usefulness, these themes were informed by views taken from a nonrepresentative sample that included the majority of White British, English-speaking mothers of teenage girls. A more representative sample, including fathers, other ethnicities, or underrepresented groups, can provide deeper insights into future research. A multi-site, cluster randomized approach was considered a major strength and did not incur additional intervention costs. However, potential contamination of the control group could be considered if the participants come in contact with the community study recruitment information. This may present some obstacles for the research team if the control group gained access to the intervention. In addition, a web app was chosen over a native mobile app. As a result, parents were not required to download or install it from an app store. Therefore, PUfP did not occupy space on the user's phone. It functioned as a website that is suitable for mobile devices and is usually cheaper to build, maintain, and update than native mobile apps [60].

Another major strength was the ability to gather use data via Google Analytics. Although the research team attempted to share the intervention only with the intervention arm, it was possible that site collaborators, HCPs, and parents could have shared the link with nonstudy participants. This may have affected the accuracy of the use data, and therefore, caution should be exercised when interpreting these types of data. Future studies may need to collect both use and self-report data to present a more reliable picture.

Finally, a mixed method design was also viewed as a strength at the feasibility phase of the intervention. Outcome measures provided valuable information that is of importance to the full RCT, providing a basis for estimating sample size calculations and selecting appropriate measures. Similarly, it provided estimates of the time required to complete outcome measures and gain access to the intervention. Although obtaining qualitative and quantitative data from the participants may triangulate the findings, this approach may potentially add a burden to parents. However, the mixed method approach can provide a better understanding of efficacy and efficiency and strengthen the findings of future RCTs [61].

Implications for Clinicians and Policymakers

Although these findings are preliminary, they suggest potential areas of clinical application. First, PUfP was found to be acceptable, as suggested by the parents and HCPs in our sample. The positive feedback surrounding the theme of perceived need and the general usefulness of PUfP highlighted a desire to obtain support if SDM was to be successfully applied in CYPMH. These findings are in line with those of other researchers, suggesting that policy guidelines should be considered to support parents who report feeling uninformed and excluded from services [54]. Notwithstanding the acknowledgment of the *Gillick competency* principle [62], policy guidelines specific to CYPMH could be informative for HCPs working with families of young people who are still considered being *under the care* of their parents. Finally, the findings also highlighted moderate to high levels of anxiety among parents. This may provide HCPs with a knowledge base for the emotional state of the parent population accessing CYPMH care.

Future Directions

The generalizability of our findings is unclear. However, the findings suggest that PUfP has the potential to be evaluated in future research. First, it is recommended that PUfP be upgraded and refined in line with the suggestions provided by HCPs and parents before being tested further. These suggestions can impact the usefulness and usability of the intervention. For example, incorporating mindfulness techniques or other techniques can provide additional support to parents during difficult moments. Just as important are the suggestions to include a crisis section and features to facilitate optional communication between HCPs, children and young people, and parents or parent-to-parent interactions. These improvements should also be made in collaboration with end users to ensure the suitability of the components.

In terms of the study design, it is recommended that future trials maintain a multicenter randomized controlled study design. However, a 2-arm approach may be sufficient, as growing evidence suggests that parents involved in mental health decisions may benefit from additional support [2]. Therefore, if a 3-arm study design is to be maintained, the existing body of knowledge may benefit from insights into different modes of delivery. In addition, clustered randomization is recommended to control for site-level activities that can impact family involvement in SDM [14]. However, if community recruitment is also used, comparisons can be made between samples to strengthen the findings or considerations can be made to stratify the community sample into existing clusters.

Another recommendation is to revise the eligibility criteria to allow parents with an existing mental health diagnosis to be included in future trials. These parents may actually benefit from the additional support, and therefore, future trials can control for and benefit from these statistical comparisons. It is also recommended that the future trial adopts an *intent-to-treat* analytic approach to draw accurate (unbiased) conclusions regarding the effectiveness of the intervention [59]. This approach will also be beneficial in light of the retention rates observed in this trial.

The study also benefited from the input of enthusiastic parent partners who contributed to the intervention design and study recruitment strategies. Future studies could use this PPI approach, as it possibly contributed to the smooth running of this feasibility study. Future trials can also explore extending an invitation and training to parent experts so that they can be part of the research process as interviewers or participate in the identification and recruitment process at CYPMH services. Finally, it may also be possible to estimate NHS provider costs for usual care and other interventions, in addition to the parent-reported costs to access services. Taken together, these costs can be explored to fully capture any savings to be estimated if the future trial incorporates economic evaluation to explore the cost-effectiveness of the intervention.

Conclusions

This feasibility pilot trial was designed and conducted to test the essential aspects of the research design and acceptability of the intervention to examine the potential for conducting a future fully powered RCT. Despite evidence suggesting the acceptability of PUfP, the findings suggest that recruitment modifications are needed to enhance the feasibility of collecting follow-up data before scaling up to a full RCT. Possible implications for practice and policy were discussed alongside recommendations for future research. One important recommendation is that the future RCT may benefit from incorporating a mechanism to explore the cost-effectiveness of implementing PUfP. Furthermore, in recognition of the age and capacity of young people and the promotion of standards of care to empower young service users, considerations for refining PUfP to interact with other versions of Power Up may be valuable.

Acknowledgments

The authors thank the staff at Create Health for the technical development of PUfP. The authors also thank the parent experts and the PUfP steering committee for their input and guidance on the intervention and the study design. The authors also thank the research and development team and staff at all NHS sites for their patience and cooperation in setting up the study at the respective sites. The authors thank James Sinclair at the lead site North East London NHS Foundation Trust for his support and guidance. The authors thank Professors Miranda Wolpert and Peter Fonagy for supervision and expert advice at various stages of the research project. They also thank Miles Weekes for his support at the time of writing the manuscript. This project has received funding from the European Union's Horizon 2020 Research and Innovation Program under the Marie Sklodowska-Curie grant agreement No. 722561.

Authors' Contributions

JE developed the initial idea for the intervention and was involved in the funding application process. SL worked with the staff at Create Health to develop, adapt, and refine PUfP. SL drafted the manuscript, and JE contributed to the revisions. Both authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of findings against 8 predetermined Go-No-Go criteria for feasibility research.
[DOCX File , 17 KB - formative_v5i3e25235_app1.docx]

Multimedia Appendix 2
Characteristics of parents participating in interviews and focus group discussions.
[DOCX File , 15 KB - formative_v5i3e25235_app2.docx]

Multimedia Appendix 3
Characteristics of HCPs participating in interviews and focus group discussions.
[DOCX File , 14 KB - formative_v5i3e25235_app3.docx]

Multimedia Appendix 4
Demographic characteristics of parent participants in stage 2 of the feasibility trial.
[DOCX File , 16 KB - formative_v5i3e25235_app4.docx]

Multimedia Appendix 5
Summary of outcome data.
[DOCX File , 17 KB - formative_v5i3e25235_app5.docx]

Multimedia Appendix 6
CONSORT-EHEALTH checklist (V.1.6.1).
[PDF File (Adobe PDF File), 1139 KB - formative_v5i3e25235_app6.pdf]

References

1. Elwyn G, Durand MA, Song J, Aarts J, Barr PJ, Berger Z, et al. A three-talk model for shared decision making: multistage consultation process. Br Med J 2017 Nov 06;359:j4891 [FREE Full text] [doi: 10.1136/bmj.j4891] [Medline: 29109079]

2. Wyatt KD, List B, Brinkman WB, Prutsky Lopez G, Asi N, Erwin P, et al. Shared decision making in pediatrics: a systematic review and meta-analysis. Acad Pediatr 2015 Nov;15(6):573-583. [doi: 10.1016/j.acap.2015.03.011] [Medline: 25983006]
3. Shay LA, Lafata JE. Where is the evidence? A systematic review of shared decision making and patient outcomes. Med Decis Making 2014 Oct 28;35(1):114-131. [doi: 10.1177/0272989x14551638]
4. Gondek D, Edbrooke-Childs J, Velikonja T, Chapman L, Saunders F, Hayes D, et al. Facilitators and barriers to person-centred care in child and young people mental health services: a systematic review. Clin Psychol Psychother 2017 Jul;24(4):870-886. [doi: 10.1002/cpp.2052] [Medline: 27910173]
5. Weiss EM, Clark JD, Heike CL, Rosenberg AR, Shah SK, Wilfond BS, et al. Gaps in the implementation of shared decision-making: illustrative cases. Pediatrics 2019 Mar 05;143(3):e20183055 [FREE Full text] [doi: 10.1542/peds.2018-3055] [Medline: 30824603]
6. Gabe J, Olumide G, Bury M. It takes three to tango? A framework for understanding patient partnership in paediatric clinics. Social Science and Medicine 2004 2004. [doi: 10.1111/j.1365-2214.2005.00500_4.x]
7. "There for you": the role of parents in supporting young people with mental health problems : results of a survey. Association of Young People's Health (AYPH). 2016. URL: http://www.youngpeopleshealth.org.uk/wp-content/uploads/2016/11/AYPH-Parenting-briefing-11-nov-2016.pdf [accessed 2021-02-05]
8. Mandarelli G, Sabatello U, Lapponi E, Pace G, Ferrara M, Ferracuti S. Treatment decision-making capacity in children and adolescents hospitalized for an acute mental disorder: the role of cognitive functioning and psychiatric symptoms. J Child Adolesc Psychopharmacol 2017 Jun;27(5):462-465 [FREE Full text] [doi: 10.1089/cap.2016.0092] [Medline: 27935747]
9. Baker-Ericzén MJ, Jenkins MM, Haine-Schlagel R. Therapist, parent, and youth perspectives of treatment barriers to family-focused community outpatient mental health services. J Child Fam Stud 2013 Aug 01;22(6):854-868 [FREE Full text] [doi: 10.1007/s10826-012-9644-7] [Medline: 24019737]
10. Boland L, Kryworuchko J, Saarimaki A, Lawson ML. Parental decision making involvement and decisional conflict: a descriptive study. BMC Pediatr 2017 Jun 13;17(1):146 [FREE Full text] [doi: 10.1186/s12887-017-0899-4] [Medline: 28610580]
11. Simmons MB, Hetrick SE, Jorm AF. Making decisions about treatment for young people diagnosed with depressive disorders: a qualitative study of clinicians' experiences. BMC Psychiatry 2013 Dec 12;13(1). [doi: 10.1186/1471-244x-13-335]
12. Allan LQC. The children act 1989. Medicolegal Essentials in Healthcare, Second Edition 2004. [doi: 10.1017/cbo9780511545559.018]
13. Haine-Schlagel R, Walsh NE. A review of parent participation engagement in child and family mental health treatment. Clin Child Fam Psychol Rev 2015 Jun 1;18(2):133-150 [FREE Full text] [doi: 10.1007/s10567-015-0182-x] [Medline: 25726421]
14. Edbrooke-Childs J, Jacob J, Argent R, Patalay P, Deighton J, Wolpert M. The relationship between child- and parent-reported shared decision making and child-, parent-, and clinician-reported treatment outcome in routinely collected child mental health services data. Clin Child Psychol Psychiatry 2016 Apr;21(2):324-338. [doi: 10.1177/1359104515591226] [Medline: 26104790]
15. Boland L, Graham ID, Légaré F, Lewis K, Jull J, Shephard A, et al. Barriers and facilitators of pediatric shared decision-making: a systematic review. Implement Sci 2019 Jan 18;14(1):7 [FREE Full text] [doi: 10.1186/s13012-018-0851-5] [Medline: 30658670]
16. Liverpool S, Pereira B, Hayes D, Wolpert M, Edbrooke-Childs J. A scoping review and assessment of essential elements of shared decision-making of parent-involved interventions in child and adolescent mental health. Eur Child Adolesc Psychiatry 2020 Apr 16. [doi: 10.1007/s00787-020-01530-7] [Medline: 32300893]
17. Liverpool S, Webber H, Matthews R, Wolpert M, Edbrooke-Childs J. A mobile app to support parents making child mental health decisions: protocol for a feasibility cluster randomized controlled trial. JMIR Res Protoc 2019 Aug 14;8(8):e14571 [FREE Full text] [doi: 10.2196/14571] [Medline: 31414665]
18. Edbrooke-Childs J, Edridge C, Averill P, Delane L, Hollis C, Craven MP, et al. A feasibility trial of power up: smartphone app to support patient activation and shared decision making for mental health in young people. JMIR Mhealth Uhealth 2019 Jun 04;7(6):e11677 [FREE Full text] [doi: 10.2196/11677] [Medline: 31165709]
19. Chapman L, Edbrooke-Childs J, Martin K, Webber H, Craven MP, Hollis C, et al. A mobile phone app to support young people in making shared decisions in therapy (power up): study protocol. JMIR Res Protoc 2017 Oct 30;6(10):e206 [FREE Full text] [doi: 10.2196/resprot.7694] [Medline: 29084708]
20. Liverpool S, Mota C, Sales C, Čuš A, Carletto S, Hancheva C, et al. Engaging children and young people in digital mental health interventions: systematic review of modes of delivery, facilitators, and barriers. J Med Internet Res 2020 Jun 23;22(6):e16317 [FREE Full text] [doi: 10.2196/16317] [Medline: 32442160]
21. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A systematic review of the technology acceptance model in health informatics. Appl Clin Inform 2018 Jul 15;9(3):604-634 [FREE Full text] [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
22. Deaton A, Cartwright N. Understanding and misunderstanding randomized controlled trials. Soc Sci Med 2018 Aug;210:2-21 [FREE Full text] [doi: 10.1016/j.socscimed.2017.12.005] [Medline: 29331519]

23. Gillies K, Chalmers I, Glasziou P, Elbourne D, Elliott J, Treweek S. Reducing research waste by promoting informed responses to invitations to participate in clinical trials. Trials 2019 Oct 28;20(1):613 [FREE Full text] [doi: 10.1186/s13063-019-3704-x] [Medline: 31661029]
24. Akobeng AK. Understanding randomised controlled trials. Arch Dis Child 2005 Aug;90(8):840-844 [FREE Full text] [doi: 10.1136/adc.2004.058222] [Medline: 16040885]
25. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010 Jan 06;10:1 [FREE Full text] [doi: 10.1186/1471-2288-10-1] [Medline: 20053272]
26. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol 2010;10:67 [FREE Full text] [doi: 10.1186/1471-2288-10-67] [Medline: 20637084]
27. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. Br Med J 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
28. United Nations, Department of Economic and Social Affairs Youth Internet. United Nations. 1985. URL: https://www.un.org/development/desa/youth/what-we-do/faq.html [accessed 2020-03-20]
29. Anwyl-Irvine A, Massonnié J, Flitton A, Kirkham N, Evershed J. Gorilla in our midst: an online behavioral experiment builder. bioRxiv 2018 2018. [doi: 10.1101/438242]
30. Power up for parents: a pilot study to enhance shared decision-making in CAMH. Association of Child and Adolescent Mental Health. URL: https://www.acamh.org/blog/power-parents-pilot-study/ [accessed 2020-04-06]
31. Pyke-Grimm KA, Degner L, Small A, Mueller B. Preferences for participation in treatment decision making and information needs of parents of children with cancer: a pilot study. J Pediatr Oncol Nurs 1999 Jan;16(1):13-24. [doi: 10.1177/104345429901600103] [Medline: 9989013]
32. Kriston L, Scholl I, Hölzel L, Simon D, Loh A, Härter M. The 9-item Shared Decision Making Questionnaire (SDM-Q-9). Development and psychometric properties in a primary care sample. Patient Educ Couns 2010 Jul;80(1):94-99. [doi: 10.1016/j.pec.2009.09.034] [Medline: 19879711]
33. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making 2016 Jul 02;15(1):25-30. [doi: 10.1177/0272989x9501500105]
34. Brown A, Ford T, Deighton J, Wolpert M. Satisfaction in child and adolescent mental health services: translating users' feedback into measurement. Adm Policy Ment Health 2014 Jul;41(4):434-446. [doi: 10.1007/s10488-012-0433-9] [Medline: 22829193]
35. Lewis JR. Psychometric evaluation of the pssuq using data from five years of usability studies. Int J Human-Comp Intera 2002 Sep;14(3-4):463-488. [doi: 10.1080/10447318.2002.9669130]
36. Quek K, Low W, Razack A, Loh C, Chua C. Reliability and validity of the Spielberger State-Trait Anxiety Inventory (STAI) among urological patients: a Malaysian study. Medical J Malay 2004;59(2):267. [doi: 10.1002/9780470479216.corpsy0943]
37. Chan A, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. Br Med J 2013 Jan 08;346:e7586 [FREE Full text] [doi: 10.1136/bmj.e7586] [Medline: 23303884]
38. Braun V, Clarke V. Using thematic analysis in psychology. Quali Res in Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
39. Silverman D. Interpreting qualitative data? Methods for analysing talk, text and interaction. In: 2nd ed. London: SAGE Publications Ltd; 2001.
40. QSR International Pty Ltd. What is NVivo? QSR International 2015 [FREE Full text] [doi: 10.1016/j.bbapap.2010.04.001.Molecular]
41. Plaza B. Google Analytics for measuring website performance. Tour Manag 2011 Jun;32(3):477-481. [doi: 10.1016/j.tourman.2010.03.015]
42. IBM Corp. IBM SPSS Statistics for Windows, version 24. IBM Corp. 2016. URL: https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-24 [accessed 2021-02-05]
43. Cousineau D. Varieties of confidence intervals. Adv Cogn Psychol 2017 Jun 30;13(2):140-155 [FREE Full text] [doi: 10.5709/acp-0214-z] [Medline: 28729890]
44. Carbonell P, Rodriguez Escudero AI, Munuera Aleman JL. Technology newness and impact of go/no-go criteria on new product success. Marketing Letters 2004 Jul;15(2/3):81-97. [doi: 10.1023/b:mark.0000047386.33282.51]
45. Bower P, Wallace P, Ward E, Graffy J, Miller J, Delaney B, et al. Improving recruitment to health research in primary care. Fam Pract 2009 Oct;26(5):391-397. [doi: 10.1093/fampra/cmp037] [Medline: 19549623]
46. Guest G, Bunce A, Johnson L. How many interviews are enough? Field Methods 2016 Jul 21;18(1):59-82. [doi: 10.1177/1525822X05279903]
47. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]

48. Ahmed R, McCaffery KJ, Silove N, Butow P, Clarke S, Kohn M, et al. The evaluation of a question prompt list for attention-deficit/hyperactivity disorder in pediatric care: a pilot study. Res Social Adm Pharm 2017 Jan;13(1):172-186. [doi: 10.1016/j.sapharm.2016.01.009] [Medline: 27086063]
49. Hayes D, Town R, Lemoniatis E. Developing and piloting i-THRIVE grids internet. 2018. URL: http://www.implementingthrive.org/i-thrive-grids/ [accessed 2021-02-05]
50. Ossebaard HC, van Gemert-Pijnen JEWC, Sorbi MJ, Seydel ER. A study of a Dutch online decision aid for parents of children with ADHD. J Telemed Telecare 2010 Jan 19;16(1):15-19. [doi: 10.1258/jtt.2009.001006] [Medline: 20086262]
51. Ahmed R, Raynor DK, McCaffery KJ, Aslani P. The design and user-testing of a question prompt list for attention-deficit/hyperactivity disorder. BMJ Open 2014 Dec 16;4(12):e006585 [FREE Full text] [doi: 10.1136/bmjopen-2014-006585] [Medline: 25515843]
52. Brinkman WB, Hartl Majcher J, Poling LM, Shi G, Zender M, Sucharew H, et al. Shared decision-making to improve attention-deficit hyperactivity disorder care. Patient Educ Couns 2013 Oct;93(1):95-101 [FREE Full text] [doi: 10.1016/j.pec.2013.04.009] [Medline: 23669153]
53. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract 2004 May;10(2):307-312. [doi: 10.1111/j..2002.384.doc.x] [Medline: 15189396]
54. Wolpert M, Hoffman J, Abrines N, Feltham A, Baird L, Law D, et al. Closing the gap through changing relationships: final report for closing the gap through changing relationships (award holders). London: The Health Foundation; 2012. URL: https://www.health.org.uk/sites/default/files/SharedDecisionMakingCAMHS_report.pdf [accessed 2021-02-05]
55. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
56. Degner LF, Sloan JA. Decision making during serious illness: what role do patients really want to play? J Clin Epidemiol 1992 Sep;45(9):941-950. [doi: 10.1016/0895-4356(92)90110-9] [Medline: 1432023]
57. Levinson W, Kao A, Kuby A, Thisted RA. Not all patients want to participate in decision making. A national study of public preferences. J Gen Intern Med 2005 Jun;20(6):531-535 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.04101.x] [Medline: 15987329]
58. Blandford A, Gibbs J, Newhouse N, Perski O, Singh A, Murray E. Seven lessons for interdisciplinary research on interactive digital health interventions. Digit Health 2018;4 [FREE Full text] [doi: 10.1177/2055207618770325] [Medline: 29942629]
59. Sedgwick P. Intention to treat analysis versus per protocol analysis of trial data. Br Med J 2015 Feb 06;350:h681. [doi: 10.1136/bmj.h681] [Medline: 25663096]
60. Charland A, Leroux B. Mobile application development. Commun. ACM 2011 May 01;54(5):49-53. [doi: 10.1145/1941487.1941504]
61. Regnault A, Willgoss T, Barbic S, International Society for Quality of Life Research (ISOQOL) Mixed Methods Special Interest Group (SIG). Towards the use of mixed methods inquiry as best practice in health outcomes research. J Patient Rep Outcomes 2017 Apr 11;2(1):19 [FREE Full text] [doi: 10.1186/s41687-018-0043-8] [Medline: 29757311]
62. Great Britain. England. Court of Appeal, Civil Division. Gillick v West Norfolk and Wisbech Area Health Authority. All Engl Law Rep 1984;1985(1):533-559. [Medline: 11648530]

Abbreviations

CPS: Control Preference Scale
CYPMH: children's and young people's mental health
DCS: Decisional Conflict Scale
ESQ: Experience of Service Questionnaire
FGD: focus group discussion
NHS: National Health Service
PPI: patient and public involvement
PSSUQ: Poststudy System Usability Questionnaire
PUfP: Power Up for Parents
RCT: randomized controlled trial
REC: research ethics committee
RQ: research question
SDM: shared decision-making
TAM: Technology Acceptance Model

Edited by G Eysenbach; submitted 23.10.20; peer-reviewed by K Harrington, AV Das; comments to author 25.11.20; revised version received 27.12.20; accepted 17.01.21; published 02.03.21.

Please cite as:
Liverpool S, Edbrooke-Childs J
Feasibility and Acceptability of a Digital Intervention to Support Shared Decision-making in Children's and Young People's Mental Health: Mixed Methods Pilot Randomized Controlled Trial
JMIR Form Res 2021;5(3):e25235
URL: https://formative.jmir.org/2021/3/e25235
doi:10.2196/25235
PMID:33650973

Original Paper

A Digital Health Tool to Understand and Prevent Cannabis-Impaired Driving Among Youth: A Cross-sectional Study of Responses to a Brief Intervention for Cannabis Use

Georgina Moreno[1,2*], BA, MA, PhD; Trevor van Mierlo[3*], BA(Hons), MScCH, MBA, GEMBA, MScBMR, DBA

[1]Department of Economics, California State University, Long Beach, Long Beach, CA, United States

[2]South Bay Economics LLC, Hermosa Beach, CA, United States

[3]Evolution Health Systems, Toronto, ON, Canada

[*]all authors contributed equally

Corresponding Author:
Georgina Moreno, BA, MA, PhD
Department of Economics
California State University, Long Beach
1250 Bellflower Blvd MS-4607
Long Beach, CA, 90840-4607
United States
Phone: 1 310 600 1767
Email: gigi.moreno@csulb.edu

Abstract

Background: Cannabis legalization has raised concern about an increased risk of cannabis-impaired driving, particularly among youth. Youth advocates and policy makers require cost-effective tools to target educational resources to promote responsible cannabis use.

Objective: The objective of this paper is threefold. First, it describes how a youth advocacy organization disseminated a low-cost digital brief intervention to educate and inform young people about responsible cannabis use. Second, it illustrates how digital tools can help promote understanding about attitudes and behaviors toward cannabis while simultaneously offering tailored education. Finally, this paper contributes to examining behavioral factors associated with youth cannabis-impaired driving by quantifying relationships between cannabis users' willingness to drive impaired and self-reported demographic and behavioral factors.

Methods: This paper analyzed data from 1110 completed Check Your Cannabis (CYC) brief interventions between March 2019 and October 2020. The CYC asks respondents a brief set of questions about their cannabis use and their personal beliefs and behaviors. Respondents receive comprehensive feedback about their cannabis use and how it compares with others. They also receive a summary of reported behaviors with brief advice. An ordered probit model was used to test relationships between cannabis use, demographics, and driving behaviors to gain further insights.

Results: The vast majority (817/1110, 73.6%) of respondents reported using cannabis. However, a much smaller share of respondents reported problems associated with their cannabis use (257/1110, 23.2%) or driving after cannabis use (342/1110, 30.8%). We found statistically significant relationships between driving after cannabis use and age; Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST) risk score; and polysubstance use. However, we did not find gender to be a significant determinant of driving after cannabis use. We estimated that every 10-point increase in the ASSIST score increased the probability of sometimes driving after cannabis use by 7.3% (P<.001). Relative to respondents who reported never drinking alcohol or using other substances with cannabis, those who sometimes drink or use other substances with cannabis were 13% (P<.001) more likely to sometimes or always drive after using cannabis.

Conclusions: The digital health tool cost the youth advocacy organization approximately Can $0.90 (US $0.71) per use. Due to the tool's unlimited use structure, the per-use cost would further decrease with increased use by the organization’s target population. Based on our results, public health campaigns and other interventions may consider tailoring resources to frequent cannabis users, youth with high ASSIST scores, and those with polysubstance abuse. The cost-effectiveness of delivering digital brief interventions with unlimited use is attractive, as increased use decreases the per-user cost. Further research examining the efficacy of digital health interventions targeting problematic cannabis use is required.

(JMIR Form Res 2021;5(3):e25583) doi:10.2196/25583

KEYWORDS

cannabis use; driving after cannabis use; internet; intervention; online intervention; digital health; cannabis; drug; online tool; youth; adolescent; Canada

Introduction

Background

The legalization of cannabis in Canada and several US states has raised concerns over cannabis-impaired driving. For example, a 2019 Gallup poll found that the top concern among opponents of legalization in the United States was the belief that legalization would increase vehicle accidents [1]. A 2019 survey found that 71% of Canadians were concerned about the impacts of cannabis legalization on road safety [2]. At the same time, many younger drivers do not perceive driving after cannabis use to be a safety risk, and some believe that cannabis use improves driving performance [3-5]. Although cannabis use impairs driving [6-11], practical and feasible means for testing for cannabis impairment are limited [8,11].

Hall [11] acknowledged that a major problem in addressing cannabis-impaired driving is the lack of a practical test for impairment and called for changing attitudes toward cannabis-impaired driving. Wadsworth and Hammond [12] compared patterns of cannabis use among youth in Canada, England, and the United States [12]. They found higher rates of driving after cannabis use among US youth relative to youth in Canada and England and attributed this finding to greater accessibility of cannabis and lower perceived harm among US youth [12]. Wadsworth and Hammond's [12] findings suggest that understanding cannabis risk perceptions may help development of cannabis harm reduction programs.

Using an online survey of US cannabis users, Borodovsky et al [5] examined the relationship between perceived intoxication levels and driving after cannabis use and found evidence that the perception of a safe level of cannabis intoxication, not the typical level of intoxication, is associated with driving after using cannabis use. These findings support the need for identifying perceptions of harm from cannabis use and suggest that tools that can identify these perceptions may inform preventative messaging and screening for harmful cannabis use. One approach for identifying and mitigating risks associated with cannabis use is implementing a digital health screener, such as the Check Your Cannabis (CYC) brief intervention, an anonymous digital health brief intervention designed for personal computers, tablets, and smartphones [13].

The CYC, developed by Evolution Health Systems and independent academic researchers, asks respondents questions about their cannabis use and provides respondents with personalized feedback about how the severity of their cannabis use compares with others of the same age and gender. The CYC screener provides a nonjudgmental approach for educating and positively influencing respondents' awareness of harm from cannabis consumption.

The primary theoretical behavior change constructs used in the CYC design are normative feedback, harm reduction, and motivational interviewing. The CYC's design, workflow, and technical infrastructure are based on a similar brief intervention for addressing alcohol use, Check Your Drinking (CYD). Evolution Health Systems and independent academic researchers also developed this tool [14]. Like the CYC, the CYD provides respondents with a brief tailored feedback report summarizing personal substance use and comparing it with others of the same age, sex, and country of residence [14]. The CYD has been subject to 7 randomized controlled trials demonstrating support of its efficacy in reducing alcohol consumption [14-20].

The CYC is publicly available at no cost at a number of URLs [21,22]. The intervention consists of a 13-item questionnaire (Multimedia Appendix 1) and an output report tailored to the user (Multimedia Appendix 2). The first 3 questions of the CYC collect demographic data used to tailor the report to the individual (first name or anonymous nickname, gender, age). Questions 4 and 5 address personal use of cannabis in the past 3 months. Questions 6 to 9 address personal negative consequences of cannabis use. Question 10 addresses driving behavior following cannabis use, and question 11 addresses polysubstance abuse. Questions 12 and 13 request information about the user's cannabis expenditure patterns. Before the user submits their responses, they must actively endorse a checkbox that acknowledges that their nonpersonal information will be used for improving the tool, and they receive a link to the intervention's privacy policy.

Upon completing the questionnaire, the user receives a tailored personal report divided into either 3 or 4 sections, depending on user inputs and output algorithms. The first section contains normative feedback based on the user's age, gender, country of residence (United States or Canada), and cannabis use patterns, presented in text and graphical format. The second section reports the user's estimated annual expenditure on cannabis and compares this expenditure to purchases of movie passes and pizza slices. The third section reports the user's Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST) score through text and graphical representation, which can be used to evaluate whether a person's cannabis use is problematic. If users indicate that they drive after cannabis use or use other substances with cannabis, their report includes a fourth section with information about the risks of cannabis-impaired driving, polysubstance use, or both.

In 2019, Parent Action on Drugs (PAD), an Ontario-based community youth advocacy group, licensed a white-label version of CYC from Evolution Health Systems at the standard nonprofit annual rate of Can $1000 (US $787.01), including hosting and reporting fees. Use of the tool is unlimited, meaning that there is no restriction on the number of respondents or reports generated. PAD also entered into a research partnership with Evolution Health Systems in which Evolution Health Systems assisted with the intervention dissemination strategy, data analysis, and technical guidance. PAD's white-label version of

the digital intervention is available for public use on PAD's website [23].

The primary purpose of PAD licensing a white-label version of the CYC was to offer high school students in the Greater Toronto Area information and advice about cannabis use, which Canada had recently legalized. A secondary purpose was to collect data on cannabis consumption patterns, polysubstance abuse, and driving.

Objective

This paper analyzes the data collected from PAD's licensed version of the CYC and focuses on the behavioral factors associated with cannabis-impaired driving from the cannabis user's perspective. The objective of this paper is threefold. First, it illustrates how low-cost brief interventions can be broadly implemented to educate young people about responsible cannabis use. Second, data from these tools can help promote understanding about attitudes and behaviors toward cannabis while anonymously offering tailored education. Finally, this paper contributes to examining behavioral factors associated with cannabis-impaired driving by quantifying relationships between cannabis users' willingness to drive impaired and self-reported demographic and behavioral factors.

Methods

This paper's analysis relies on data obtained between March 2019 and October 2020 from the CYC brief intervention. Identical versions of the CYC intervention are available from the Evolution Health website or PAD's websites [21-23]. PAD promoted their version of the CYC on their website and at speaking engagements at several high schools in the Greater Toronto Area. Participants were ad libitum and anonymous, and there were no incentives for completing the intervention. All study participants consented to the use of their anonymous data for research purposes. This study relied on convenience nonprobabilistic sampling, in which respondents self-selected to respond to the digital brief intervention questionnaire [24]. Multimedia Appendices 1 and 2 include images of the CYC interface.

Between March 28, 2019, and October 23, 2020, we collected 1553 completed CYC questionnaires. With the goal of informing responsible cannabis use among young adults in Canada, we limited the sample to participants aged 25 years and younger from Canada, thus reducing the sample to 1175. We eliminated from the analysis 34 responses that appeared to be duplicates. We further reduced the sample by dropping from the analysis 6 responses with extreme outlier values for the reported expenditures on cannabis and 25 responses with inconsistencies (eg, reported not using cannabis but reported positive expenditures on cannabis). The final study data set included 1110 responses. Table 1 summarizes the age and gender distribution for the included responses.

Table 1. Age and gender distribution for included respondents.

Gender	Minimum age (years)	Median age (years)	Maximum age (years)	Observations, n
Female	12	17	25	447
Male	12	17	25	593
Transgender or nonbinary	14	16	25	50
Not reported	14	16	18	20
Overall	12	17	25	1110

Data were extracted from the intervention platform's precoded custom structure query language database. All responses were anonymous, and data collection procedures adhered to American and Canadian privacy guidelines [25-27]. Respondents actively endorsed a checkbox consenting to the use of their nonpersonal responses to analyze and improve the intervention and received a link to the intervention's privacy policy. As the study relies on unidentifiable archival data, the study was exempt from further review.

After reporting gender and age, respondents answered 8 questions related to their use of cannabis. The questionnaire allowed respondents to report their gender as female, male, transgender, nonbinary, or not reported. Respondents also reported age as an open-ended question. The brief intervention asks respondents about their expenditures on cannabis and measures expenditures in 2 ways: the average monthly expenditure on cannabis over the past year and the largest single-day expenditure over the past year. Average monthly expenditures capture typical use. Expenditures are reported in Canadian dollars.

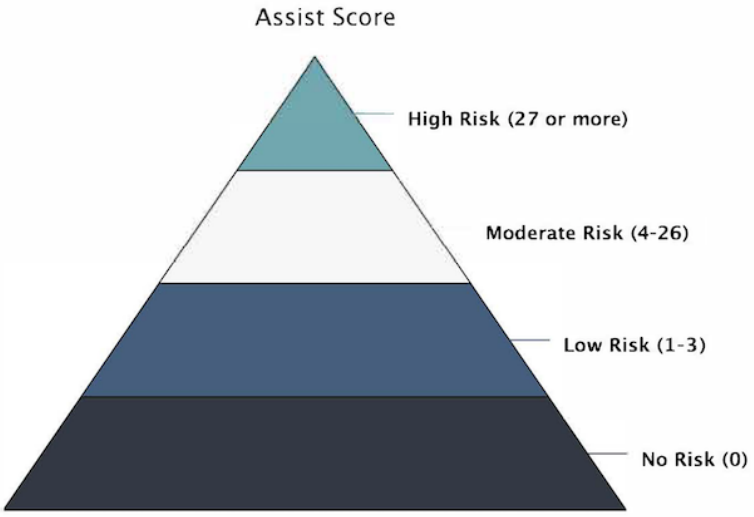
A key component of the intervention is a calculated score from the 6-item ASSIST developed by the World Health Organization (WHO) [28]. The ASSIST evaluates an individual's use of cannabis and classifies individuals into risk categories, with high risk indicating the experience of problems resulting from cannabis consumption. Figure 1 is an illustration explaining the ASSIST score after completion of the CYC questionnaire. The WHO ASSIST score is the sum of weights assigned to the 6 questions listed in Table 2 [28]. The questionnaire asks respondents to answer the 6 ASSIST questions. We computed the ASSIST score based on the response weights, also shown in Table 2.

Figure 1. Check Your Cannabis ASSIST score. ASSIST: Alcohol, Smoking, and Substance Involvement Screening Test.

Your ASSIST Score?

The ASSIST score was developed by the World Health Organization to evaluate a person's use of cannabis. The ASSIST score shows whether a person's cannabis use should be considered a problem. Higher scores usually mean serious problems. The chart is in the shape of a pyramid to show that there are more people with low ASSIST scores than high ones.

Your ASSIST score is **19**. The white area of the chart shows where your score falls.



Score	Definition	Probable Life Consequences
0	No risk	No cannabis use in the last 3 months.
1-3	Low risk	You are at low risk of health and other problems from your current pattern of cannabis use.
4-26	Moderate risk	You are at risk of health and other problems from your current cannabis use.
27 or more	High risk	You are at high risk of experiencing severe problems (health, social, financial, legal, relationship) as a result of your current pattern of use and are likely to be dependent on cannabis.

You noted that you **usually** drive after using cannabis, or ride with a driver who has. Research has shown that driving high causes you to weave in and out of lanes, slow down and then speed up, miss signs, misjudge time and distance, exhibit slow reaction time, and it can double your risk for a serious crash. Treat cannabis like alcohol and do not drive, or arrange a ride from a sober driver.

You noted that you **usually** drink alcohol or use other substances when using cannabis. Here are some common risks:

Table 2. ASSIST score questions, possible responses, and WHO weights [28]. ASSIST: Alcohol, Smoking, and Substance Involvement Screening Test; WHO: World Health Organization.

Question	Response (WHO weight)
1. During the past 3 months, how often have you used cannabis?	Never (0); once or twice (4); monthly (5); weekly (6); daily or almost daily (7)
2. During the past 3 months, how often have you had a strong desire or urge to use cannabis?	Never (0); once or twice (4); monthly (5); weekly (6); daily or almost daily (7)
3. During the past 3 months, how often has your use of cannabis led to health, social, legal, or financial problems?	Never (0); once or twice (4); monthly (5); weekly (6); daily or almost daily (7)
4. During the past 3 months, how often have you failed to do what was normally expected of you because of your use of cannabis?	Never (0); once or twice (4); monthly (5); weekly (6); daily or almost daily (7)
5. Has a friend or relative or anyone else ever expressed concern about your use of cannabis?	No, never (0); yes, but not in the past 3 months (3); yes, in the past 3 months (6)
6. Have you ever tried and failed to control, cut down, or stop using cannabis?	No, never (0); yes, but not in the past 3 months (3); yes, in the past 3 months (6)

Respondents also answered questions about driving after using cannabis and polysubstance use. These questions were “How often do you drive after using cannabis, or ride with someone who has?” and “When you use cannabis, how often do you drink alcohol or use other substances?” The possible responses for these questions were “Never,” “Sometimes,” “Usually,” or “Always.”

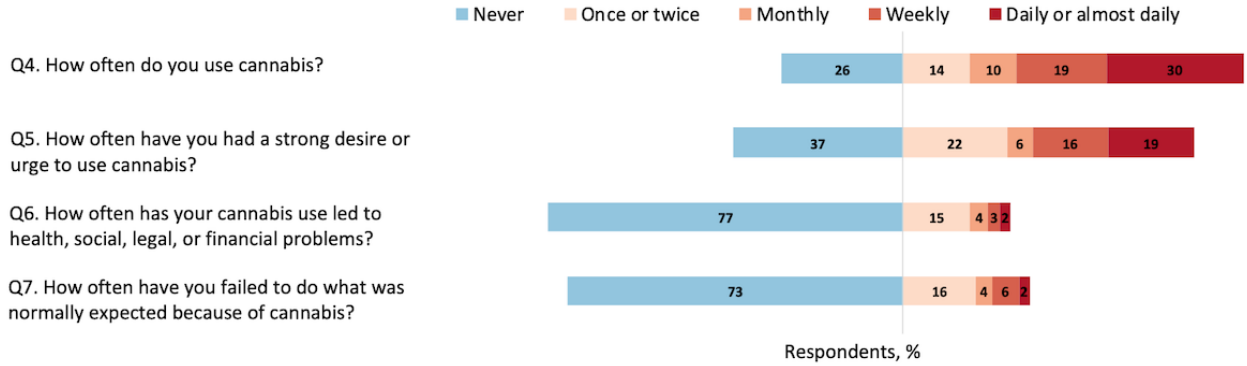
To better understand the factors associated with driving after cannabis use, we modeled driving after cannabis use as a function of age, gender, ASSIST score, whether the respondent drinks or uses other substances with cannabis (ie, polysubstance use), and cannabis expenditures. Given the qualitative and ordered nature of the responses to the question about driving after cannabis use, we estimated the probability of driving after using cannabis using an ordered probit regression [29]. We combined the “usually drives” and “always drives” categories for the ordered probit model. Therefore, the dependent variable in the model included 3 categories: (1) never drives after using cannabis, (2) sometimes drives after using cannabis, and (3) usually or always drives after using cannabis. For this analysis, we limited the data to respondents of driving age (n=913). All analysis was conducted using Stata 16 (StataCorp) [30].

We hypothesized that age has an increasing and diminishing effect on driving after using cannabis and included age as a quadratic in the ordered probit model. Because the vast majority of participants in the analytical data reported their gender as female (447/1110, 40.3%) or male (593/1110, 53.4%), we defined the gender variable in the model as female, male, or other. Based on past studies, we expected differences by gender in cannabis use and attitudes [31]. We also hypothesized that respondents with higher ASSIST scores are more likely to drive after using cannabis, since a higher ASSIST score indicates a higher risk of harmful cannabis use. We hypothesized that polysubstance use is positively correlated with driving after cannabis use, as it reflects risky behaviors. To the extent that higher expenditures reflect greater use, we hypothesized that higher expenditures are associated with heavier cannabis use and therefore a higher probability of driving after using cannabis.

Results

Figures 2-4 report the distribution of responses for the 3 categories of questions included in the brief intervention: (1) cannabis use behaviors in the past 3 months, (2) concerns about cannabis use, and (3) driving after cannabis use and polysubstance use. While 73.6% (817/1110) of respondents reported using cannabis and 63.1% (700/1110) reported having a desire or urge to use cannabis, over 76.8% (853/1110) reported never having problems with cannabis use (Figure 2, question 6) and 72.7% (807/1110) reported never failing to meet expectations because of their cannabis use (Figure 2, question 7). However, among the heaviest cannabis users or those using daily or almost daily (n=330), 48.8% (161/330) reported having incidents of health, social, legal, or financial problems associated with cannabis use.

Figure 2. Percentage of respondents aged 12 to 25 years who reported cannabis use and problems in the past 3 months (N=1110).



As seen in Figure 3, 22.0% (244/1110) of respondents reported having problems controlling their cannabis use (question 9). Although only 12.6% (140/1110) reported having such problems in the past 3 months, 79.3% (111/140) of the respondents in this category were aged between 12 and 18 years, below the legal age for cannabis consumption in Canada. When asked about others' concerns over their cannabis use (question 8), a smaller proportion, 62.4% (693/1110), reported "never." Figure 4 shows that the vast majority of driving-age respondents, 66.5% (607/913), reported never driving after using cannabis (question 10), and 61.0% (557/913) reported using alcohol and other substances with cannabis (question 11).

Figure 3. Percentage of respondents aged 12 to 25 years who reported concerns about their cannabis use (N=1110).

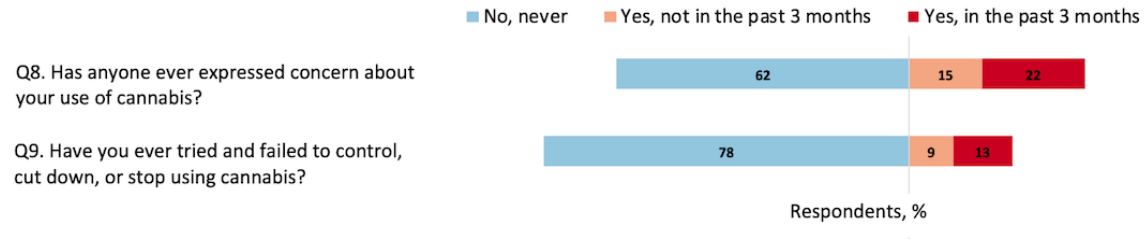
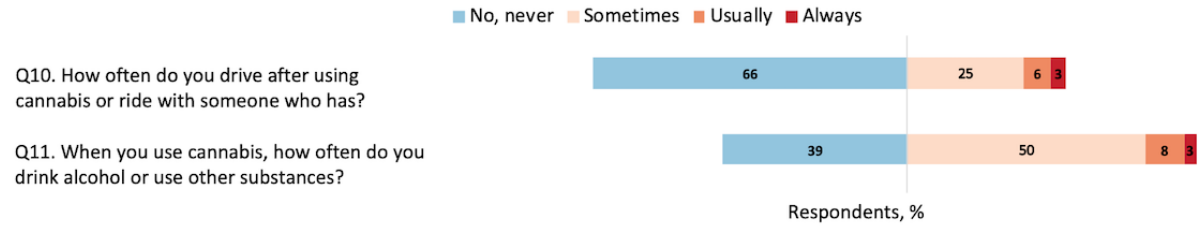


Figure 4. Percentage of driving-age respondents aged 16 to 25 years who reported risky behaviors (n=913).



In our sample, the ASSIST score varied from 0 to 40, with 66.3% (736/1110) of respondents falling in the moderate risk category. The average ASSIST score was 12 among teenaged respondents (aged 12-18 years) and 18 for young adult respondents (aged 19-25 years). ASSIST scores for teenagers were statistically significantly smaller than ASSIST scores for young adults (P<.001).

Table 3 reports the ordered probit coefficient estimates. Male gender (P=.28), other gender (P=.36), and monthly expenditures (P=.07) were not statistically significant at the .05 level. Age (P<.001), polysubstance use sometimes (P<.001) and usually or always (P<.001), ASSIST score (P<.001), and maximum expenditure (P=.02) were statistically significant.

Table 3. Ordered probit estimates for the question "How often do you drive after using cannabis or ride with someone who has?" (n=913).

Covariate	Coefficient estimate[a]	*P* value
Age[b]	0.6586	.04
Age^2	–0.0162	.051
Gender (base=female)		
Male	0.1105	.23
Other gender	–0.2876	.20
ASSIST[c] score	0.0450	<.001
Polysubstance use (base=never)		
Sometimes	0.4221	<.001
Usually or always	0.5942	<.001
Average monthly expenditure	0.0006	.11
Maximum single expenditure	0.0004	.11

[a]The log-likelihood for this model was –640.1 and the pseudo-R^2 was 0.1569.

[b]Age and Age^2 are jointly significant with a *P* value of .10.

[c]ASSIST: Alcohol, Smoking, and Substance Involvement Screening Test.

Table 4 reports the average marginal effects for each of the outcomes of the driving after cannabis use variable. In Table 4, ASSIST scores and polysubstance use have the largest association with driving after using cannabis. The model predicts that every 10-point increase in the ASSIST score increases the probability of sometimes driving after cannabis use by 7.3% (*P*<.001) and increases the probability of usually or always driving after using cannabis by 5.8% (*P*<.001).

Relative to respondents who reported never using other substances with cannabis, respondents who reported sometimes using other substances with cannabis were 13% (*P*<.001) more likely to report driving after using cannabis. Those who reported usually or always using other substances were 18% (*P*<.001) more likely to report driving after using cannabis. This suggests that using alcohol and other substances with cannabis increases the probability of driving after cannabis use, consistent with other studies in the literature [8].

Table 4. Average marginal effects for ordered probit outcomes (n=913).

Covariate	"How often do you drive after using cannabis or ride with someone who has?"					
	Never		Sometimes		Usually or always	
	Effect	*P* value	Effect	*P* value	Effect	*P* value
Age[a]	–0.0253	.03	0.0152	.03	0.0101	.03
Gender (base=female)						
Male	–0.0327	.23	0.0182	.23	0.0144	.22
Other gender	0.0781	.17	–0.0482	.20	–0.0299	.14
ASSIST[b] score	–0.0132	<.001	0.0073	<.001	0.0058	<.001
Polysubstance use (base=never)						
Sometimes	–0.1245	<.001	0.0767	<.001	0.0478	<.001
Usually or always	–0.1811	<.001	0.1062	<.001	0.0750	.001
Average monthly cannabis expenditure	–0.0002	.11	0.0001	.11	0.0001	.11
Maximum single cannabis expenditure	–0.0001	.11	0.0001	.11	0.0001	.11

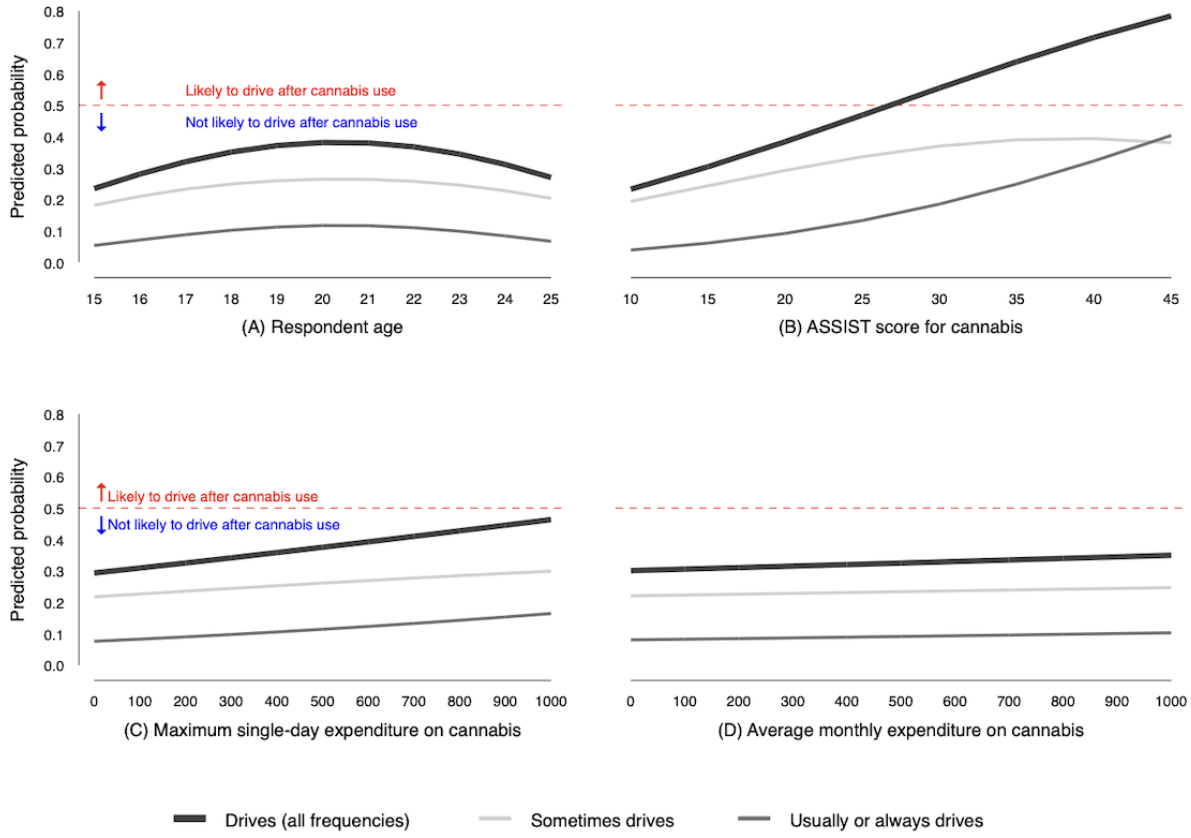
[a]The marginal effect on age includes the full quadratic effect of age on the probability of driving after cannabis use.

[b]ASSIST: Alcohol, Smoking, and Substance Involvement Screening Test.

Figure 5 plots the predicted probability of driving after using cannabis at varying levels of age, ASSIST scores, maximum expenditures, and average expenditures. This figure plots the predicted probability by respondents who sometimes drive and those who usually or always drive after cannabis use and the overall predicted probability of driving after cannabis use (ie, the probability of sometimes driving plus the probability of usually or always driving). In Figure 5, we explored how the predicted probability of driving after cannabis use changes with respondent characteristics. If the predicted probability exceeds

0.5, the model predicts that respondents are more likely than not to drive after cannabis use, indicated with a red dashed line in the figure.

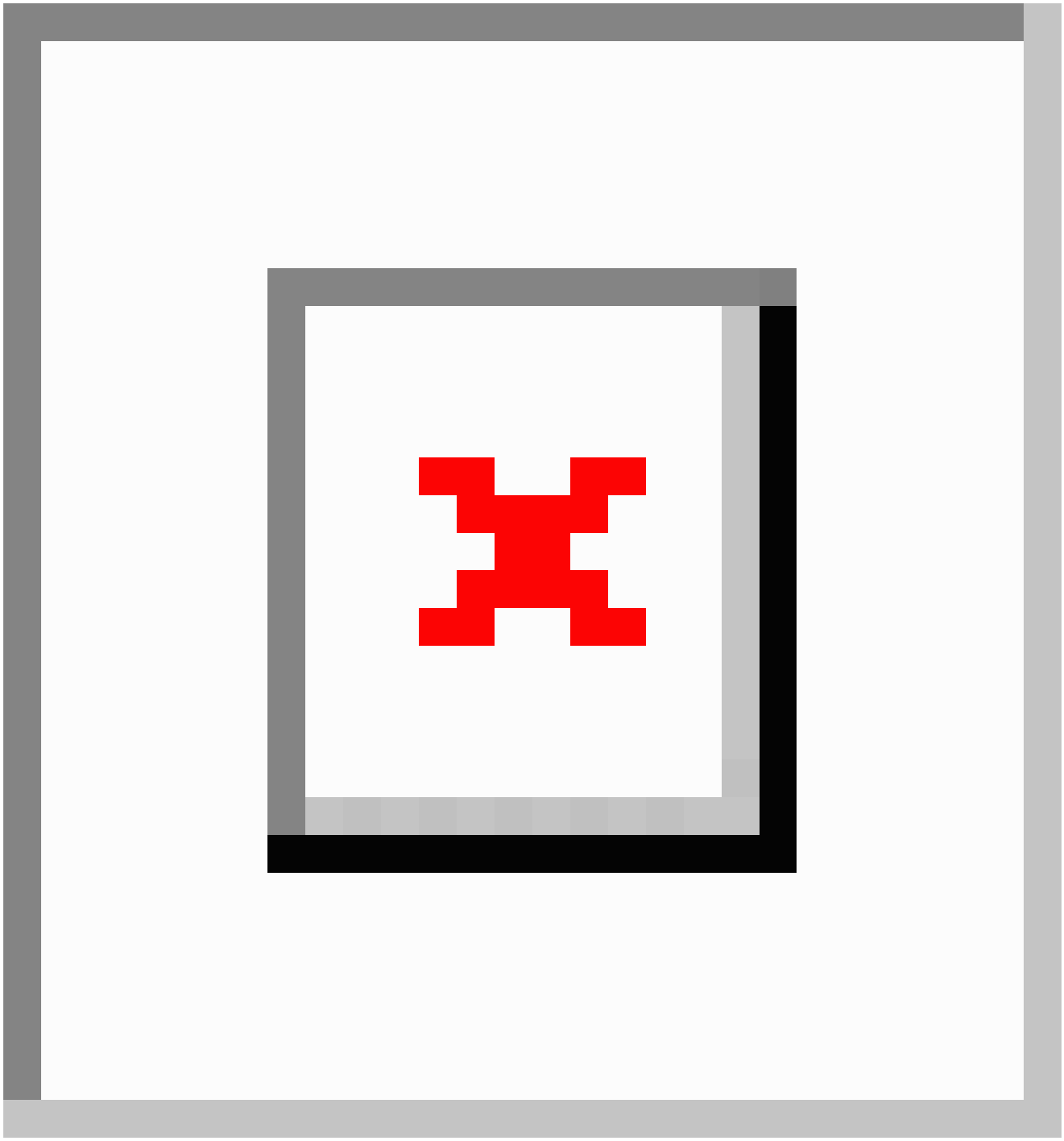
Figure 5. Probability of driving after cannabis use is low over a range of respondent characteristics (n=913). ASSIST: Alcohol, Smoking, and Substance Involvement Screening Test.



Panel A of Figure 5 shows that holding constant gender, ASSIST score, and cannabis expenditures, the overall probability of driving after cannabis use was 0.28 for the youngest respondents in the sample, peaked at 0.35 for 20-year-olds, and dropped to approximately 0.20 for the oldest respondents. Panel B shows that the overall predicted probability of driving after using cannabis, all else constant, increased with ASSIST score, suggesting that respondents with ASSIST scores greater than 25 are more likely than not to drive after using cannabis. As seen in panel C, the predicted probability of driving after cannabis use increased with maximum single-day expenditures. However, the model predicted that even the respondents with the highest expenditures, those most likely to be the heaviest users, are unlikely to drive after cannabis use (ie, the predicted probability of driving after cannabis use is less than 0.5). Similarly, panel D shows that the predicted probability of driving after cannabis use increased with average monthly expenditures, holding all other factors constant. Only at extreme monthly expenditures (over Can $800 [US $630.09]) did the model predict that respondents are more likely than not to drive after cannabis use.

Figure 6 plots the probability of responding "usually" or "always" drives after cannabis use by gender. We found that gender was not a statistically significant factor in determining the probability of driving after using cannabis after controlling for age, ASSIST score, polysubstance use, and expenditures on cannabis. The figure plots the 95% CI for the probability of usually or always driving after cannabis use for male respondents and shows that the probability of usually or always driving after cannabis use for female respondents is within the 95% CI for men at all ages. Respondents younger than 24 years who reported being of another gender have a slightly lower predicted probability of usually or always driving after cannabis use than male respondents.

Figure 6. Probability of reporting usually or always driving after cannabis does not differ by gender (n=913).



Our results show that except for individuals with high ASSIST scores, after controlling for age, gender, and polysubstance use, respondents in our sample were unlikely to drive after using cannabis. Contrary to the literature [31,32], we found that gender is not a statistically significant factor in determining the probability of driving after using cannabis. We also found that polysubstance use and ASSIST scores were strongly correlated with driving after using cannabis. While the level of cannabis expenditures was statistically significant, the effects on driving after cannabis use were small.

Discussion

Principal Findings

The brief intervention questionnaire sheds light on self-reported harmful cannabis use behaviors. We observed that the vast majority of respondents reported cannabis use; however, they also reported not experiencing problems with cannabis use. However, daily cannabis users did report experiencing issues resulting from their use of cannabis. More respondents reported having family or others express concerns about their cannabis use than reported their own concerns about cannabis use. This may suggest that while respondents may not acknowledge problems with their cannabis use, they report an awareness of others' concerns about their use of cannabis.

At an implementation cost of Can $1000 (US $787.62), the digital health tool cost PAD approximately Can $0.90 (US $0.71) per use. Due to the unlimited use structure, the per-use cost would further decrease with PAD's target population's increased use. The digital brief intervention is an example of a low-cost tool that public health campaigns can leverage to tailor resources toward the most at-risk populations. While we focused on cannabis-impaired driving concerns, the tool is flexible and customizable to explore specific cannabis use concerns.

Strengths and Limitations

This study provides insights into how young cannabis users perceive their cannabis use and examines the relationship between cannabis use and driving behaviors from their perspectives. A particular strength is that the CYC intervention has limited use barriers, is anonymous, and was disseminated by a well-known youth advocacy group (PAD). However, as is common with self-reported digital health interventions, there is no way to guarantee that user responses are accurate or honest. Nevertheless, we have shown that a simple, low-cost digital health tool such as the CYC can provide insights to guide cannabis education programs and assist policy makers and youth health advocates targeting efforts for preventing cannabis abuse specific to a community.

The population analyzed in this study can be described as self-seeking and may be problematic cannabis users who actively sought help or more information on their behavior. As a result, they may not be considered representative of the general population. However, if this is the case, it may add strength to our principal finding that the largest amount spent on any given day, higher ASSIST scores, and polysubstance use were positively and significantly associated with driving under the influence of cannabis. As a result, these preliminary results should be interpreted with caution, and there is a need for replication studies that may or may not confirm our results.

Future Research Directions

The CYC has recently been enhanced to include questions addressing the source of cannabis acquisition (primarily dispensaries or licensed services versus informal connections) as well as employment status and the user's geographic location. These additional questions are designed to generate further insight into cannabis use patterns and behavior. Geospatial analysis will allow us to gather further insights and compare use patterns across specific jurisdictions. The CYC has also been modified to include users' first 3 characters in their zip or postal code, and ongoing research is focusing on geospatial analysis to help assess regional patterns.

Based on the ability of digital health interventions to collect self-reported demographic and behavioral data and compare and contrast these data for specific geographic areas, there is the potential for these tools to examine associations between cannabis use and driving at a regional level. This may give further insight on potential predictors of increased risk, which could support or validate findings from other, nondigital studies.

Conclusion

To our knowledge, this is the first study to examine associations between self-reported cannabis use and driving behaviors through the use of a digital brief intervention, which has the dual purpose of educating cannabis users and collecting data to help inform and shape responsible cannabis use programs. Our analysis indicates that the largest amount spent on any given day, higher ASSIST scores, age, and polysubstance use were positively and significantly associated with driving under the influence of cannabis. Gender was not a significant factor.

We have shown how a low-cost digital health tool can inform programs and policies for educating young people about responsible cannabis use. Based on these results, public health campaigns or other interventions may have a more significant impact if they focus resources on problematic cannabis users rather than the general population. The largest amount spent variable may give insight into those who purchase their cannabis from nonretail sources. Further research is required.

Conflicts of Interest

TvM is the CEO and founder of Evolution Health Systems. Evolution Health owns and manages digital health interventions, including the application analyzed in this study. No financial compensation was provided for his participation in this study. GM has no conflicts of interest to declare.

Multimedia Appendix 1
CYC Interface.
[PDF File (Adobe PDF File), 90 KB - formative_v5i3e25583_app1.pdf]

Multimedia Appendix 2
CYC Sample Report.
[PDF File (Adobe PDF File), 1048 KB - formative_v5i3e25583_app2.pdf]

References

1. Jones JM. In U.S., Medical Aid Top Reason Why Legal Marijuana Favored. Gallup News. 2019 Jun 12. URL: https://news.gallup.com/poll/258149/medical-aid-top-reason-why-legal-marijuana-favored.aspx [accessed 2021-02-09]

2. Abbott. Canadians Indicate Growing Concern About Road Safety at One-year Anniversary of Recreational Cannabis Legalization, New Survey Finds. Cision Canada. 2019. URL: https://tinyurl.com/26jzxdds [accessed 2021-02-20]
3. Arterberry BJ, Treloar HR, Smith AE, Martens MP, Pedersen SL, McCarthy DM. Marijuana use, driving, and related cognitions. Psychol Addict Behav 2013 Sep;27(3):854-860 [FREE Full text] [doi: 10.1037/a0030877] [Medline: 23276319]
4. Robertson RD, Woods-Fry H, Morris K. Traffic Injury Research Foundation. Ottawa, Ontario: Traffic Injury Research Foundation; 2016. URL: https://tirf.ca/wp-content/uploads/2016/12/TIRF_DruggedDriving_Policy-Challenges_13_published.pdf [accessed 2020-12-27]
5. Borodovsky JT, Marsch LA, Scherer EA, Grucza RA, Hasin DS, Budney AJ. Perceived safety of cannabis intoxication predicts frequency of driving while intoxicated. Prev Med 2020 Feb;131:105956 [FREE Full text] [doi: 10.1016/j.ypmed.2019.105956] [Medline: 31863787]
6. Brands B, Mann RE, Wickens CM, Sproule B, Stoduto G, Sayer GS, et al. Acute and residual effects of smoked cannabis: Impact on driving speed and lateral control, heart rate, and self-reported drug effects. Drug Alcohol Depend 2019 Dec 01;205:107641. [doi: 10.1016/j.drugalcdep.2019.107641] [Medline: 31678833]
7. Hartman R, Huestis M. Cannabis effects on driving skills. Clin Chem 2013 Mar;59(3):478-492 [FREE Full text] [doi: 10.1373/clinchem.2012.194381] [Medline: 23220273]
8. Kleiman M, Jones T, Miller C, Halperin R. Driving While Stoned: Issues and Policy Options. J Drug Policy Anal 2018;11(2) [FREE Full text] [doi: 10.1515/jdpa-2018-0004]
9. Mann RE, Stoduto G, Ialomiteanu A, Asbridge M, Smart RG, Wickens CM. Self-reported collision risk associated with cannabis use and driving after cannabis use among Ontario adults. Traffic Inj Prev 2010 Apr;11(2):115-122. [doi: 10.1080/15389580903536704] [Medline: 20373229]
10. Cook S, Shank D, Bruno T, Turner NE, Mann RE. Self-reported driving under the influence of alcohol and cannabis among Ontario students: Associations with graduated licensing, risk taking, and substance abuse. Traffic Inj Prev 2017 Jul 04;18(5):449-455. [doi: 10.1080/15389588.2016.1149169] [Medline: 28095034]
11. Hall W. How should we respond to cannabis-impaired driving? Drug Alcohol Rev 2018 Jan;37(1):3-5. [doi: 10.1111/dar.12651] [Medline: 29345082]
12. Wadsworth E, Hammond D. Addict Behav 2019 Mar;90:171-175 [FREE Full text] [doi: 10.1016/j.addbeh.2018.10.050] [Medline: 30412908]
13. Cunningham JA, van Mierlo T. The check your cannabis screener: a new online personalized feedback tool. Open Med Inform J 2009 May 07;3(1):27-31 [FREE Full text] [doi: 10.2174/1874431100903010027] [Medline: 19587808]
14. Cunningham J, Wild T, Cordingley J, van Mierlo T, Humphreys K. A randomized controlled trial of an internet-based intervention for alcohol abusers. Addiction 2009 Dec;104(12):2023-2032 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02726.x] [Medline: 19922569]
15. Cunningham JA, Koski-Jännes A, Wild TC, Cordingley J. Treating alcohol problems with self-help materials: a population study. J Stud Alcohol 2002 Nov;63(6):649-654. [doi: 10.15288/jsa.2002.63.649] [Medline: 12529064]
16. Cunningham JA, Godinho A, Kushnir V. Can Amazon's Mechanical Turk be used to recruit participants for internet intervention trials? A pilot study involving a randomized controlled trial of a brief online intervention for hazardous alcohol use. Internet Interv 2017 Dec;10:12-16 [FREE Full text] [doi: 10.1016/j.invent.2017.08.005] [Medline: 30135748]
17. Cunningham JA, Murphy M, Hendershot CS. Treatment dismantling pilot study to identify the active ingredients in personalized feedback interventions for hazardous alcohol use: randomized controlled trial. Addict Sci Clin Pract 2014 Dec 10;10(1):1 [FREE Full text] [doi: 10.1186/s13722-014-0022-1] [Medline: 25539597]
18. Doumas DM, Hannah E. Preventing high-risk drinking in youth in the workplace: a web-based normative feedback program. J Subst Abuse Treat 2008 Apr;34(3):263-271. [doi: 10.1016/j.jsat.2007.04.006] [Medline: 17600650]
19. Doumas DM, McKinley LL, Book P. Evaluation of two Web-based alcohol interventions for mandated college students. J Subst Abuse Treat 2009 Jan;36(1):65-74. [doi: 10.1016/j.jsat.2008.05.009] [Medline: 18657941]
20. Doumas DM, Haustveit T, Coll KM. Reducing Heavy Drinking Among First Year Intercollegiate Athletes: A Randomized Controlled Trial of Web-Based Normative Feedback. J Appl Sport Psychol 2010 Jul 23;22(3):247-261. [doi: 10.1080/10413201003666454]
21. Check Your Cannabis Screener. Evolution Health. URL: http://www.CheckYourCannabis.net [accessed 2021-02-21]
22. Evolution Health Check Your Screener Questionnaire. Evolution Health. URL: https://evolutionhealth.care/cyc [accessed 2021-02-21]
23. What's With Weed Questionnaire. Evolution Health Care. URL: https://whatswithweed.evolutionhealth.care [accessed 2021-02-21]
24. Shorten A, Moorley C. Selecting the sample. Evid Based Nurs 2014 Apr;17(2):32-33. [doi: 10.1136/eb-2014-101747] [Medline: 24561511]
25. Personal Information Protection and Electronic Documents Act. Office of the Privacy Commissioner of Canada. 2000. URL: https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/ [accessed 2021-02-08]
26. Summary of the HIPAA Privacy Rule. US Department of Health & Human Services. 2003. URL: https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html [accessed 2021-02-09]

27. Directive 2002/58/E.C. of the European Parliament and of the Council of 12 July 2002 Concerning the Processing of Personal Data and the Protection of in the Electronic Communications Sector (Directive on Privacy and Electronic Communications). European Parliament and the Council of the European Union. 2002. URL: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32002L0058 [accessed 2021-02-09]
28. Humeniuk R, Henry-Edwards S, Ali R, Poznyak V, Monteiro MG. The Alcohol, Smoking and Substance involvement Screening Test (ASSIST) : manual for use in primary care. World Health Organization. 2010. URL: https://apps.who.int/iris/handle/10665/44320 [accessed 2021-02-09]
29. Greene WH, Hensher DA. Modeling ordered choices: A primer. Cambridge, England: Cambridge University Press; 2010.
30. Stata Statistical Software: Release 16. In: StataCorp. College Station, TX: StataCorp; 2019.
31. Carliner H, Mauro PM, Brown QL, Shmulewitz D, Rahim-Juwel R, Sarvet AL, et al. The widening gender gap in marijuana use prevalence in the U.S. during a period of economic change, 2002-2014. Drug Alcohol Depend 2017 Jan 01;170:51-58 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.10.042] [Medline: 27875801]
32. Brubacher J, Chan H, Staples J. Cannabis-impaired driving and Canadian youth. Paediatr Child Health 2020 Jun;25(Suppl 1):S21-S25. [doi: 10.1093/pch/pxaa017] [Medline: 32581627]

Abbreviations

ASSIST: Alcohol, Smoking, and Substance Involvement Screening Test
CYC: Control Your Cannabis
CYD: Control Your Drinking
PAD: Parent Action on Drugs
WHO: World Health Organization

Edited by G Eysenbach; submitted 07.11.20; peer-reviewed by S Moore; comments to author 27.11.20; revised version received 07.01.21; accepted 29.01.21; published 02.03.21.

Please cite as:
Moreno G, van Mierlo T
A Digital Health Tool to Understand and Prevent Cannabis-Impaired Driving Among Youth: A Cross-sectional Study of Responses to a Brief Intervention for Cannabis Use
JMIR Form Res 2021;5(3):e25583
URL: https://formative.jmir.org/2021/3/e25583
doi:10.2196/25583
PMID:33650982

Original Paper

Acceptability, Safety, and Resonance of the Pilot Digital Suicide Prevention Campaign "Better Off With You": Qualitative Study

Elise Rose Carrotte[1], BPsych (Hons), MPsych; Marianne Webb[1], PhD; Anna Flego[1], BPhysio, MPH; Bonnie Vincent[1], BA; Jack Heath[1], BA (Hons), LLB; Michelle Blanchard[1,2], GradDip, Dip, BA (Hons), MAICD, PhD

[1]Anne Deveson Research Centre, SANE Australia, South Melbourne, Australia

[2]Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Australia

Corresponding Author:
Elise Rose Carrotte, BPsych (Hons), MPsych
Anne Deveson Research Centre
SANE Australia
PO Box 226
South Melbourne, 3205
Australia
Phone: 61 3 9682 5933
Email: elise.carrotte@sane.org

Abstract

Background: The Interpersonal Theory of Suicide posits that there are three key elements of suicidal behavior: perceived burdensomeness, thwarted belongingness, and the acquired capability for suicide. The digital campaign *Better Off With You* was developed to directly challenge the idea of perceived burdensomeness among people who are contemplating suicide in 2 communities within Australia.

Objective: The aim of this study is to explore the needs and preferences of people with lived experience of suicidal thoughts and actions to inform the development of *Better Off With You.*

Methods: This study involved a series of focus groups that aimed to discuss campaign messaging, scope, and approach. People with lived experience of suicidal thoughts and actions attended the focus groups. After the completion of initial focus groups, the results informed the creation of campaign collateral by creative agencies. Early versions of the campaign collateral were then presented in the user testing sessions. Transcriptions were analyzed via thematic analysis.

Results: In total, 13 participants attended the focus groups and 14 attended the user testing sessions. The following three overarching themes were presented: *acceptability*, *safety*, and *resonance*. Participants believed that suicide is a serious and ongoing issue in their communities and welcomed a localized suicide prevention focus via peer-to-peer storytelling. The idea of perceived burdensomeness required clarification but was perceived as acceptable and relevant. Participants seemed drawn toward peer narratives that they perceived to be authentic, genuine, and believable as given by real people with lived experience. Campaign messaging needs to be clear and empathetic while directly talking about suicide. Participants did not anticipate any significant negative or harmful impact from any campaign videos and highlighted the importance of providing appropriate help-seeking information.

Conclusions: This iterative study provided important insights and knowledge about peer-to-peer storytelling in suicide prevention campaigns. Future campaigns should involve simple messaging, be validating and empathetic, and consider including a lived experience perspective.

(JMIR Form Res 2021;5(3):e23892) doi:10.2196/23892

KEYWORDS

suicide; interpersonal theory of suicide; social media; co-design; lived experience

Introduction

Background

In 2016, there were approximately 800,000 deaths by suicide globally [1]. In Australia, approximately 2.1 million people aged 16 to 85 years have thought about taking their own lives at some point [2]. There are a number of theories that provide various insights into suicidal thoughts and behavior [3]. To understand antecedents to suicide, one of these key theories—the Interpersonal Theory of Suicide—posits that there are 3 key elements of suicidal behavior: perceived burdensomeness, thwarted belongingness, and the acquired capability for suicide [4]. Both perceived burdensomeness and thwarted belongingness are significantly correlated with suicidal thoughts and actions [5].

Digital platforms and social media are increasingly being used to drive health promotion campaigns [6,7]. Indeed, social media has a range of potential benefits for suicide prevention, further providing an anonymous, accessible, and nonjudgmental forum for sharing experiences [8]. However, only a few studies have reported on the development of a safe campaign using social media for suicide prevention purposes [8,9]. To date, only one known campaign has published findings with regard to this approach, which was focused on youth [10]. Critically, there is some concern that the misuse of social media could possibly lead to the risk of contagion [11], where suicidal behavior, thoughts, or deaths spread through a group or community. However, following appropriate media guidelines and designing campaigns with people with lived experience is likely to mitigate these risks [10] alongside appropriate social media moderation. In fact, aligned with the Papageno effect, media can play an important, protective role in sharing alternative approaches to suicide, such as through suicide prevention campaigns [12]. Suicide prevention may involve various messages, including encouraging those at risk of suicide to access the appropriate mental health support, connect with others, and challenge their suicidal thoughts to reduce morbidity and mortality [13-16].

Better Off With You: A Suicide Prevention Initiative

In response to this need, *Better Off With You* was developed, a digital suicide prevention campaign underpinned by the Interpersonal Theory of Suicide [4]. The campaign specifically aims to challenge perceived burdensomeness: the common schema experienced by people contemplating suicide that they are a burden on their family, friends, and other people. Perceived burdensomeness may relate to many areas, including worry about being an emotional or financial burden on others. Such a schema may lead people to believe that their family, friends, and community are *better off without them* or that their death would be a relief to those around them. Perceived burdensomeness was chosen as a key focus of the campaign, as several reviews have identified statistically significant relationships between perceived burdensomeness, suicidal thoughts, and suicide attempts across a range of populations [5,17,18]. In fact, some studies have found perceived burdensomeness to be a stronger predictor of suicidal behavior than thwarted belongingness [18]. Furthermore, no known campaigns have focused on this schema, meaning that it is a novel focus for such a campaign.

Although digital suicide prevention campaigns currently exist in Australia (eg, R U OK?, #YouCanTalk campaigns, [19,20]), these campaigns generally target thwarted belongingness rather than perceived burdensomeness and typically focus on educating the community to reach out to those at suicide risk. In contrast, *Better Off With You* directly targets people who may be contemplating suicide. The campaign is delivered through video stories featuring people with lived experience of suicidal thoughts and actions, encouraging people to challenge their cognitions and visit a campaign website for more information about the campaign, burdensomeness, and help seeking [21]. This focus on people with lived experience is important, as previous evidence has found that peer-led interventions can significantly impact areas such as quality of life, decreasing intensity of suicidal thoughts, and improving a sense of community [22,23].

Better Off With You was developed in partnership with 2 primary health networks (PHNs), which are organizations funded by the Australian Government Department of Health to deliver an efficient and effective primary health care system, including commissioning community-based suicide prevention activities. The 2 PHN partners, Northern Sydney PHN and Northern Queensland PHN, selected targeted areas within their regions to undertake the campaign pilot: the Northern Beaches of Sydney and the Whitsunday, Isaac, and Mackay area of Northern Queensland. These areas were chosen based on the following criteria: areas of greatest need (based on suicide rates and community concern); existing local networks where suicide prevention action planning and community engagement were already underway (including collaboration with local hospital networks, service providers, emergency services, and other community groups); strong local coordination, including capacity of staff within the specific PHN area to be a key contact for supporting campaign implementation; and preparedness of local services to respond to people seeking help in crisis, including emergency department referral pathways and follow-up care after a suicide attempt.

The campaign was developed iteratively with a direct involvement from the local communities and integrating best practice media guidelines for suicide prevention–focused communications. This included having a targeted, clear message with an emphasis on hope [15,16]. In addition, the approach was designed to engage the local community and provide opportunities to be a part of the change, aligned with evidence-informed, community-led suicide prevention approaches [24].

Objectives

The aim of this study is to explore the needs and preferences of people with lived experience of suicidal thoughts and actions to inform the development of a digital peer-to-peer suicide prevention campaign pilot, *Better Off With You*, targeting perceived burdensomeness in 2 Australian communities.

Methods

Design

This formative study was qualitative in nature and involved focus groups to discuss initial campaign messaging, scope, and approach. Focus groups were chosen so that participants could share and reflect on each other's beliefs, opinions, and experiences in the local community. The results informed the creation of the campaign collateral by creative agencies. Early versions of the campaign collateral were then presented in user testing sessions and then further refined before the campaign launch and impact evaluation.

Participants

Focus groups used criterion-based sampling [25]—in this case, people who have experienced suicidal thoughts and actions living within 2 PHN areas. As such, the research included individuals who were most likely to benefit from (or be impacted by) *Better Off With You*.

Targeted recruitment occurred through local health and community organizations that are already engaged in suicide prevention efforts. These organizations were contacted and asked to share a flyer with relevant contacts, such as local suicide prevention champions and advocates within businesses and community organizations. Potential participants were also contacted through social media advertisements. Participants were recontacted for user testing workshops alongside additional contact with local stakeholders and snowballing (asking the existing focus group participants to share a recruitment notice with their networks).

Participants were required to meet the following inclusion criteria: be aged 20 years or older, currently or previously resided in the area, have a relevant lived experience (identified through the yes or no question: *Have you ever experienced a personal crisis or mental health issues relating to suicide [that is, experienced suicidal thoughts or actions yourself]*) but felt well enough to discuss their experiences or speak about this topic in a workshop context, and had not attempted or seriously contemplated suicide in the last 12 months. The inclusion criterion related to lived experience was to ensure that the participants were able to comment on the campaign messaging and materials with respect to their own personal experiences, rather than speaking theoretically or from the perspective of a family member or friend. The inclusion criterion relating to age (≥20 years) was specifically chosen after a review of suicide rates among different demographics in the local communities. This review identified that people aged 20-60 years were particularly at risk and were thus an important target audience of the campaign.

Ethical Considerations

The study was reviewed and approved by Human Research Ethics Committee, Bellberry Limited (2019-03-230). As the study was iterative, the latter stages of the study were reviewed via amendments to the original application. Research was conducted by a multidisciplinary team with expertise in mental health research, suicide prevention, public health, qualitative research methods, and psychology.

Participant safety was prioritized by incorporating strategies used in previous suicide-related research [9]. Participants prepared a wellness plan to support them if they experienced distress during the workshop. A local mental health clinician was present to support participants should they experience distress during the workshop. Participants were requested to follow *safe sharing guidelines* and not disclose the details of suicide attempts in the session (either their own or those that have occurred in their community). Furthermore, participants were provided with crisis service contact details before focus groups through email contact, consent forms, and in follow-up emails within 48 hours of participation.

Measures and Procedure

Focus groups and user testing sessions were attended by 3 female members of the project team: MW and EC, researchers who cofacilitated the discussion and who have experience in qualitative data collection and analysis, and BV, project lead, who provided oversight and support during discussion and took field notes. Sessions ran for 1.5-2 hours and were conducted in spaces such as local clinic meeting rooms and community centers.

Each session commenced with participants completing consent forms and an ice breaker. Each session was audio-recorded, and the participants were reimbursed with a Aus $50 (US $38) gift voucher. Participants were asked to complete feedback forms covering participants' perceptions of the session's safety at the end of each session, whether the session was interesting, whether they felt heard during the session, or whether they felt *Better Off With You* could make a difference in their community. Each of these was rated on a 5-point Likert scale from strongly disagree to strongly agree. Participants were also asked if they were interested in being recontacted for user testing sessions in the coming months.

Content was informed by a discussion guide. Focus groups involved an in-depth exploration of individual experiences, needs, and preferences related to content, format, and channels to assist in designing *Better Off With You*. Discussion questions pertained to memorable health campaigns, important messages around suicide prevention, discussion around perceived burdensomeness and its relevance, and format of the campaign. Region-specific questions were asked, including questions about the impact of suicide on their local community and the types of imagery or messaging that would appeal to the community.

User testing sessions were held around 3 months after the workshops. These involved reviewing the near-final campaign materials to ensure that they met the design requirements identified in the focus groups. Participants were shown 30-second and 3-min versions of the campaign videos featuring local people with lived experience of suicidal thoughts and actions and filmed in local settings. There were 3 video stories per region; each 30-second version was shown and discussed, and one 3-min version was shown and discussed. Discussion questions pertained to immediate reactions to the videos, likes and dislikes, perceived messages in the videos, and perceived appropriateness for the local community. Participants also discussed website design and architecture (*wireframes*).

Data Analysis

Descriptive statistics were calculated for written feedback forms (percentage agreement and means or SDs). Audio files were transcribed verbatim. Transcriptions, field notes, and written feedback were analyzed via thematic analysis using the methodology outlined by Braun and Clarke [26] using NVivo 12 (QSR International). This was a recursive process that was both deductive and inductive in nature. The initial steps involved immersion in the data by 2 authors (MW and EC), via reading, rereading, and listening to recordings. These authors then independently coded each transcript line by line, and met to discuss and generate overarching themes based on patterns identified between the codes. Latter transcripts were then coded using these initial themes as a guide. The authors regularly met to discuss the thematic analysis and to modify and refine the themes, subthemes, and codes.

Results

Overview

The results are presented chronologically below, with each subphase structured according to 3 themes: *acceptability*, *safety*, and *resonance*. These overarching themes were applied on both the focus groups and user testing sessions and are discussed in relevance to each iterative stage. Subthemes, including *clarity of messaging* and *balancing hope and realism*, were discussed where appropriate. Descriptive statistics for feedback forms are presented in Multimedia Appendix 1.

Focus Groups

In total, 13 people with lived experience of suicidal thoughts and actions participated in focus groups in June 2019. The first workshop was held in Northern Queensland and included 6 women aged 27 to 59 years (mean 48.7, SD 11.6). Owing to recruitment difficulties, a second workshop could not be held in this PHN area; instead, 2 individual interviews were conducted with 2 women aged 48 and 64 years. A second workshop was held a fortnight later in New South Wales. This workshop included 3 women and 2 men aged 27-60 years (mean 44.8, SD 19.2). For workshop participants, written feedback (Multimedia Appendix 1) indicated a very high agreement (>90%, 10-11/11) that the workshop was safe, interesting, participants' voices were heard, and the campaign could make a difference in the local communities.

Acceptability

Participants believed that suicide is a serious, ongoing issue in their communities and welcomed a localized suicide prevention focus. However, most participants did not immediately identify perceived burdensomeness as an important schema for someone contemplating suicide. Participants also had difficulty understanding the tagline *Better Off With You* and its relationship to burden without context. Participants tended to initially conceptualize the message as being said by other people (family or friend) to a suicidal person, rather than being said by people with a lived experience. They felt that this perspective would have a negative impact on someone contemplating suicide and that they would feel guilty and ashamed of having suicidal thoughts:

> *I just got the feeling of guilt [from the tagline "Better Off With You"]…that it's not about the person who is suicidal, it's about the other people; we are better off with you. We're not focusing on the person and trying to find out what's going on for you, and that goes into that guilt tripping: "We want you here so don't you dare kill yourself".* [Female participant, New South Wales session]

Interpretation of the tagline was discussed in detail; hence, the first subtheme was identified: *clarity of messaging*. Participants found the tagline to be clearer once they understood the broader messaging of the campaign. Once participants understood that the message would be delivered by someone with a lived experience rather than by a third party and that the message was free from judgment or shame, they generally accepted the tagline. Similarly, participants accepted targeting perceived burdensomeness once this was explained clearly and contextualized, even though participants might not have experienced this schema as a part of their own lived experience with suicidal thoughts and actions:

> *I think [the message is] actually very, very powerful. I think the idea of if you can just get yourself over this point – and there is a point and it is going to change. It may not get better straight away, but if you can just get yourself over this hurdle and out of this rut again maybe you can start helping yourself a little bit.* [Female participant, Queensland session]

Safety

Participants did not identify any component of the campaign that was identified as inherently triggering or unsafe, as long as the messaging was clear and considered. However, they spoke to the nuance of maximizing emotional safety in the campaign—that is, minimizing distress, avoiding critical incidents, and ensuring viewers feel comfortable with the campaign.

The importance of validating people's lived experience through the narratives was highlighted, with participants noting that being suicidal "is a big deal and it is hard" (female participant, New South Wales session). They warned against trivializing experiences with presumptive messaging, such as *you will get over it.*

Participants reported the importance of including help-seeking information as a part of the campaign in the event that a viewer of the campaign becomes distressed or triggered (ie, re-experiencing a traumatic event): "I think if you've going to have a powerful ad like that, at the end you need to have something about if it triggered you, here's the number to call" (female participant, Queensland session). They highlighted the importance of including a range of local, national, and web-based services, including 24/7 services. They reported that the website accompanying the campaign should also include support for carers, families, or friends. Participants also reported that help seeking can be challenging; they suggested encouraging campaign viewers to keep trying even if they did not find the right support immediately.

Resonance

Real stories were perceived as likely to resonate with the campaign's target audience. Participants identified that an authentic and honest tone would have a real impact through voices of those with lived experience of suicidal thoughts and actions and who have walked the path themselves (rather than actors or celebrities):

> *It's actually using people who have been through it and have survived and have turned their lives around...there is nothing worse than listening to some paid actor and it's like "you have no idea…you haven't been in my shoes. I can't have any faith in what you're telling me because you're not even real.*" [Female participant, Queensland session]

It was identified that the campaign was most likely to resonate if the message was simple, without being too simplistic, and did not trivialize experiences; participants highlighted the importance of acknowledging the complexity of issues and of people's lives.

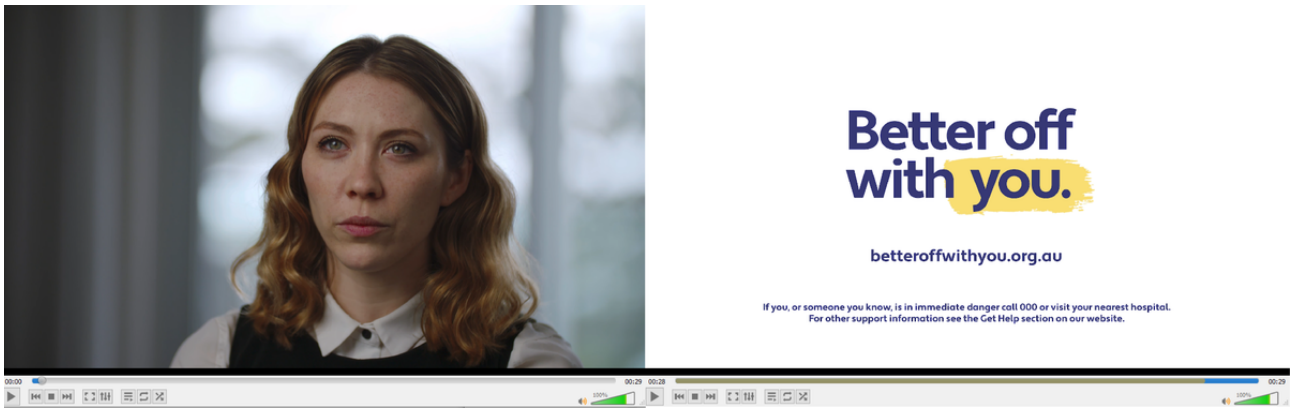
Participants reported that for the campaign to resonate, it must be engaging. Participants recommended that the campaign should have a serious but hopeful tone and use conversational language rather than formal language, which might ostracize some viewers. They identified that the video content should be brief and eye catching to hold the audience's attention, should not be overproduced (ie, should not include overly dramatic music), and should talk directly about suicide rather than just alluding to it.

Generally, participants were enthusiastic about the campaign being localized. They reported that known geographic locations and landmarks could be shown to be locally identifiable. Participants also highlighted the importance of diverse representations in the campaign, thereby targeting a range of audiences, such as people from farming or mining communities, young people who are not engaged in popular local social activities or cultures such as surfing and sport, and people from culturally and linguistically diverse backgrounds.

User Testing

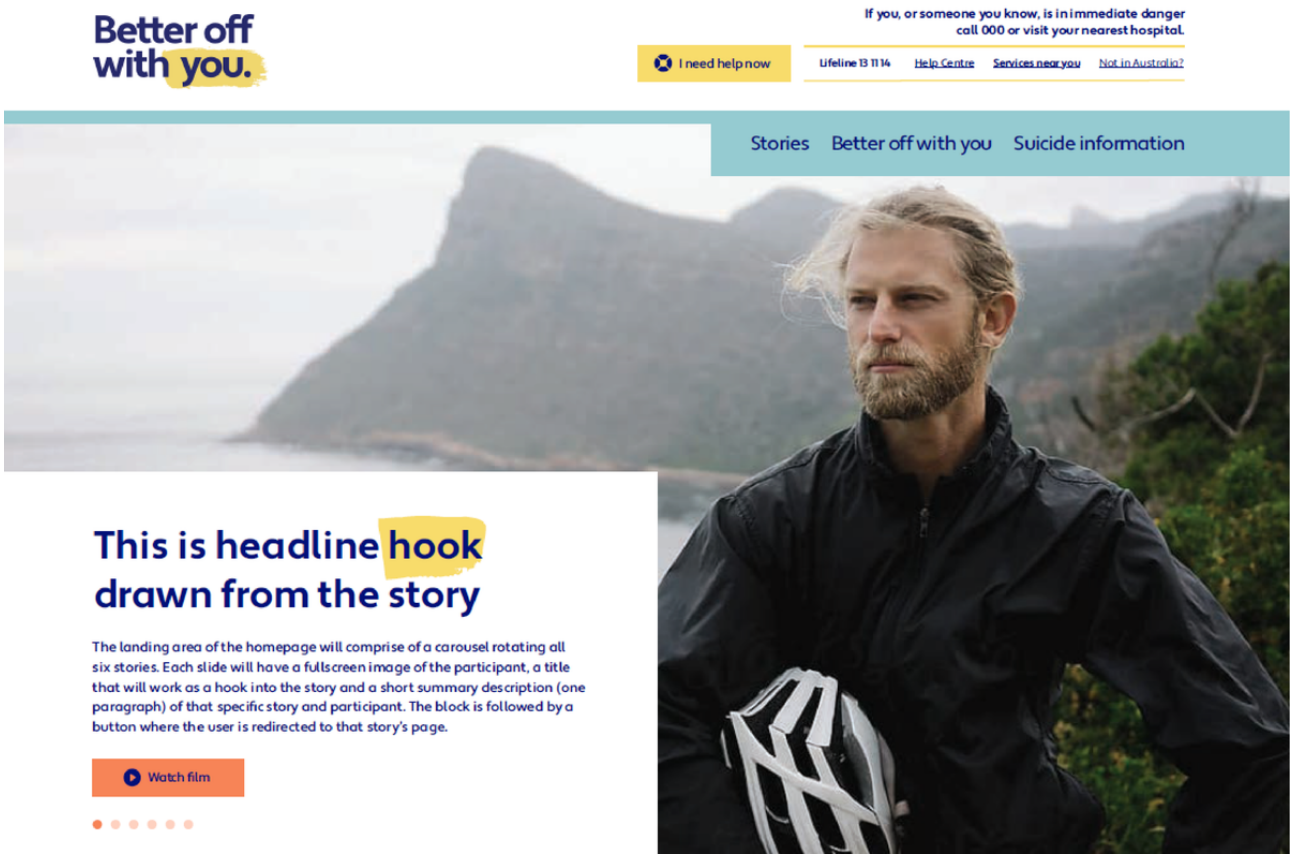
The focus group findings informed the development of the campaign collateral while considering tone, language, length of videos, clarity of messaging, and opportunities for dissemination. Draft collateral was presented to participants in user testing sessions (examples are presented in Figures 1 and 2). Videos featured a person with personal experience of suicidal thoughts and actions discussing events, emotions, and thought processes that preceded or contributed to their suicidal thoughts, how they coped, and what life has been like since the crisis had passed. Thoughts around perceived burdensomeness were highlighted in each story. In each video, the storyteller spoke directly to the camera, with local overlay footage edited throughout the video (eg, the participant walking along the beach, spending time with friends, or walking their dog). The final shot of each video was the *Better Off With You* logo, website link, and crisis support numbers. A 30-second teaser version was produced for each story for social media, designed to draw in the viewer and then direct the viewer to the *Better Off With You* website to watch the longer 3-min version of the video and access support information.

Figure 1. Screenshots of draft Better Off With You videos (video 1/6; Phoebe's story; 30-second version) discussed in user testing sessions.



A total of 3 user testing sessions occurred across September to October 2019: 2 in New South Wales and 1 in Northern Queensland. In total, 14 participants attended these sessions. A total of 10 participants attended in New South Wales, comprising 4 men and 6 women aged 23-68 years (mean 42.2, SD 15.1); 4 participants participated in Northern Queensland, with 1 man and 3 women aged 43-64 years (mean 51.5, SD 9.0). Written feedback (Multimedia Appendix 1) indicated a very high agreement (14/14, 100%) that the workshop was safe and interesting, and participants' voices were heard. Feedback was mixed for the items pertaining to the campaign materials, with the strongest agreement for the videos being engaging (12/14, 86%), likely to encourage help seeking (14/14, 100%), and likely to give viewers a sense of hope for the future (13/14, 93%). The lowest percentage agreement was for the item: *the campaign could distress viewers*, with only 29% (4/14) agreement.

Figure 2. Draft Better Off With You website landing page design discussed in user testing sessions.



Acceptability

Messaging was frequently discussed, both prompted and unprompted by the facilitators. As seen in the focus groups, participants accepted the general message behind *Better Off With You*. The message was generally perceived as fairly clear in the 30-second videos but became diluted during the longer 3-min videos, which involved more of the storyteller's narrative. Participants identified key (but unintended) messages they noticed in the longer videos, which included *there is hope*, *find the right diagnosis*, and *talk about it*. These were not necessarily received as inappropriate messages for a suicide prevention campaign but were not the intended primary focus of *Better Off With You*. Participants felt that the *you are not a burden* message often came through as secondary to these other messages in the 3-min videos, was too subtle, or was lost in the narrative:

> *I found the 30 seconds captured me, whereas this...it was just your classic sort of suicide video. Because I loved the message so much about this burden on family, and it was only really mentioned very briefly...* [Female participant, New South Wales session]

Participants were most accepting of the 30-second videos, generally identifying them as engaging, real, and honest. These clips were very rapid, and some participants commented that they could be simplified slightly to improve comprehension. Regardless, they reported that the message in these short videos was clear and strong. Participants often identified a sense of hope in these videos and praised this tone, for example: "(The video) showed the future, from a dark place...it showed a possible future" (female participant, Queensland session).

Feedback around the acceptability of the longer, 3-minute versions of the videos was more mixed. Although some participants identified these as powerful, "open, and honest" (male participant, New South Wales), others reported that the videos were not as engaging or clear. Creative flow significantly impacted this result, particularly the choice of music and rapid editing of overlay or dialog, which were seen as disjointed, distracting, or inappropriate to the topic at times. Suggestions were made to simplify the editing and alter the music to ensure that the narrative was engaging and the message was clearer.

The website design was generally viewed favorably. Although some felt the design was "a bit clinical" (female participant, Northern Queensland session), the graphic design was generally well received. Participants recommended that the website content have a personal, informal, and direct tone:

> *It's a clean website; it's not cluttered. And it seems to be easy to navigate, and the colour scheme is the same. So I think it's a good one.* [Female participant, New South Wales session]

Safety

Participants did not anticipate any significant negative impact from any campaign videos. A few participants experienced an emotional response to the videos within the sessions, but no participants identified being emotionally triggered or feeling unsafe after seeing the campaign materials.

Participants identified that help-seeking messages in the videos were necessary to ensure that viewers knew where to seek help if they were experiencing suicidal thoughts. However, specific

descriptions of help seeking from the storytellers often brought about comparative responses. A second subtheme was identified through this discussion: *balancing hope and realism*. Participants spoke to the challenge of needing to balance the message of *seek help if you are feeling suicidal* with a realistic assessment of the challenges associated with seeking help in local mental health services. If participants interpreted the storyteller's help-seeking journey as being a straightforward process, then this was met with rejection:

> *The fact that [the storyteller] said, "You can actually talk to your doctor and you won't be judged," but there will be people out there who say, "Well, how would you know? My doctor judged me badly and I'm never going back to this doctor."* [Female participant, New South Wales session]

Many participants identified that they had personally experienced challenges while seeking help for mental health or suicidal thoughts; however, focusing too much on this in the videos could be seen as off-putting and detracting from the intended messages of the campaign. This meant that if participants interpreted the storyteller as being privileged with access to a range of health care professionals and self-care resources, then this interpretation could also be taken in the wrong way:

> *I feel [the message around help seeking] is a little bit risky... It's like, just keep on trying and maybe you'll get there, and you've got to do ballet and see a psychiatrist and see a psychologist and go to the gym, stay away from the screens and ask someone to help you... And even when you do all that, you might not click with the right person.* [Female participant, New South Wales session]

Participants highlighted the need for the website to be easy to navigate, particularly in the event that someone accesses the website in a state of distress or suicidal crisis. They suggested a clear signposting of *Better Off With You* as a suicide prevention initiative, including more direct information about managing thoughts around burdensomeness on the website:

> *Just having something like some practical skills, like how to talk to someone about the burden...so just some questions you could ask. Because that's really what it's about, isn't it? Like there's all different layers of help, but if you do feel that you are a burden to others, just being able to sit down, say with a mother, and say, "These are the feelings I'm having," and what questions can you ask.* [Female participant, Queensland session]

Resonance

Participants seemed drawn to storytellers who were perceived to be authentic, genuine, and honest. The sense of authenticity was enhanced by seeing storytellers express clear emotions (eg, becoming teary when discussing a difficult experience), ensuring that the camera was on the storyteller's face when they were discussing sensitive elements of their stories, and showing storytellers in their own environments. This contributed to some storytellers being viewed as engaging, likable, and *quite authentic* (male participant, New South Wales session).

Participants also commented on the relatability of the participants' stories. Some participants reported relating to the elements of the storytellers' narratives, such as their thought processes. Relatability was enhanced because of some storytellers who do not look like they fit the stereotypes of people experiencing mental health issues or suicidal thoughts, for example: "I think she (the storyteller) looked like someone who you would not imagine would struggle" (male participant, New South Wales session).

Points of contention were identified in the longer videos, which explored more of the storyteller's journey. Aligned with the subtheme of *balancing hope and realism* discussed above, when more specifics of a diagnosis or recovery journey were provided, this tended to invoke a more comparison-based response from the participants, who were likely to comment along the lines of *not everyone can access that support* or *in my experience that did not work*. Similarly, if participants mentioned specific symptoms of mental health issues (such as mania if they experienced bipolar disorder), then the participants tended to find this less relatable if they did not reflect their experience. The more specific details provided, the more likely participants were to focus on the details and lose a sense of the overall intended messages due to perceived irrelevance or being *invalidating* (male participant, New South Wales). Participants tended to respond critically to the instances of advice-giving or blanket reassurance provided by the storytellers:

> *I think for young people [the message] might be a bit too strong...Kind of like a dad saying, "You're not thinking straight if you're thinking about killing yourself." And it's like, "Well, what would you know?"* [Male participant, New South Wales session]

Discussion

Principal Findings

This qualitative, iterative study explored the needs and preferences of people with lived experience of suicidal thoughts and actions for a digital peer-to-peer suicide prevention campaign pilot, *Better Off With You*. We conducted a series of focus groups to discuss initial campaign messaging and approaches. On the basis of this feedback, campaign collateral was then developed before being presented in user testing sessions with local participants with lived experience for further feedback and recommendations. To the best of our knowledge, this is the first study to report on the design and development of a digital suicide prevention campaign focusing on perceived burdensomeness.

The findings of this study highlight the challenges in developing a suicide prevention campaign that successfully conveys a clear, simple message. The intended burdensomeness message did not immediately resonate with participants, with participants also perceiving a variety of other appropriate, positive campaign messages. This was important, as participants highlighted the need for a hopeful tone. Concerns were also raised about the key message being misinterpreted. Even though research has

shown a strong correlation between perceived burdensomeness and suicidal behavior [18], there may be a low awareness in the general public about this relationship. Given that we have not found any other suicide prevention campaigns that focused on perceived burdensomeness, this finding is not surprising. It reinforces the need to have ongoing end user involvement in the development of any suicide prevention campaign to ensure that the intended message is understood by the target audience and communicated in a clear, validating manner. It also highlights the potential role of the media in conveying a sense of hope and alternative pathways in the context of suicidal thoughts and actions.

A major finding of the study was the participants' strong preference to include first-person perspectives in the campaign. This resonance is aligned with a previous research study that suggests that peer-to-peer mental health and suicide programs with relatable, credible lived experience perspectives are an effective way to decrease stigma and increase help seeking [27]. It also reflects a recent research study by Thorn et al [10], who found that young people wanted to see real people in videos and photographs for a social media suicide prevention campaign, rather than professional actors or models. Not only was the peer-to-peer approach seen to be an authentic and acceptable, participants felt that a campaign where the message was delivered by another person, such as a family member or friend, would not be as powerful. Our findings contribute new insights and knowledge about the design elements required for developing effective lived experience suicide prevention videos, as shown in recent studies reporting on the effectiveness of videos targeting farmers [28] and lesbian, gay, bisexual, transgender, intersex, and queer communities [29]. These findings also suggest that a third-person approach, while well intentioned, may not engage the target audience of those contemplating suicide effectively.

Regardless, the study identified the importance of ensuring that these real stories are perceived to be realistic and relatable while acknowledging the inequities of access to appropriate support. For example, participants were critical of the stories they perceived to be downplaying the realities of seeking help for suicidal thoughts and actions; however, they agreed that help-seeking messages were important. This result highlights the need to strike a balance between being true to a person's lived experience while ensuring that their story is told safely and realistically.

Ensuring the material does not have any negative or unexpected impact is critical for any suicide prevention intervention, particularly for campaigns intended for a mass media distribution across a region or community. The findings of this study suggest that it is possible to develop a safe suicide prevention campaign that addresses perceived burdensomeness. Although there are existing evidence-based guidelines for developing suicide prevention campaigns [15,16], this study provides new insights into the nuances required to develop safe suicide prevention messages without alienating the target audience; for example, acknowledging the challenges of seeking help, without discouraging people or suggesting that there is only one *right* way to manage suicidal thoughts. Furthermore, the written feedback and lack of adverse incidents indicate that there were no negative or unexpected impacts on participants because of participating in the study or seeing the draft campaign materials. This supports research indicating that people with lived experience of suicidal thoughts and actions can be safely involved in this type of research [10,30,31].

Research Translation

Findings from this study were communicated to the creative agencies involved in designing the website, video stories, and other campaign collaterals. Edits included changing dialog (including removing some lines that were highlighted by participants as invalidating or challenging), slowing down the pace, removing some of the overlay to simplify the videos, and replacing some of the music cues. Website changes included a clearer branding of the website as a suicide prevention initiative, adding more information about perceived burdensomeness, and various language changes. A future publication will report on the effectiveness of this suicide prevention campaign pilot when delivered via social media.

Limitations

This study has some limitations. As a small pilot conducted in only 2 regions of Australia, the generalizability of the findings may be limited. Notably, each workshop ranged between only 4 to 6 participants. Difficulty with recruitment is common in suicide-related research [9], and it was difficult to find people in the community who fit the inclusion criteria and were willing and able to participate within a short recruitment timeline. Recruitment was restricted to specific geological areas and budget or overarching project timeline constraints, likely impacting upon the participation rates.

Younger participants and men were underrepresented, which is not unusual in health research [32]. However, with an increased buy-in from local stakeholders to help target local men more directly, we were able to improve the gender ratio for the user testing sessions, with 36% (5/14) being male participants. Future research should expand on this work to improve the generalizability to the specific needs of populations who are overrepresented in suicide statistics, including men, Aboriginal and Torres Strait Islanders, those from cultural and linguistically diverse backgrounds, and those with diverse sexual and gender identities.

Despite the small sample and recruitment challenges, the feedback was rich. Indeed, given the sensitive nature of the topic, a smaller and more intimate size may have facilitated a more meaningful contribution in each session with the opportunity to discuss ideas in greater depth in smaller groups.

Finally, it is possible that using direct payment incentivized participants to sign up for the study who may not have had sufficient experience or local expertise; however, considering the depth of the feedback, the authors do not believe this impacted upon the quality of the data.

Conclusions

This iterative study provided important new insights and knowledge about how the contribution of people with lived experience of suicidal thoughts and actions is essential for the development of an acceptable, safe, and resonant suicide

prevention campaign. The messaging must be clear, simple, and validating, and the inclusion of a lived experience perspective is particularly valuable. This feedback was critical for developing the *Better Off With You* pilot, which has now been rolled out. Subsequently, opportunities for expanding the campaign are being explored. Impact evaluation will be discussed in a future publication.

Acknowledgments

The authors wish to thank all the participants who provided valuable contributions to this study. The authors acknowledge the support and contribution from the Northern Sydney PHN and Northern Queensland PHN, HotGlue and ThinkHQ. They also wish to thank Pauline Kenny for her contribution to the project.

Better Off With You is a pilot initiative delivered by SANE Australia and is supported by funding from the Australian Government Department of Health and the John & Myriam Wylie Foundation.

Conflicts of Interest

JH is the director of SANE Australia and The Dax Centre. MB is a nonexecutive director of youth mental health charity, batyr Australia; this paper is unrelated to this role. The other authors have no conflicts to declare.

Multimedia Appendix 1
Feedback form summary.
[DOCX File , 18 KB - formative_v5i3e23892_app1.docx]

References

1. World Health Statistics 2019: Monitoring Health for the SDGs, Sustainable Development Goals. World Health Organization. 2019. URL: https://www.who.int/gho/publications/world_health_statistics/2019/en/ [accessed 2021-02-09]
2. Slade T, Johnston A, Teesson M, Whiteford H, Burgess P, Pirkis J, et al. The Mental Health of Australians 2: Report on the 2007 National Survey of Mental Health and Wellbeing. Canberra; Australia. URL: https://www1.health.gov.au/internet/main/publishing.nsf/Content/mental-pubs-m-mhaust2 [accessed 2021-02-09]
3. Klonsky ED. The role of theory for understanding and preventing suicide (but not predicting it): A commentary on Hjelmeland and Knizek. Death Studies 2019 Apr 15;44(7):459-462. [doi: 10.1080/07481187.2019.1594005]
4. Van Orden KA, Witte TK, Cukrowicz KC, Braithwaite SR, Selby EA, Joiner TE. The interpersonal theory of suicide. Psychol Rev 2010 Apr;117(2):575-600 [FREE Full text] [doi: 10.1037/a0018697] [Medline: 20438238]
5. Chu C, Buchman-Schmitt JM, Stanley IH, Hom MA, Tucker RP, Hagan CR, et al. The interpersonal theory of suicide: A systematic review and meta-analysis of a decade of cross-national research. Psychological Bulletin 2017 Dec;143(12):1313-1345. [doi: 10.1037/bul0000123]
6. Lim MSC, Wright C, Hellard ME. The medium and the message: fitting sound health promotion methodology into 160 characters. JMIR mHealth uHealth 2014 Nov 03;2(4):e40. [doi: 10.2196/mhealth.3888]
7. Lim M, Wright C, Carrotte E, Pedrana A. Reach, engagement, and effectiveness: a systematic review of evaluation methodologies used in health promotion via social networking sites. Health Promot J Austr 2016 Oct 06;27(3):187-197. [doi: 10.1071/he16057]
8. Robinson J, Cox G, Bailey E, Hetrick S, Rodrigues M, Fisher S, et al. Social media and suicide prevention: a systematic review. Early Interv Psychiatry 2016 Apr;10(2):103-121. [doi: 10.1111/eip.12229] [Medline: 25702826]
9. Bailey E, Mühlmann C, Rice S, Nedeljkovic M, Alvarez-Jimenez M, Sander L, et al. Ethical issues and practical barriers in internet-based suicide prevention research: a review and investigator survey. BMC Med Ethics 2020 May 13;21(1). [doi: 10.1186/s12910-020-00479-1]
10. Thorn P, Hill N, Lamblin M, Teh Z, Battersby-Coulter R, Rice S, et al. Co-designing a suicide prevention social media campaign with young people: The #chatsafe project. JMIR Ment Heal 2019;7. [doi: 10.2196/preprints.17520]
11. Luxton DD, June JD, Fairall JM. Social media and suicide: a public health perspective. Am J Public Health 2012 May;102 Suppl 2:S195-S200. [doi: 10.2105/AJPH.2011.300608] [Medline: 22401525]
12. Sisask M, Värnik A. Media roles in suicide prevention: a systematic review. IJERPH 2012 Jan 04;9(12):123-138. [doi: 10.3390/ijerph9010123]
13. Hom MA, Stanley IH, Joiner TE. Evaluating factors and interventions that influence help-seeking and mental health service utilization among suicidal individuals: A review of the literature. Clin Psychol Rev 2015 Aug;40:28-39. [doi: 10.1016/j.cpr.2015.05.006] [Medline: 26048165]
14. McLaughlin C, McGowan I, Kernohan G, O'Neill S. The unmet support needs of family members caring for a suicidal person. Journal of Mental Health 2015 Dec 11;25(3):212-216. [doi: 10.3109/09638237.2015.1101421]
15. Acosta J, Ramchand R, Becker A. Best practices for suicide prevention messaging and evaluating California's 'Know the Signs' media campaign. Crisis 2017 Feb 23:1-13. [doi: 10.1027/0227-5910/a000446] [Medline: 28228062]

16. National Action Alliance for Suicide Prevention. Actional alliance framework for successful messaging. 2018. URL: http://suicidepreventionmessaging.org/ [accessed 2021-02-09]
17. Hill RM, Pettit JW. Perceived burdensomeness and suicide-related behaviors in clinical samples: current evidence and future directions. J Clin Psychol 2014 Jul 13;70(7):631-643. [doi: 10.1002/jclp.22071] [Medline: 24421035]
18. Ma J, Batterham PJ, Calear AL, Han J. A systematic review of the predictions of the Interpersonal–Psychological Theory of Suicidal Behavior. Clinical Psychology Review 2016 Jun;46:34-45. [doi: 10.1016/j.cpr.2016.04.008]
19. R U OK? Foundation. R U OK? campaign website. 2020. URL: https://www.ruok.org.au/ [accessed 2021-02-09]
20. Life in Mind Australia. #YouCanTalk campaign website. 2020. URL: https://lifeinmind.org.au/youcantalk [accessed 2021-02-09]
21. SANE Australia. Better Off With You campaign website. 2020. URL: https://betteroffwithyou.org.au/ [accessed 2021-02-09]
22. Corrigan PW, Sokol KA, Rüsch N. The impact of self-stigma and mutual help programs on the quality of life of people with serious mental illnesses. Community Ment Health J 2013 Feb;49(1):1-6 [FREE Full text] [doi: 10.1007/s10597-011-9445-2] [Medline: 22038373]
23. Schlichthorst M, Ozols I, Reifels L, Morgan A. Lived experience peer support programs for suicide prevention: a systematic scoping review. Int J Ment Health Syst 2020 Aug 12;14(1). [doi: 10.1186/s13033-020-00396-1]
24. Shand F, Torok M, Cockayne N, Batterham PJ, Calear AL, Mackinnon A, et al. Protocol for a stepped-wedge, cluster randomized controlled trial of the LifeSpan suicide prevention trial in four communities in New South Wales, Australia. Trials 2020 Apr 15;21(1):332 [FREE Full text] [doi: 10.1186/s13063-020-04262-w] [Medline: 32293516]
25. Onwuegbuzie A, Collins K. A typology of mixed methods sampling designs in social science research. Qual Rep 2007;12(2):281-316 [FREE Full text]
26. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
27. Jones N, Corrigan PW, James D, Parker J, Larson N. Peer support, self-determination, and treatment engagement: A qualitative investigation. Psychiatric Rehabilitation Journal 2013;36(3):209-214. [doi: 10.1037/prj0000008]
28. Kennedy AJ, Brumby SA, Versace VL, Brumby-Rendell T. The ripple effect: a digital intervention to reduce suicide stigma among farming men. BMC Public Health 2020 May 29;20(1). [doi: 10.1186/s12889-020-08954-5]
29. Kirchner S, Till B, Plöderl M, Niederkrotenthaler T. Videos for suicide prevention in LGBTIQ+ adolescents. Crisis 2020 May 19:1-8. [doi: 10.1027/0227-5910/a000691]
30. Andriessen K, Krysinska K, Draper B, Dudley M, Mitchell PB. Harmful or helpful? a systematic review of how those bereaved through suicide experience research participation. Crisis 2018 Sep;39(5):364-376. [doi: 10.1027/0227-5910/a000515]
31. Maple M, Wayland S, Sanford R, Spillane A, Coker S. Carers' motivations for, and experiences of, participating in suicide research. Int J Environ Res Public Health 2020 Mar 06;17(5) [FREE Full text] [doi: 10.3390/ijerph17051733] [Medline: 32155819]
32. Ryan J, Lopian L, Le B, Edney S, Van Kessel G, Plotnikoff R, et al. It's not raining men: a mixed-methods study investigating methods of improving male recruitment to health behaviour research. BMC Public Health 2019 Jun 24;19(1). [doi: 10.1186/s12889-019-7087-4]

Abbreviations

PHN: primary health network

Edited by G Eysenbach; submitted 03.09.20; peer-reviewed by J Mayoh, J Robinson; comments to author 11.10.20; revised version received 14.12.20; accepted 17.01.21; published 03.03.21.

Please cite as:
Carrotte ER, Webb M, Flego A, Vincent B, Heath J, Blanchard M
Acceptability, Safety, and Resonance of the Pilot Digital Suicide Prevention Campaign "Better Off With You": Qualitative Study
JMIR Form Res 2021;5(3):e23892
URL: https://formative.jmir.org/2021/3/e23892
doi:10.2196/23892
PMID:33656441

Original Paper

Medical Food Assessment Using a Smartphone App With Continuous Glucose Monitoring Sensors: Proof-of-Concept Study

Hector Roux de Bézieux[1,2,3], MA, MSc; James Bullard[1], PhD; Orville Kolterman[1], MD; Michael Souza[1], BSc; Fanny Perraudeau[1], PhD

[1]Pendulum Therapeutics, Inc, San Francisco, CA, United States

[2]Division of Biostatistics, School of Public Health, University of California, Berkeley, Berkeley, CA, United States

[3]Center for Computational Biology, University of California, Berkeley, CA, United States

Corresponding Author:
Fanny Perraudeau, PhD
Pendulum Therapeutics, Inc
933 20th Street
San Francisco, CA, 94107
United States
Phone: 1 650 276 6517
Email: fanny.perraudeau@pendulum.co

Abstract

Background: Novel wearable biosensors, ubiquitous smartphone ownership, and telemedicine are converging to enable new paradigms of clinical research. A new generation of continuous glucose monitoring (CGM) devices provides access to clinical-grade measurement of interstitial glucose levels. Adoption of these sensors has become widespread for the management of type 1 diabetes and is accelerating in type 2 diabetes. In parallel, individuals are adopting health-related smartphone-based apps to monitor and manage care.

Objective: We conducted a proof-of-concept study to investigate the potential of collecting robust, annotated, real-time clinical study measures of glucose levels without clinic visits.

Methods: Self-administered meal-tolerance tests were conducted to assess the impact of a proprietary synbiotic medical food on glucose control in a 6-week, double-blind, placebo-controlled, 2×2 cross-over pilot study (n=6). The primary endpoint was incremental glucose measured using Abbott Freestyle Libre CGM devices associated with a smartphone app that provided a visual diet log.

Results: All subjects completed the study and mastered CGM device usage. Over 40 days, 3000 data points on average per subject were collected across three sensors. No adverse events were recorded, and subjects reported general satisfaction with sensor management, the study product, and the smartphone app, with an average self-reported satisfaction score of 8.25/10. Despite a lack of sufficient power to achieve statistical significance, we demonstrated that we can detect meaningful changes in the postprandial glucose response in real-world settings, pointing to the merits of larger studies in the future.

Conclusions: We have shown that CGM devices can provide a comprehensive picture of glucose control without clinic visits. CGM device usage in conjunction with our custom smartphone app can lower the participation burden for subjects while reducing study costs, and allows for robust integration of multiple valuable data types with glucose levels remotely.

Trial Registration: ClinicalTrials.gov NCT04424888; http://clinicaltrials.gov/ct2/show/NCT04424888.

(JMIR Form Res 2021;5(3):e20175) doi:10.2196/20175

KEYWORDS

clinical trials; continuous glucose monitoring; lifestyle modification; mobile app; telemedicine; diabetes

Introduction

Diabetes is one of the most prevalent chronic diseases in the world, impacting over 422 million people worldwide [1]. In the United States alone, type 2 diabetes (T2D) affected over 32 million individuals as of 2018 [2]. Western lifestyle, including diet, has been shown to play a clear role in the development of the disease [3-5]. Although a large and growing number of drugs have been approved for the treatment of T2D, no cure or universally effective pharmacological intervention yet exists.

By contrast, studies where sustained dietary and behavioral changes were made (eg, caloric restriction) have shown profound improvements in glucose control in subjects who are able to adhere to the regimen [6]. For example, in a study conducted by Lean et al, 306 individuals with T2D were given an 850-calorie diet. While only 24% achieved the primary end goal of weight loss of ≥15 kg, 86% of subjects who achieved the goal saw remission of their diabetes by 12 months [7]. Such low compliance, which is routinely observed in trials with dietary and behavioral changes, limits the interpretability of the intervention. In addition, rigorously accounting for such attrition in a randomized controlled trial (RCT) can dramatically increase the cost and complexity of conducting such studies, especially over long periods of time. As a consequence, rigorous trials of feasible dietary changes in routine clinical practice settings are frequently not performed, which in turn limits the body of clinical evidence available to guide treatment [8,9]. Studies that are less invasive in terms of the participants' usual daily activities may enable collection of endpoint measurements matching the quality of traditional clinical study assessments while decreasing dropouts and improving the tracking of compliance.

The growing use of mobile technologies, wearable biosensors, and telemedicine in RCTs allows for closer patient monitoring and increased engagement while limiting overhead [10]. If a sensor can provide high-quality measurements of relevant clinical metrics, it may be leveraged to conduct robust trials outside the traditional clinical trial setting. In the context of diabetes, continuous glucose monitoring (CGM) devices represent a widely used type of sensor. A new generation of devices [11] has overcome earlier drawbacks of previous CGM devices, including cost and the continuing need for calibration [12], and has gained regulatory approval [13-15]. Consequently, the number of clinical trials using CGM devices to measure primary or secondary endpoints has increased [16-22], permitting new innovative approaches and clinical trial designs. For example, in a previous report [23], 36 subjects were enrolled remotely and received CGM devices via shipments. Instructions on CGM application and usage were given remotely, and 34 subjects used the devices correctly more than 95% of the time. Such recent results may enable and support the conduction of robust investigations outside the usual clinical trial setting.

Although CGM deployment in clinical research is now well established, the advent of this new generation of glucose monitors has led to a large increase in studies relying mostly or solely on CGM data as primary outcome measures. Despite this, most studies do not fully leverage the potential of CGM devices, as they are deployed only for specific time periods in the study, for example, the first and last days of the trial [17-21], and when conducting interventions and monitoring while attending a clinical center [24,25]. Furthermore, integrating and augmenting the data provided by CGM devices with other data types, such as food logs and activity data, are still relatively unexplored.

In this manuscript, we present the results from a 6-week, double-blind, randomized, placebo-controlled, exploratory 2×2 cross-over study (n=6). A custom smartphone app was used to record food intake, exercise, and alcohol consumption and integrate these data with CGM data to construct a visual diet log in sync with concurrent glucose levels. In a concomitant study [26], we observed changes in postprandial glucose response in patients with T2D who consumed a five-strain synbiotic medical food manufactured by the study sponsor. The postprandial glucose response was measured following a standardized meal-tolerance test (MTT) administered at the clinic. In parallel, we were also interested in following subjects using the medical food while they pursue their daily activities unencumbered by the disruptions introduced by required visits to a research clinic. Thus, in this exploratory study, subjects consumed the medical food, and the postprandial glucose response was measured using CGM devices via a self-administered MTT at home. The aim of this study was to demonstrate the feasibility of measuring clinically relevant data for research purposes using CGM devices and a smartphone app while subjects pursue their normal daily activities.

Methods

Study Design

A double-blind, randomized, placebo-controlled cross-over design was used to compare responses to the following two experimental interventions: a synbiotic medical food product, which was provided by the study sponsor, in the form of three capsules taken twice a day, and an inactive placebo present in similar capsules (Multimedia Appendix 1). Subjects wore CGM (Freestyle Libre) devices through the entirety of the study according to the manufacturer's instructions for use. At baseline, beginning/end of each period, and washout, subjects were asked to collect a stool sample. The concentrations of the bacterial strains, which were ingested in the medical food, in those stool samples were estimated using quantitative polymerase chain reaction (qPCR). Strain concentrations were averaged across three replicates. One subject (Subject 4) did not provide stool samples, so this part of the analysis was performed in five subjects. Anthropometric characteristics were also measured.

Throughout the study, subjects used a custom smartphone app for the Android platform. This app was developed to facilitate the real-time collection and integration of glucose levels, a visual food log, an alcohol log, and an exercise log, as well as record scanning compliance (how often do subjects scan their CGM devices) and study events (such as MTT). Through the app, subjects could take pictures and annotate them to describe their food consumption. They could also submit text annotations without pictures to log events such as exercise. Screenshots of the mobile app can be found in Multimedia Appendix 2 and Multimedia Appendix 3.

The goal of this exploratory study was to determine how to collect and analyze appropriate glucose metrics when conducting a medical food study with CGM devices and a novel smartphone app. This includes measuring operating characteristics (sensitivity and specificity) and sources of variance for those metrics when determined with Freestyle Libre glucose sensors, as well as identifying proper statistical tests of glucose metrics, for example, to test for changes in areas under the curve (AUCs) for glucose derived from self-administered MTTs between active and placebo groups. Additional endpoints of interest were change in body mass, change in fecal probiotic strain

concentration, expected lifespan of CGM sensors, CGM sensor scanning, photo logging of compliance rates, and usability feedback.

Subjects were randomly assigned to the two study arms (study product then placebo or vice versa). Subjects were assigned in pairs so that each arm contained the same number of subjects. Study staff and subjects were blinded to the identity of the study product. Randomization was performed by a statistician not directly involved in study conduct. Both periods lasted 13 to 17 days and were separated by a washout period of 3 to 6 days. Details of the durations of the study periods for each subject can be found in Multimedia Appendix 4.

This study was approved on March 13, 2018, by the Aspire Institutional Review Board (IRB) (protocol number WB01-205). Informed consent was obtained from each subject prior to participation. Since it was a prospective study, it was retrospectively registered on ClinicalTrials.gov (registration number NCT04424888).

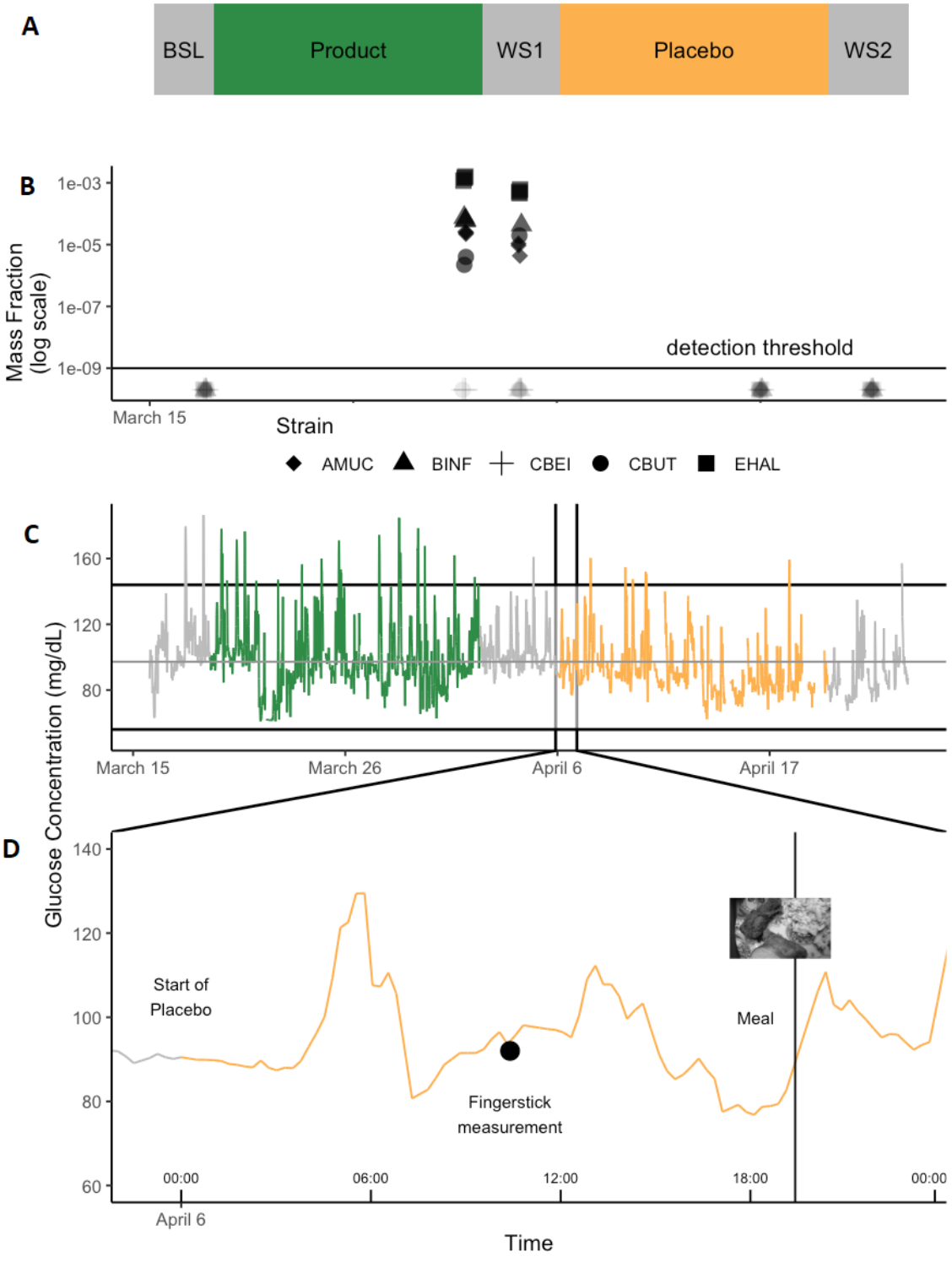
Subject Enrollment and Characteristics

Since this was an exploratory study, the main inclusion criterion was a minimum age of 17 years. Key exclusion criteria were pregnancy; continuous parenteral nutrition; and current use or planned use of antibiotic, antifungal, antiparasitic, or antiviral treatment during the study. Subjects were recruited by word of mouth from members of the biotechnology research community in San Francisco, California. A total of six subjects were enrolled in the study. Subjects were remunerated in the form of an Amazon gift card of US $100 for successful completion of each of the two study periods. Subjects had a run-in of 3 to 5 days before the first period (product/placebo) to familiarize themselves with the CGM devices and demonstrate that they could employ the devices as instructed. Subjects were shown how to use their CGM devices and how to log their meals and exercise and take pictures of their food intake through the smartphone app. Throughout the baseline period, they could ask questions about CGM usage. All subjects mastered correct use of the CGM devices and the smartphone app under observation by the study coordinator at the clinic. This consisted of properly collecting 24 hours of data and demonstrating the ability to take pictures and make annotations through the app.

Data Collection

The Freestyle Libre CGM devices used in this study provide estimates of interstitial fluid glucose levels. At 15-minute intervals, the estimates are recorded in the device and stored for up to 8 hours. Whenever the CGM receiver is wanded over the device, the 8 hours of data in the device at that time are downloaded for storage in an external database. These data can be used to display a subject's glucose levels at various time intervals. For a subject who started with the study product followed by washout (WS1), crossed over to placebo, and ended with a second washout (WS2) (Figure 1A), for example, we can visualize the glucose level throughout the study (Figure 1C) or zoom in on a specific day (Figure 1D). Glucose levels as estimated by the CGM device can be complemented by other data sources. A visual food log was constructed as subjects logged pictures of their meals through the smartphone app. As evidenced in Figure 1D, the recorded meal was associated with a spike in glucose levels. To facilitate interpretation of CGM data, we developed a peak detection algorithm based on methodology from Palshikar [27] that can automatically detect glucose peaks (Multimedia Appendix 1).

Figure 1. Glucose and dietary data collected for one subject. All data types collected for one subject. (A) Subject 1 starts in the active product arm (14 days) and then moves to the placebo arm (14 days). (B) At baseline, the end of each period, and study end, the subject provides stool samples. Strains present in the product are detected at the end of the active product period and a little after (arrow mark). (C) Continuous glucose monitoring (CGM) glucose levels are also tracked throughout the 6 weeks. (D) The values can be compared to the fingerstick measurements or used in coordination with pictures taken (here chicken wings at 7 PM) to detect meal-related glucose excursions. AMUC: *Akkermansia muciniphila*; BINF: *Bifidobacterium infantis*; BSL: baseline; CBEI: *Clostridium beijerinckii*; CBUT: *Clostridium butyricum*; EHAL: *Anaerobutyricum hallii*; WS: washout.



To ensure that their CGM devices were appropriately inserted, subjects also used conventional fingerstick devices (FreeStyle or CVS Health Blood Glucose Meter) to measure capillary glucose levels several times during the study and reported the values through the app. Fingerstick and CGM measurements were in good agreement (Multimedia Appendix 5) and were within the expected range [13].

At the end of the study, after unblinding, the clinical coordinator conducted an interview with all subjects, excluding Subject 5.

This interview was guided by a predefined template (Multimedia Appendix 6). Questions focused on the following categories: study design, product, CGM device, mobile app, behavior changes, future studies, and closing remarks, and there was a mix of open-ended and closed-ended questions. Subjects were also asked if they would participate in another similar study and if they would recommend others to a study like this, on a scale ranging from 1 (no) to 10 (absolutely).

Finally, to assess the presence of our strains, stool samples were collected throughout the study, and strain concentrations were measured by qPCR of frozen fecal samples. qPCR analysis was conducted in five of the six subjects who provided stool samples. In Figure 1B, strain concentration information is overlaid with arm assignment. As expected, we only detected the presence of the strains after study product administration.

Measuring the Response to a Fasted MTT

The fasted MTT is a gold standard test to measure glucose control in subjects. A standardized meal was to be consumed in the morning after fasting for at least 6 hours, and subjects were asked to ingest no additional calories for one additional hour. All subjects ingested the same standardized meal as follows: a Boost nutritional drink containing 45 g of carbohydrates including 22 g of sugars [28]. The MTT was performed four times throughout the study by each subject (at the start and end of both the placebo and product periods).

Subjects recorded when they ingested the Boost drink through the smartphone app, and corresponding glucose levels were obtained from the CGM devices. The incremental AUC for glucose levels was calculated for each time point and subject [29]. The baseline for the incremental AUC was defined as the glucose level at t=0, when the Boost ingestion began, as recorded by the CGM device.

The AUCs at the beginning and end of the placebo/product periods were compared using ΔAUC as follows:

$$\Delta AUC = AUC_{End} - AUC_{Beginning} \textbf{(1)}$$

Finally, the cross-over design allowed the comparison of ΔAUC between the placebo and active products (ΔΔAUC). We performed a one-sided Wilcoxon signed-rank test to test the null hypothesis (H0: ΔΔAUC ≥0 versus H1: ΔΔAUC <0).

Results

Subjects and Sensors

All six enrolled subjects completed the study. No adverse events or major deviations from the protocol were reported. All hypoglycemic events (defined as excursion below 70 mg/dL for at least two consecutive data points, ie, 30 minutes) are summarized in Multimedia Appendix 7, according to the recommendations of Schnell et al [30]. In accordance with manufacturer recommendations, CGM sensors were worn for an average of 13 consecutive days. The first sensor was placed under the supervision of the study coordinator. Subjects then placed their second and third sensors by themselves. Four subjects (1, 3, 5, and 6) had to replace one sensor each after only 10 days, mostly because the sensor was dislodged while dressing. Multimedia Appendix 8 recapitulates a few glucose-based subject-level metrics, according to the report by Schnell et al [30].

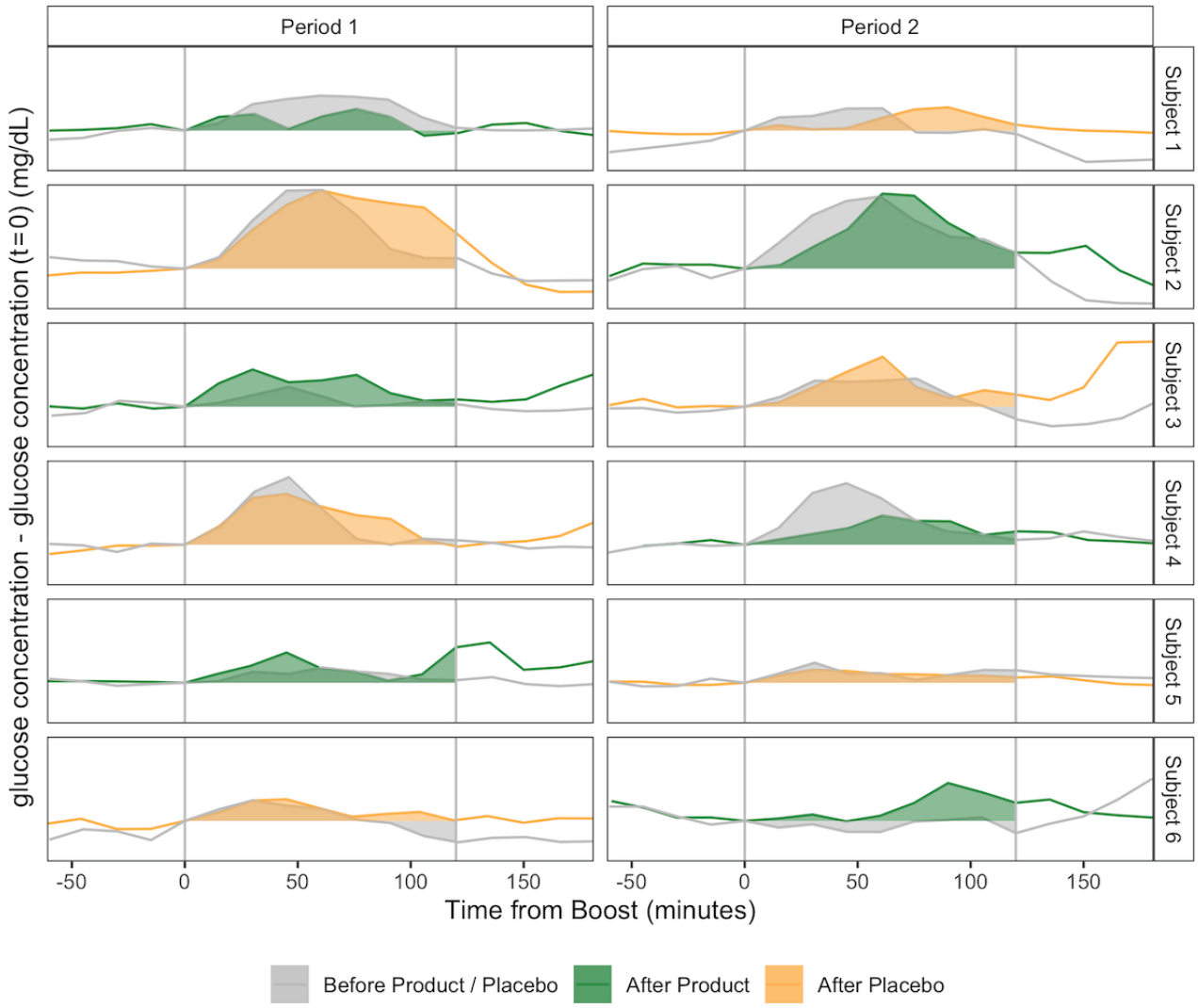
During interviews, subjects reported that CGM usage was a positive experience (Multimedia Appendix 9). However, two subjects inadvertently dislodged sensors because of clothing snags. Two subjects also reported minor soreness shortly after CGM implantation.

Fasted MTT Using CGM Devices

At four time points throughout the study, subjects performed MTTs as described above. MTTs were performed just before the start and at the end of the placebo/product periods. The standardized MTT data were used to compare a potential change in the glucose response for a fixed nutrient intake. If the study product had the effect of improving a subject's glucose metabolism, a decrease in the peak glucose level and/or AUC would be expected.

As shown in Figure 2, most subjects displayed an increase in glucose levels after taking the standardized meal. Each individual performed four MTTs (one at the beginning of each period, one at the end of the placebo period, and one at the end of the product period). Since the CGM devices used in this study generated a glucose measurement every 15 minutes, the curves are quite smooth, which allowed a robust computation of the AUC. The MTT data were obtained without additional subject engagement other than ingesting the standardized meal at home.

Figure 2. Results of the meal tolerance test (MTT). Glucose response to standardized meals for all six subjects in the study. Area under the curve values for the beginning of placebo/product, end of placebo, and end of product periods are in gray, yellow, and green, respectively. The MTTs are annotated by the subjects at time *t*=0 using the smartphone app. Glucose levels are normalized by the glucose level at *t*=0.



For Subjects 1, 2, and 4, the AUC decreased on ingesting the active study product, and $\Delta AUC_{Product}$ was smaller than $\Delta AUC_{Placebo}$. For Subjects 3, 5, and 6, the AUC increased on ingesting the product, and $\Delta AUC_{Product}$ was greater than $\Delta AUC_{Placebo}$ (Table 1 and Multimedia Appendix 10). Owing to the scale (n=6) and design of this pilot study, there is very limited power. In addition, none of the subjects were diagnosed with T2D, so we do not expect to find statistically significant results. Accordingly, the one-sided Wilcoxon signed-rank test had a *P* value of .28.

Table 1. Subjects ranked by baseline value of area under the glucose curve.

Subject[a]	Baseline AUC[b] (mg/dL/min)	$\Delta AUC_{Placebo}$ (mg/dL/min)	$\Delta AUC_{Product}$ (mg/dL/min)	ΔΔAUC (mg/dL/min)
6	265.3	1035.1	1683.2	648.1
3	737.0	1027.1	1598.3	571.2
5	834.8	102.8	552.8	450.0
4	1557.6	427.7	−933.0	−1360.7
1	2084.2	−221.6	−1213.9	−992.3
2	3931.8	1309.1	−978.1	−2287.2

[a]For Subjects 1, 2, and 4, the AUC decreased when taking the product and $\Delta AUC_{Product}$ was smaller than $\Delta AUC_{Placebo}$.

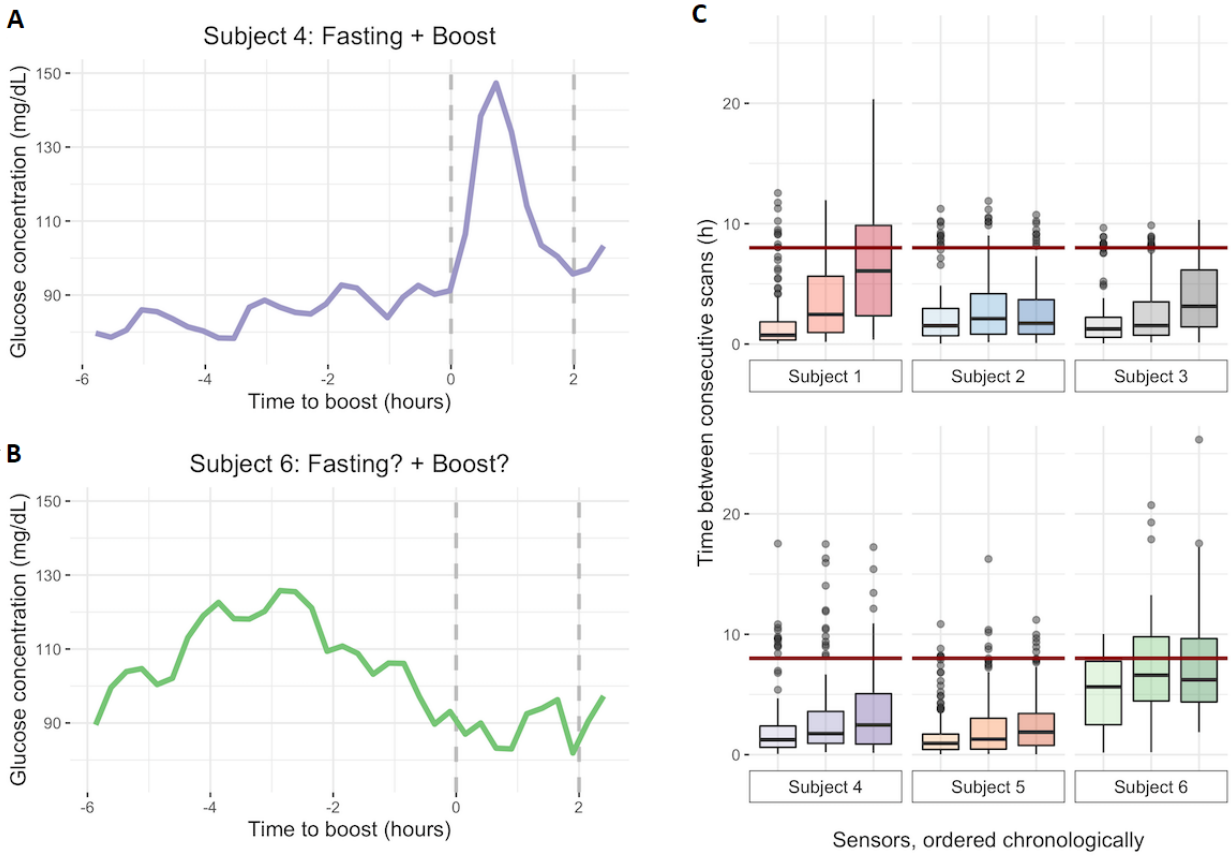
[b]AUC: area under the curve.

As this was a small pilot study, we did not use any baseline metric to stratify subjects based on their initial glucose control levels. However, a post-hoc analysis revealed that baseline AUC stratified the subjects into two clear and distinct groups (Table 1). Interestingly, those groups exactly matched the responder (AUC decrease)/nonresponder (AUC increase) status. Baseline AUC is an indication of the subject's glucose control level at the onset of the study. This is an interesting observation that should be explored in follow-up studies designed and powered appropriately. Using the mean ΔΔAUC value of the three responders as a reasonable effect size and considering the variance of the ΔΔAUC value from all subjects, a future study employing the same study design with 35 subjects is estimated to have a power above 90%.

CGM Enables Novel Measures of Compliance

CGM devices can be utilized to perform MTTs in the subject's living environment instead of the artificial setting of a research clinic. In this study, subjects were asked to use a custom Android app to create a visual food diary and log workouts and the start times of four standardized meal consumption tests. As such, we know precisely when the MTT was conducted. Subjects were asked to fast for 6 hours before taking the MTT, and Subject 4 seemed to have adhered to this. Indeed, the glucose level curve of this subject was rather flat for the 6 hours preceding the test. Then, right after ingesting the standardized meal, the glucose level spiked in a characteristic manner before returning to the baseline level after 2 hours (Figure 3A). On the other hand, Subject 6's response was quite different. For the test at the beginning of the product period, Subject 6's glucose level curve did not stay flat for the 6 hours before the glucose challenge (Figure 3B). Instead, it spiked from 90 mg/dL to 130 mg/dL and back over the course of 5 hours. Afterwards, glucose levels stayed flat during the MTT. Several explanations are possible. The MTT was performed at 7 AM, so the preceding peak might reflect a dawn effect [31], poor quality of sleep, alcohol consumption during the previous evening, or nighttime snacking. The subject's last recorded annotation before the Boost intake was at 8.44 PM the previous evening, so there is no indication of what could have caused this. Nonetheless, the fact that such an event was detectable before the MTT was performed is valuable information. Indeed, in such a case, it is easy to probe the subject for additional details to understand potentially confounding circumstances and ask the subject to take the MTT again the following day. This would lower variability by improving the quality of the MTT data, resulting in a more precise estimate of the product effect. A single estimate of fasting glucose at t=0, either via a fingerstick measurement or a blood draw as traditionally performed in clinical trials, would have failed to detect this.

Figure 3. Continuous glucose monitoring (CGM) devices provide insight into compliance. (A) CGM data are concordant between when Subject 4 mentioned fasting/consuming the standardized meal for the glucose challenge and what was detected from the CGM device. (B) CGM data can also be used to detect anomalies. There is no concordance between when Subject 6 recorded fasting/consumption of the standardized meal and what was observed from the CGM data. (C) Each subject is asked to obtain data at regular intervals and at least every 8 hours (red line), which is the limit for the sensor memory to store data.



A CGM device could also serve as an easy and efficient device for measuring engagement in the study since subjects needed to scan their devices regularly. In particular, since the devices used in the study could only store 8 hours of data, consecutive scans were not supposed to be more than 8 hours apart. Measuring the time intervals between scans and comparing them to the 8-hour mark provides a measure of the subject's compliance. In Figure 3C, time intervals between all consecutive scans are displayed. As the study progressed, all but one subject tended to let more time go between scans. This is consistent with interviews where subjects reported fatigue as the study progressed. Moreover, the number of scans with more than 8 hours in between also increased. This expected study fatigue affected subjects differently. While Subjects 3, 4, and 5 had small increases in their median times between scans, Subject 2 actually saw a small decrease by the end of the study. Subject 1 went from being the most compliant to the least over the course of the study, whereas Subject 6 displayed a relative lack of compliance throughout the study. Access to these data in real time could guide and target efforts to enhance compliance where needed.

Subjects could also record pictures and annotations through the smartphone app. The number of logs every day varied greatly between subjects, where the average number of logs per day ranged from 2.9 for Subject 4 to 9.73 for Subject 5, and throughout the study, as the variance for the number of logs per day was around 41% of the average. This reflects various behaviors that are captured by the smartphone app and that complement the CGM devices. For example, Subject 4 was one of the most compliant in terms of scanning requirements but logged the fewest number of meals. Study fatigue was also far less apparent than that in scanning requirements, as the average number of annotations per day decreased only slightly with time, suggesting that engagement with the smartphone was easier to maintain. However, compliance was also impacted by social norms, as reported in interviews. Subjects were less likely to log pictures in social settings and thus annotated after the fact, although recent trends, such as posting pictures of meals on social networks, have lowered the barrier.

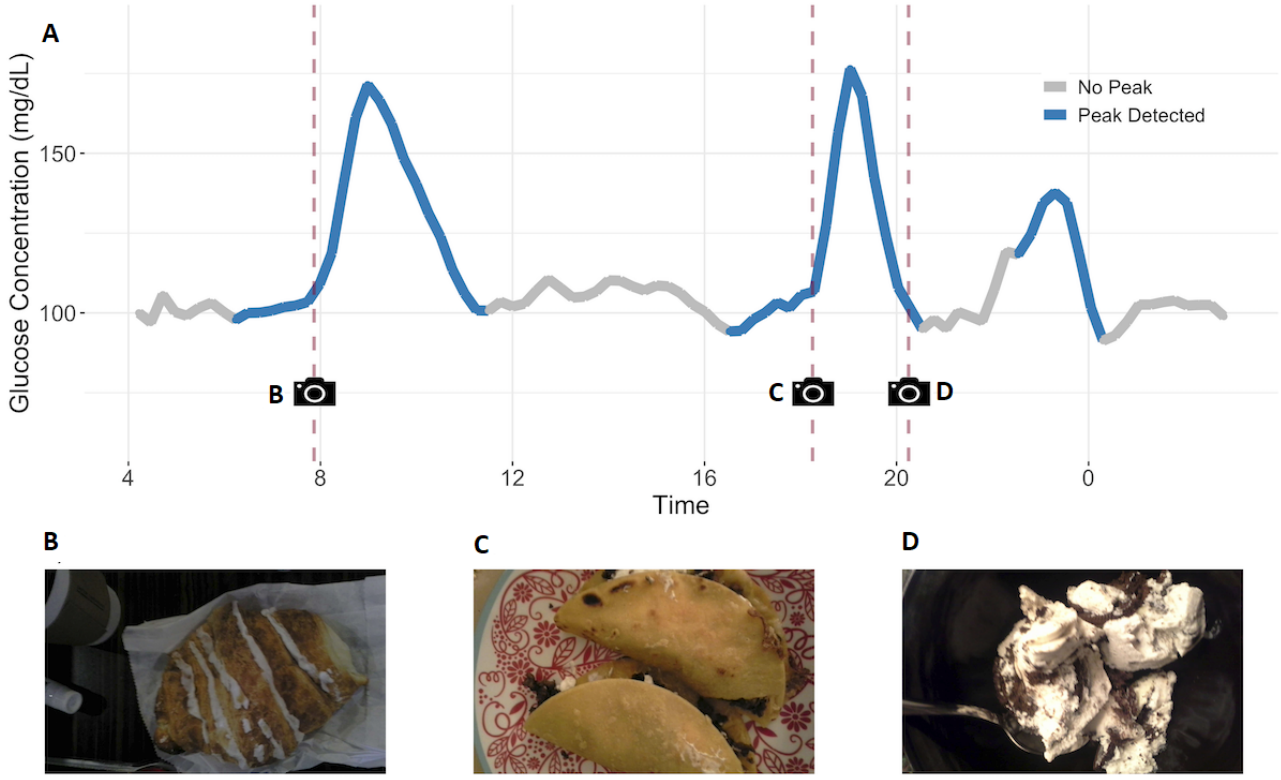
Integrating CGM Data With the Smartphone App

With a measurement every 15 minutes for 6 weeks, the CGM devices produced about 3000 data points per subject throughout the study. An MTT has nine CGM measurements (a measurement every 15 minutes for 2 hours), so all four MTTs represented a total of 36 points per subject throughout the study (1% of the data collected). Here, we show that the CGM devices do capture a trove of information that is not available in current clinical designs without CGM devices.

One feature of interest in continuous monitoring of glucose levels is glucose peaks. Those peaks are induced by a diverse set of behaviors, mostly food consumption and, in some subjects, exercise. Glucose peaks are characterized by their range and duration. Glucose levels of 140 mg/dL or above 2 hours after a meal are indicative of a lack of glucose tolerance [32]. Using our peak detection algorithm, we found that the subjects spent an average of 5 hours and 47 minutes of their day in glucose peaks, representing roughly 25% of all the data collected, whereas MTTs comprised only 1% of the data. We detected between two and three peaks a day, consistent with a routine of three meals a day.

An example of the types of peaks detected by the algorithm can be seen in Figure 4, where the results for March 20, 2018, for one subject are shown. The algorithm detected three peaks during this day, clearly differentiating between noisy fluctuations of glucose levels and glucose peaks induced by a meal. While most peaks are easy to distinguish visually, the amount of data means that automated annotation is necessary.

Figure 4. Continuous glucose monitoring (CGM) devices capture more data than the standard meal-tolerance test. (A) Zooming on a specific day allows one to display an entire day of glucose estimates (March 20, 2018, for Subject 1). Peaks are detected with the in-house algorithm (Multimedia Appendix 1) and colored in blue. Annotations logged in the smartphone app by the subject are marked with a dashed red vertical line and a picture logo. (B-D) Pictures submitted by Subject 1 through the app on March 20, 2018.



Integrating the CGM data with the visual diet log provides rich and insightful information. In particular, since the smartphone app was used for both recording glucose levels and taking pictures, we can precisely match the times of one with the other. On March 20, 2018, the subject recorded three meal pictures (Figure 4B-D). The start of the first and second peaks matched very well with the pictures taken. Interestingly however, the third picture (Figure 4D) and the third peak did not match. This picture was recorded over an hour before the start of the next glucose peak. The high fat content visible in the meal (Figure 4D) is consistent with this delay.

The app was only available for Android-based devices, and for this exploratory study, subjects were provided with compatible smartphones. The need to carry a second smartphone was highlighted as a burden in interviews, especially by users of alternative smartphone operating systems. Development of a general app that can be used directly on a subject's smartphone would lessen the burden on subjects in a future study. Subjects also mentioned several improvements to the design that would facilitate their interactions with the app.

qPCR-Based Secondary Endpoints

At baseline, the beginning/end of each period, and washout, subjects were asked to provide a stool sample. Analysis of stool samples was performed in five of the six subjects; one subject did not provide stool samples. Using fecal qPCR, the concentrations of the strains in the formulation were estimated. In the results shown in Multimedia Appendix 11, strain concentrations were averaged across three replicates. As expected, at baseline, during placebo, and during placebo washout, the strains in the formulation were not detected. This was true even when the placebo period followed the product period. Most strains were detected at the end of the product period in all subjects (Multimedia Appendix 11). For the subjects with a washout duration of 1 to 3 days (Multimedia Appendix 4), the strains in the formulation could still be detected, although generally at lower levels, after the product washout, suggesting that a washout period of 3 days is too short. Additionally, for the subjects on the active product during period 1 (Subjects 1, 3, and 5), strains were not detected at the end of the placebo period, suggesting that a washout duration between 5 days and 2 weeks is sufficient.

Code and Data Availability

All the codes and data necessary to reproduce the figures and tables in this paper are available at GitHub [33].

Discussion

Principal Findings

This manuscript describes a pilot study exploring the potential of CGM devices with a smartphone app that collects additional data as an avenue for measuring clinically relevant endpoints, assessing compliance in realistic settings, and decreasing the burden of study participation while collecting valuable data that can easily be integrated with glucose levels. Using CGM devices, it is possible to accurately measure important metrics, such as the outcome of a fasted MTT, the AUC of meal-induced

glucose peaks, and the compliance to a fasting regimen. We explored this framework in the context of a 6-week, double-blind, placebo-controlled, 2×2 cross-over pilot study (n=6) comparing a twice daily synbiotic medical food designed to improve glucose control. We demonstrated the promise of CGM devices as a means to assess clinical outcomes of nonpharmacological interventions in rigorous studies for research purposes.

Although accurate automatic detection of glucose peaks is straightforward and can provide valuable data for the clinician and individual, how to use those peaks as clinical endpoints is challenging as the glucose peaks represent a response to real-world behaviors and not to a prequantified study-wide intervention. Commonly, diet logs are used to control for dietary variation, but they have reporting bias and increase subject burden considerably [34]. In this study, we tried to address the issues of subject burden and reporting bias by employing a smartphone app that allowed subjects to take pictures of their meals. This allows the creation of a visual diet log that is in sync with the glucose levels measured with the CGM devices, which provides a rich and valuable source of data using a solution that is simpler to implement and less burdensome to participants. However, further work is required to bridge the gap between image data and actionable carbohydrate content estimates. Variation in the regularity of eating habits or meal content could potentially lead to higher between-subject variance. Even for a within-subject contrast, the lack of regularity in eating and exercise habits increases the variance of the estimate of any treatment effect. Future studies could also leverage those types of visual logs to monitor adherence to a specific diet type (eg, high fiber diet). Complementing the photos with more precise annotations could also improve estimates. Smartphone apps that rely on drop-down menus provide fair estimates of carbohydrate consumption [35,36] and greatly outperform self-assessment [37].

On the other end of the spectrum, the MTT protocol employed in traditional clinical studies represents a very controlled setting. This test yields a much more consistent measure, but it may not reflect the real-world effectiveness of an intervention as accurately. As is expected and has been reported in various studies [38,39], compliance decreases when subjects are increasingly inconvenienced. MTTs, as required in this study, necessitated consuming a specific shake, undergoing a 6-hour fast, and monitoring for 2 to 3 hours following beverage consumption. As this disrupts an individual's daily routine, the number of observations per subject per week needs to be limited. Focusing on breakfast-induced glucose peaks might represent a middle ground. As can be seen in Multimedia Appendix 12, all subjects, except Subject 5, had a high number of peaks at a consistent time most mornings. This denotes a regular breakfast routine at the subject level. This was confirmed in interviews, where subjects reported having either a consistent weekday breakfast or a rotation among a limited set of menus. Defining a consistent breakfast for all subjects a few days a week would provide a robust measure of the level of glucose control in a realistic setting with reduced variability.

We expect that future studies will leverage the novel capabilities of CGM devices to provide real-time compliance data. For example, a reminder to subjects whenever compliance is detected to be declining, in the form of a push notification or a phone call, could ensure that data capture remains sufficient, with the aim to improve the robustness of the inference. Subjects unanimously said in the final interviews that reasonable notifications would increase compliance, and such methods have been successfully deployed in previous studies [40]. CGM devices could also increase subjects' engagement. In this study, subjects were presented with an interactive dashboard recapitulating their glucose and food behaviors (Figure 4). Subjects were excited about this and spent time looking at the effects of individual foods on their glucose levels. Presenting an example dashboard to the subjects before baseline, with the promise of more complete results at the end, could increase engagement in the study and therefore compliance. While such usage of CGM devices shows clear promise, care needs to be taken to ensure sound experimental design, including maintaining subject and researcher blinding to the arm assignment. In spite of these issues, proper usage of CGM devices can provide just-in-time actionable measures of compliance.

The CGM devices deployed in this study measured glucose-related metrics, but they can be complemented with other types of data collection depending on the setting. In this exploratory study, stool samples were collected. While collection of this type of data is feasible at home, the samples still need to be delivered to a clinic for analysis. However, the difficulties in collecting this type of data are limited by the fact that they are not needed for all patients or in confirmatory studies. CGM devices can also be used in coordination with other sensors (eg, smartwatches and mobility data from the smartphone) that could be complementary with little added burden on the patient [41,42]. The general principles of compliance measurements discussed above in the context of CGM devices are also applicable to any other metric relying on a connected device. Privacy concerns should also be addressed properly, as most subjects raised this issue during the interviews.

Limitations

Previous studies [43] have linked the regularity of CGM device scanning to improvements in the "time in range" endpoint (glucose level between 70 mg/dL and 180 mg/dL). While we did not observe it in this study given the general lack of glucose intolerance in the study subjects, differences in engagement between subjects could be linked to differences in outcomes in larger studies. Controlling for adherence, directly measured through scanning and recording annotations, could limit this phenomenon and lead to less noisy estimates of the treatment effect.

Given the scale of the study, a cross-over design was employed as a means to guard against a number of potential confounders. In particular, anecdotal observations while developing the app suggested that individuals responded to the sensor in a distinct and consistent manner. Coupled with the fact that the MTTs were self-administered, a cross-over design seemed logical. Despite these factors, cross-over designs can be of limited utility in the presence of carryover or spillover effects. Indeed, strains were detected at the beginning of the placebo period in some

subjects crossing over from the active product period, indicating the potential for carryover effects. No strains were detected at the end of the placebo period for those same subjects, suggesting that a longer washout period will be required in a future cross-over study.

Conclusions

We have shown that CGM devices can provide high-quality measurements of relevant endpoints and can enable granular monitoring of subject compliance with low clinical overhead. While this was a small pilot study, we observed a meaningful effect of the study product in subjects with diminished initial glucose control. A broader follow-up study, both in size and duration, is necessary to validate our very preliminary results.

Acknowledgments

We thank Suzanne Dufault, Nima Hejazi, and Ivana Malenica for insightful discussions on clinical trial designs. We gratefully thank the participants of this study.

Editorial Notice

This randomized study was only retrospectively registered, explained by authors with “Since it was a prospective trial, we did not register it beforehand.”. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials because the risk of bias appears low and the study was considered formative. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Authors' Contributions

FP, JB, and OK designed the study. MS, FP, and OK conducted the study. HRdB and FP analyzed the data. HRdB wrote the first draft of the manuscript, and FP, JB, OK, and MS suggested edits.

Conflicts of Interest

All authors are employees and stock/stock option shareholders of Pendulum Therapeutics, Inc (formerly known as “Whole Biome Inc”). OK owns stock in GlySens, Inc, has stock options in ViaCyte, Inc, and is a consultant to NuSirt BioPharma, Adocia, Circius, and NanoPrecision Medical.

Multimedia Appendix 1
Supplementary methods.
[PDF File (Adobe PDF File), 397 KB - formative_v5i3e20175_app1.pdf]

Multimedia Appendix 2
Screenshot of the app.
[PDF File (Adobe PDF File), 39 KB - formative_v5i3e20175_app2.pdf]

Multimedia Appendix 3
Screenshot of the app.
[PDF File (Adobe PDF File), 51 KB - formative_v5i3e20175_app3.pdf]

Multimedia Appendix 4
Number of days spent in each period. Each subject spent at least 2 weeks in each arm of the study. The washout period between the two arms is at least 3 days.
[XLS File (Microsoft Excel File), 20 KB - formative_v5i3e20175_app4.xls]

Multimedia Appendix 5
Difference between fingerstick values and continuous glucose monitoring (CGM) estimates. Each panel shows the fingerstick value logged through the mobile app (with a red dot) and the glucose levels as measured with the CGM devices. Values are in accordance with expected values [14].
[PNG File , 1050 KB - formative_v5i3e20175_app5.png]

Multimedia Appendix 6
Interview guide.
[PDF File (Adobe PDF File), 45 KB - formative_v5i3e20175_app6.pdf]

Multimedia Appendix 7
According to the recommendations of Schnell et al [30], we report all hypoglycemic events, defined as excursions below 70 mg/dL for at least two consecutive values.
[XLS File (Microsoft Excel File), 30 KB - formative_v5i3e20175_app7.xls]

Multimedia Appendix 8
According to the recommendations of Schnell et al [30], we report four summary statistics at the subject level, measuring time out of range (as number of continuous glucose monitoring scans per day), and glucose variability, reported as the average within-day standard deviation of glucose levels and the standard deviation of the average daily glucose level.
[XLS File (Microsoft Excel File), 20 KB - formative_v5i3e20175_app8.xls]

Multimedia Appendix 9
Interview summary.
[PDF File (Adobe PDF File), 33 KB - formative_v5i3e20175_app9.pdf]

Multimedia Appendix 10
Area under the curve (AUC) during the self-administered meal-tolerance tests. Incremental AUCs are computed using the glucose levels provided by the continuous glucose monitoring devices. Subjects recorded the time of the ingestion of the Boost drink using the smartphone app. Incremental AUCs are rounded to the nearest unit.
[XLS File (Microsoft Excel File), 20 KB - formative_v5i3e20175_app10.xls]

Multimedia Appendix 11
Fecal quantitative polymerase chain reaction measures of formulation strain concentration. Each subject, except for Subject 4, provided stool samples. As expected, strains in the formulation were detected when the subjects were taking the formulation and not detected when the subjects were not (ie, at baseline, during placebo, and during placebo washout). Some carryover is observed. Strains are detectable during a few days following the product period.
[PNG File , 88 KB - formative_v5i3e20175_app11.png]

Multimedia Appendix 12
Times of the tops of the peaks during the day. Peaks are detected using the algorithm described in Multimedia Appendix 1. For most subjects, a clear trend in the morning suggests a routine linked to breakfast.
[PNG File , 27 KB - formative_v5i3e20175_app12.png]

Multimedia Appendix 13
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1345 KB - formative_v5i3e20175_app13.pdf]

References

1. World Health Organization. Global Report on Diabetes. Geneva: World Health Organization; 2016.
2. Centers for Disease Control and Prevention. National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States. Atlanta, GA: Centers for Disease Control and Prevention, US Department of Health and Human Services; 2020.
3. van Dam RM, Rimm EB, Willett WC, Stampfer MJ, Hu FB. Dietary patterns and risk for type 2 diabetes mellitus in U.S. men. Ann Intern Med 2002 Feb 05;136(3):201-209. [doi: 10.7326/0003-4819-136-3-200202050-00008] [Medline: 11827496]
4. Cordain L, Eaton SB, Sebastian A, Mann N, Lindeberg S, Watkins BA, et al. Origins and evolution of the Western diet: health implications for the 21st century. Am J Clin Nutr 2005 Feb;81(2):341-354. [doi: 10.1093/ajcn.81.2.341] [Medline: 15699220]
5. Qi L, Cornelis MC, Zhang C, van Dam RM, Hu FB. Genetic predisposition, Western dietary pattern, and the risk of type 2 diabetes in men. Am J Clin Nutr 2009 May;89(5):1453-1458 [FREE Full text] [doi: 10.3945/ajcn.2008.27249] [Medline: 19279076]
6. Franz MJ, MacLeod J. Success of nutrition-therapy interventions in persons with type 2 diabetes: challenges and future directions. Diabetes Metab Syndr Obes 2018;11:265-270 [FREE Full text] [doi: 10.2147/DMSO.S141952] [Medline: 29928137]
7. Lean ME, Leslie WS, Barnes AC, Brosnahan N, Thom G, McCombie L, et al. Primary care-led weight management for remission of type 2 diabetes (DiRECT): an open-label, cluster-randomised trial. The Lancet 2018 Feb 10;391(10120):541-551. [doi: 10.1016/S0140-6736(17)33102-1] [Medline: 29221645]

8. Crichton GE, Howe PR, Buckley JD, Coates AM, Murphy KJ, Bryan J. Long-term dietary intervention trials: critical issues and challenges. Trials 2012 Jul 20;13(1):111 [FREE Full text] [doi: 10.1186/1745-6215-13-111] [Medline: 22812577]
9. Lewis JD, Albenberg L, Lee D, Kratz M, Gottlieb K, Reinisch W. The Importance and Challenges of Dietary Intervention Trials for Inflammatory Bowel Disease. Inflamm Bowel Dis 2017 Feb;23(2):181-191 [FREE Full text] [doi: 10.1097/MIB.0000000000001009] [Medline: 28079619]
10. Coran P, Goldsack JC, Grandinetti CA, Bakker JP, Bolognese M, Dorsey ER, et al. Advancing the Use of Mobile Technologies in Clinical Trials: Recommendations from the Clinical Trials Transformation Initiative. Digit Biomark 2019;3(3):145-154 [FREE Full text] [doi: 10.1159/000503957] [Medline: 32095773]
11. Waltz E. Sweet sensation. Nat Biotechnol 2019 Apr;37(4):340-344. [doi: 10.1038/s41587-019-0086-2] [Medline: 30940935]
12. Hirsch IB. History of Glucose Monitoring. In: Role of Continuous Glucose Monitoring in Diabetes Treatment. Arlington, VA: American Diabetes Association; 2018.
13. Freestyle Libre 14 Day Flash Glucose Monitoring System. US Food & Drugs Administration. 2017. URL: https://www.fda.gov/medical-devices/recently-approved-devices/freestyle-libre-14-day-flash-glucose-monitoring-system-p160030s017 [accessed 2020-03-12]
14. Declarations of Conformity. Abbott. URL: https://www.diabetescare.abbott/declarations-of-conformity.html [accessed 2020-03-25]
15. Murata T, Sakane N, Kato K, Tone A, Toyoda M. The Current Intermittent-Scanning CGM Device Situation in Japan: Only Adjunctive Use to SMBG Is Approved and the Latest Health Insurance Coverage Details. J Diabetes Sci Technol 2018 May;12(3):729-730 [FREE Full text] [doi: 10.1177/1932296817749860] [Medline: 29281891]
16. Bergenstal RM, Rosenstock J, Bastyr EJ, Prince MJ, Qu Y, Jacober SJ. Lower glucose variability and hypoglycemia measured by continuous glucose monitoring with novel long-acting insulin LY2605541 versus insulin glargine. Diabetes Care 2014;37(3):659-665. [doi: 10.2337/dc12-2621] [Medline: 24198302]
17. Koyanagawa N, Miyoshi H, Ono K, Nakamura A, Cho KY, Yamamoto K, et al. Comparative effects of vildagliptin and sitagliptin determined by continuous glucose monitoring in patients with type 2 diabetes mellitus. Endocr J 2016 Aug 31;63(8):747-753 [FREE Full text] [doi: 10.1507/endocrj.EJ16-0266] [Medline: 27321385]
18. Jiménez-Domínguez G, Ble-Castillo JL, Aparicio-Trápala MA, Juárez-Rojop IE, Tovilla-Zárate CA, Ble-Castillo DJ, et al. Effects of Acute Ingestion of Native Banana Starch on Glycemic Response Evaluated by Continuous Glucose Monitoring in Obese and Lean Subjects. Int J Environ Res Public Health 2015 Jul 06;12(7):7491-7505 [FREE Full text] [doi: 10.3390/ijerph120707491] [Medline: 26154657]
19. Nishimura R, Osonoi T, Kanada S, Jinnouchi H, Sugio K, Omiya H, et al. Effects of luseogliflozin, a sodium-glucose co-transporter 2 inhibitor, on 24-h glucose variability assessed by continuous glucose monitoring in Japanese patients with type 2 diabetes mellitus: a randomized, double-blind, placebo-controlled, crossover study. Diabetes Obes Metab 2015 Aug;17(8):800-804 [FREE Full text] [doi: 10.1111/dom.12481] [Medline: 25930989]
20. Jinnouchi H, Koyama M, Amano A, Takahashi Y, Yoshida A, Hieshima K, et al. Continuous Glucose Monitoring During Basal-Bolus Therapy Using Insulin Glargine 300 U mL(-1) and Glargine 100 U mL (-1) in Japanese People with Type 1 Diabetes Mellitus: A Crossover Pilot Study. Diabetes Ther 2015 Jun;6(2):143-152 [FREE Full text] [doi: 10.1007/s13300-015-0115-1] [Medline: 26055218]
21. Holst JJ, Buse JB, Rodbard HW, Linjawi S, Woo VC, Boesgaard TW, et al. IDegLira Improves Both Fasting and Postprandial Glucose Control as Demonstrated Using Continuous Glucose Monitoring and a Standardized Meal Test. J Diabetes Sci Technol 2015 Oct 06;10(2):389-397 [FREE Full text] [doi: 10.1177/1932296815610124] [Medline: 26443290]
22. Danne T, Philotheou A, Goldman D, Guo X, Ping L, Cali A, et al. A randomized trial comparing the rate of hypoglycemia--assessed using continuous glucose monitoring--in 125 preschool children with type 1 diabetes treated with insulin glargine or NPH insulin (the PRESCHOOL study). Pediatr Diabetes 2013 Dec;14(8):593-601 [FREE Full text] [doi: 10.1111/pedi.12051] [Medline: 23730996]
23. Study Demonstrates Remote CGM Initiation Successful at Improving Diabetes Outcomes and Quality of Life. PRWeb. URL: https://www.prweb.com/releases/study_demonstrates_remote_cgm_initiation_successful_at_improving_diabetes_outcomes_and_quality_of_life/prweb16940784.htm [accessed 2020-02-17]
24. Aberer F, Lichtenegger KM, Smajic E, Donsa K, Malle O, Samonigg J, et al. GlucoTab-guided insulin therapy using insulin glargine U300 enables glycaemic control with low risk of hypoglycaemia in hospitalized patients with type 2 diabetes. Diabetes Obes Metab 2019 Mar;21(3):584-591 [FREE Full text] [doi: 10.1111/dom.13559] [Medline: 30328252]
25. Reddy R, El Youssef J, Winters-Stone K, Branigan D, Leitschuh J, Castle J, et al. The effect of exercise on sleep in adults with type 1 diabetes. Diabetes Obes Metab 2018 Feb;20(2):443-447 [FREE Full text] [doi: 10.1111/dom.13065] [Medline: 28718987]
26. Perraudeau F, McMurdie P, Bullard J, Cheng A, Cutcliffe C, Deo A, et al. Improvements to postprandial glucose control in subjects with type 2 diabetes: a multicenter, double blind, randomized placebo-controlled trial of a novel probiotic formulation. BMJ Open Diabetes Res Care 2020 Jul;8(1) [FREE Full text] [doi: 10.1136/bmjdrc-2020-001319] [Medline: 32675291]
27. Girish P. Simple Algorithms for Peak Detection in Time-Series. ResearchGate. 2009. URL: https://www.researchgate.net/publication/228853276_Simple_Algorithms_for_Peak_Detection_in_Time-Series [accessed 2019-05-12]

28. BOOST® Original. Nestlé Health Science. URL: https://www.nestlehealthscience.com/brands/boost/boost-original [accessed 2019-11-20]
29. Potteiger J, Jacobsen D, Donnelly J. A comparison of methods for analyzing glucose and insulin areas under the curve following nine months of exercise in overweight adults. Int J Obes Relat Metab Disord 2002 Jan 15;26(1):87-89. [doi: 10.1038/sj.ijo.0801839] [Medline: 11791151]
30. Schnell O, Barnard K, Bergenstal R, Bosi E, Garg S, Guerci B, et al. Role of Continuous Glucose Monitoring in Clinical Trials: Recommendations on Reporting. Diabetes Technol Ther 2017 Jul;19(7):391-399 [FREE Full text] [doi: 10.1089/dia.2017.0054] [Medline: 28530490]
31. Schmidt MI, Hadji-Georgopoulos A, Rendell M, Margolis S, Kowarski A. The dawn phenomenon, an early morning glucose rise: implications for diabetic intraday blood glucose variation. Diabetes Care 1981;4(6):579-585. [doi: 10.2337/diacare.4.6.579] [Medline: 6751733]
32. Two-Hour Postprandial Glucose. University of Rochester Medical Center. URL: https://www.urmc.rochester.edu/encyclopedia/content.aspx?contenttypeid=167&contentid=glucose_two_hour_postprandial [accessed 2021-02-20]
33. GitHub. URL: https://github.com/wholebiome/WB01-205 [accessed 2021-02-15]
34. Thompson F, Subar A. Dietary Assessment Methodology. In: Nutrition in the Prevention and Treatment of Disease. Bethesda, MD: Elsevier; 2001:3-30.
35. Fallaize R, Zenun Franco R, Pasang J, Hwang F, Lovegrove JA. Popular Nutrition-Related Mobile Apps: An Agreement Assessment Against a UK Reference Method. JMIR Mhealth Uhealth 2019 Feb 20;7(2):e9838 [FREE Full text] [doi: 10.2196/mhealth.9838] [Medline: 30785409]
36. Ferrara G, Kim J, Lin S, Hua J, Seto E. A Focused Review of Smartphone Diet-Tracking Apps: Usability, Functionality, Coherence With Behavior Change Theory, and Comparative Validity of Nutrient Intake and Energy Estimates. JMIR Mhealth Uhealth 2019 May 17;7(5):e9232 [FREE Full text] [doi: 10.2196/mhealth.9232] [Medline: 31102369]
37. 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]
38. Spilker B. Methods of Assessing and Improving Patient Compliance in Clinical Trials. IRB: Ethics and Human Research 1992 May;14(3):1. [doi: 10.2307/3563718]
39. Claxton AJ, Cramer J, Pierce C. A systematic review of the associations between dose regimens and medication compliance. Clin Ther 2001 Aug;23(8):1296-1310. [doi: 10.1016/s0149-2918(01)80109-0] [Medline: 11558866]
40. Haas K, Hayoz S, Maurer-Wiesner S. Effectiveness and Feasibility of a Remote Lifestyle Intervention by Dietitians for Overweight and Obese Adults: Pilot Study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e12289 [FREE Full text] [doi: 10.2196/12289] [Medline: 30973338]
41. Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. BMJ Innov 2016 Jan;2(1):14-21 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000098] [Medline: 27019745]
42. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile Sensing and Support for People With Depression: A Pilot Trial in the Wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
43. Dunn TC, Xu Y, Hayter G, Ajjan RA. Real-world flash glucose monitoring patterns and associations between self-monitoring frequency and glycaemic measures: A European analysis of over 60 million glucose tests. Diabetes Res Clin Pract 2018 Mar;137:37-46 [FREE Full text] [doi: 10.1016/j.diabres.2017.12.015] [Medline: 29278709]

Abbreviations

AUC: area under the curve
CGM: continuous glucose monitoring
MTT: meal-tolerance test
qPCR: quantitative polymerase chain reaction
RCT: randomized controlled trial
T2D: type 2 diabetes

Edited by G Eysenbach; submitted 12.05.20; peer-reviewed by C Reis, S Pit, A Teles; comments to author 17.09.20; revised version received 22.10.20; accepted 24.01.21; published 04.03.21.

Please cite as:
Roux de Bézieux H, Bullard J, Kolterman O, Souza M, Perraudeau F
Medical Food Assessment Using a Smartphone App With Continuous Glucose Monitoring Sensors: Proof-of-Concept Study
JMIR Form Res 2021;5(3):e20175
URL: https://formative.jmir.org/2021/3/e20175
doi:10.2196/20175
PMID:33661120

Original Paper

Preservation of Person-Centered Care Through Videoconferencing for Patient Follow-up During the COVID-19 Pandemic: Case Study of a Multidisciplinary Care Team

Line Silsand[1*], PhD; Gro-Hilde Severinsen[1*], PhD; Gro Berntsen[1,2*], MD, PhD

[1]Norwegian Centre for E-health Research, Tromsø, Norway

[2]Institute of Community Medicine, UiT – The Arctic University of Norway, Tromsø, Norway

[*]all authors contributed equally

Corresponding Author:
Line Silsand, PhD
Norwegian Centre for E-health Research
Sykehusveien 23
Tromsø, 9019
Norway
Phone: 47 91335053
Email: line.silsand@ehealthresearch.no

Abstract

Background: The Patient-Centered Team (PACT) focuses on the transitional phase between hospital and primary care for older patients in Northern Norway with complex and long-term needs. PACT emphasizes a person-centered care approach whereby the sharing of power and the patient's response to "What matters to you?" drive care decisions. However, during the COVID-19 pandemic, videoconferencing was the only option for assessing, planning, coordinating, and performing treatment and care.

Objective: The aim of this study is to report the experience of the PACT multidisciplinary health care team in shifting rapidly from face-to-face care to using videoconferencing for clinical and collaborative services during the initial phase of the COVID-19 pandemic. This study explores how PACT managed to maintain person-centered care under these conditions.

Methods: This case study takes a qualitative approach based on four semistructured focus group interviews carried out in May and June 2020 with 19 PACT members and leaders.

Results: The case study illustrates that videoconferencing is a good solution for many persons with complex and long-term needs and generates new opportunities for interaction between patients and health care personnel. Persons with complex and long-term needs are a heterogeneous group, and for many patients with reduced cognitive capacity or hearing and vision impairment, the use of videoconferencing was challenging and required support from relatives or health care personnel. The study shows that using videoconferencing offered an opportunity to use health care personnel more efficiently, reduce travelling time for patients, and improve the information exchange between health care levels. This suggests that the integration of videoconferencing contributed to the preservation of the person-centered focus on care during the COVID-19 pandemic. There was an overall agreement in PACT that face-to-face care needed to be at the core of the person-centered care approach; the main use of videoconferencing was to support follow-up and coordination.

Conclusions: The COVID-19 pandemic and the rapid adoption of digital care have generated a unique opportunity to continue developing a health service to both preserve and improve the person-centered care approach for persons with complex and long-term needs. This creates demand for overall agreements, including guidelines and procedures for how and when to use videoconferencing to supplement face-to-face treatment and care. Implementing videoconferencing in clinical practice generates a need for systematic training and familiarization with the equipment and technology as well as for an extensive support organization. Videoconferencing can then contribute to better preparing health care services for future scenarios.

(JMIR Form Res 2021;5(3):e25220) doi:10.2196/25220

KEYWORDS

person-centered care; rapid digitalization; health care; videoconferencing; persons with complex, long-term needs; COVID-19

Introduction

The COVID-19 pandemic, which began in mid-March 2020, has created unprecedented challenges for society and health care systems worldwide. In Norway, as in many other countries, the government imposed extensive lockdowns and social distancing measures in public spaces to reduce the spread of the SARS-CoV-2 virus. Regarding health care services, most scheduled physical appointments and nonacute surgeries were cancelled to protect both health care personnel and patients. Nevertheless, it was necessary to accept new patients with acute or subacute conditions and follow up with long-term care patients. Hence, health care personnel were suddenly forced to rethink and find alternative solutions for face-to-face treatment and care of patients. This resulted in the swift implementation of telemedicine and, in particular, the increased use of videoconferencing.

Telemedicine is defined as communication over a distance in which video and audio are transmitted in near–real time. The provision of health care remotely by means of a video meeting is not new and has been well described; see for instance [1-7]. Despite this, previous research illustrates a paradox in telemedicine in terms of the low rate of adoption of the technologies despite strong policy-level strategies and many small-scale proof-of-concept examples [2,3,8]. In a study from 2017, Alami et al [3] found that the use of telemedicine in Norway from 2009 to 2015 did not exceed 0.5% of total outpatient activity at a regional level. The study identified three major themes affecting the implementation and use of telemedicine in Norway: (1) governance and strategy, (2) organizational and professional dimensions, and (3) economic and financial dimensions [3]. An example is the Norwegian Directorate of eHealth making changes to the reimbursement of videoconferencing in response to its increased and necessary use during the COVID-19 pandemic [9]. This is in accordance with previous studies that have shown that introducing video consultations is a complex change that disrupts long-established processes and routines [10,11].

Nevertheless, in recent years, telemedicine has attracted considerable interest as a means of delivering care to patients with long-term conditions [11,12]. The use of videoconferencing in the management of chronic diseases such as chronic obstructive pulmonary disease, diabetes, and depression has potential for improving health care services [1]. Almost all of this evidence pertains to selected samples of outpatients with chronic, stable conditions [10]. Consequently, the rapid spread of COVID-19 has accelerated the use of telemedicine for assessing, planning, coordinating, and performing treatment and care. The crisis seems to have lowered previous barriers to using telemedicine and to have led to the discovery of new ways of using digital health solutions in response to the crisis [12]. By using telemedicine, health care services can support not only secured care for patients with COVID-19 but also electronic prescriptions and triage of patients with COVID-19 in different phases. In addition, health care services can use telemedicine to support routine primary care when there may be a tradeoff, such as between frail older patients staying at home or coming to the hospital for an examination [10,12].

This study reports on the initial experience of a multidisciplinary health care team located at the University Hospital of North Norway (UNN) in using videoconferencing to provide treatment and care to older persons with complex long-term needs during the early months of the pandemic. In Norway, as well as in other countries, there is a rising number of older citizens with complex and long-term health-related needs: “Persons with complex long-term needs typically face multiple care providers, organizations, and specialists, and are especially vulnerable to care fragmentation” [13-16]. This group also dominates the top 5%-10% of spenders, who account for two thirds of high-level health care costs, both in Norway and internationally [17,18]. An extensive body of research indicates that the critical elements of high-quality care for persons with complex long-term needs include strong primary care with an inherent person-centered, integrated, and proactive care focus [19-26]. Despite this, current health care systems are designed for acute short-term needs, and they struggle to address the increasing number of persons with complex long-term needs. Therefore, there is significant interest in the potential of technology such as telemedicine for improving care and reducing costs in these patient populations, as well as in using technology to diminish the burden that individuals face as they manage multiple chronic conditions while adjusting to independent life in the community [10,11,27].

For the multidisciplinary health care team, hereinafter referred to as the Patient-Centered Team (PACT) [19,24], the COVID-19 pandemic has resulted in a reorientation of work practice and a need for a rapid digitalization process to maintain the delivery of person-centered treatment and care to their patients. PACT works as an intermediary team in the transfer phase between the hospital and primary health care and takes a person-centered care approach (see the Research Site section in the Methods for an extensive description of PACT). At the core of person-centered care is the sharing of power between patients and health care professionals, wherein care is driven by the patient’s answer to the question “What matters to you?” The intervention is a continuous process of trust-building, sensitive exploration, and cocreation between professionals and the patient to design and deliver a person-centered care plan.

Accordingly, PACT was concerned about how the use of videoconferencing would affect the person-centered care approach. In the literature, we find that the terms “person-centered care” and “patient-centered care” are used equally; see for instance [19,24-26,28-30]. In this paper, we use “person-centered care” throughout, together with the acronym “PACT” for “Patient-Centered Care Team” as defined in previous papers [19,24].

This case study is based on four semistructured focus group interviews with PACT health care personnel and managers, and we collected data about the experience gained through the period of rapid scaleup of the use of digital care.

The aim of the study is to generate knowledge about the role videoconferencing can play in preserving the person-centered care approach for persons with complex long-term needs. Accordingly, the following research questions have been formulated: How does PACT preserve a person-centered focus on care for persons with complex long-term needs in care

services when videoconferencing becomes the main mode of clinical communication due to the social distancing measures during the COVID-19 pandemic? What are the challenges and opportunities for health care personnel in PACT when it is necessary to make a rapid transfer from face-to-face care to video meetings?

Methods

This qualitative case study was conducted in May and June of 2020.

Research Site

To meet the so-called "silver tsunami"—that is, the increased number of persons with complex long-term needs—UNN and Tromsø Municipality, together with neighboring municipalities, established PACT in 2014 at three of UNN's geographical locations. The teams receive referrals from all health care levels, and they collaborate with health care services, including general practitioners (GPs) who focus on coordinating care in the transition phase between hospitals and municipality care to prevent rehospitalization [19]. It is the aim of PACT to establish a single comprehensive plan for patient follow-up, including hospitals, GPs, and municipal health services. A success factor of PACT is that the teams consist of a combination of hospital and municipality health care personnel (physicians, pharmacists, nurses, secretaries, physiotherapists, and ergonomists) working across organizational borders. They have access to the electronic health record systems in both the hospitals and the municipalities. This provides the team with extensive knowledge and experience to ensure the quality of the transitions; however, it also ensures a comprehensive patient pathway for this group across health care levels and organizational borders. The team takes a person-centered care approach that is aimed at ensuring a holistic understanding in which patients and health care professionals work together toward common goals that are both meaningful to the patient and aligned with what the professionals can offer. To achieve this, the team conducts a risk analysis early in the patient pathway, and the analysis is repeated later in the pathway if there is new information or if challenges occur. In the experience of PACT, the person-centered approach helps to shift the focus away from repairing and treating toward preventing new adverse events or deterioration and initiating early intervention when problems arise [19,24,31]. PACT has achieved excellent results in relation to reducing the risk of death and the number of emergency admissions and hospital stays through the increased use of elective health care services taking the person-centered care approach [19]. The hospital and municipalities share financial, management, and employment responsibilities for the team members [32].

Participants and Recruitment

This study includes four web-based focus group interviews with a total of 19 participants. The overall project, Dignity Care, had already established collaboration with PACT, and the idea for the paper was first raised in a collaboration meeting, with PACT members expressing concern about how to preserve a person-centered focus when communicating with patients exclusively through digital means. After the study protocol was completed and approved by the Norwegian Centre for E-health Research (NSE), the authors created an information leaflet about the study that included the interview topics. This was sent to the leader of PACT, who forwarded it to all the employees on the team along with an invitation to participate in a focus group interview. The participation in focus group interviews was voluntary, and 19 of the 26 employees in the team (73%) participated. No team member refused to participate; however, some team members were reallocated to other tasks during the COVID-19 pandemic, and others had time off on the day of the interview.

All participants in the study signed an informed consent form and emailed it to the first or second author. The 19 participants were health care personnel or managers of PACT in the age range of 28-62 years, and all but one were female. Many of the participants had over 20 years of experience in health care services, both in hospitals and municipal health care; some had further education in geriatrics psychiatry and interaction, and one participant held a master's degree in pharmacy. Some of the participants had specific expertise in nutrition for older persons, Parkinson disease, dementia, and management. Further details about the participants in the focus group interviews are presented in Table 1.

Table 1. Overview of the focus group interview participants (N=19).

Focus group	Participants, n (%)	Occupations of the participants	Locations	Time of interview
1	7 (37)	• Location 1: Hospital nurses (2), municipality nurse • Location 2: Physician • Location 3: Pharmacist • Location 4: Secretary • Location 5: Secretary	• Tromsø	90 minutes
2	4 (21)	• Location 1: Hospital nurse, physical therapist, municipality nurse • Location 2: Nurse	• Tromsø • Balsfjord	60 minutes
3	6 (32)	• Location 1: Physical therapist, nurses (3) • Location 2: Operations manager • Location 3: Medical department advisor	• Harstad • Senja • Narvik	60 minutes
4	2 (10)	• Location 1: Leader • Location 2: Leader	• Tromsø	60 minutes

The focus group interviews were conducted by the first and second authors, who alternated between functioning as moderator and facilitator of the interviews. The interviews were audio-recorded, and reflections were written down immediately after each interview.

We used a semistructured interview guide with three themes and 5-8 subquestions (see Multimedia Appendix 1). The subquestions were used as a checklist to ensure that all the relevant subjects were discussed during the interview. If a subquestion did not come up in conversation, we asked about it specifically. The interview guide was developed in collaboration with all three authors. The themes and subquestions were based on the research questions and on discussions with the third author, who has extensive knowledge from research within the field of patients with complex long-term needs and person-centered care.

The social distancing measures for in-person meetings also affected the design of our study. Accordingly, data were collected through video meetings, as face-to-face fieldwork of any kind was impossible. Focus group interviews conducted on the web are very demanding; some of the participants were in the same meeting room, while others were at other locations, resulting in up to five units being logged on at the same time. There were between 2 and 7 participants in the focus group interviews. Accordingly, the moderator was required to ensure that every participant had time to share his or her experience.

Analysis

The analysis resembled a systematic text condensation process, a descriptive thematic analysis strategy inspired by Malterud [33]. The four focus group interviews were recorded, transcribed, and then analyzed. First, the authors read the text to establish an overview of the data and identified the overall preliminary themes, including face-to-face practice, rearranging work processes, keeping a person-centered care focus, adapting to technology, and collaboration with other health care professionals. Second, the second author systematically reviewed all the transcribed interviews to identify meaning units and code them in accordance with the preliminary themes. Third, the first and second authors sorted the meaning units according to the preliminary themes. Some meaning units could be sorted into two themes, and these were tagged with a comment to highlight their multiple membership. Fourth, the first and second authors carried out an iterative reading of the systematized meaning units, reducing the content under each theme but maintaining quotations. The first and second authors discussed the condensates and made further adjustments to the text. Based on the previous steps, the preliminary themes were renamed as themes. Citations related to the meaning units were used to ensure that the presented results of the analysis reflected the original context [33]. In addition, in the final phase of writing the article, the first two authors presented the findings of the study to PACT, at which point PACT was able to provide feedback on the results. Prior to submission, all PACT members were offered the opportunity to read the article. We did not receive any objections to what we had presented and written.

Ethics

For several years, PACT and NSE have been engaged in a research collaboration as a work package in the project Patients and Professionals in Partnership (3P). This study is part of the 3P research project and was approved by the privacy ombudsman at UNN in 2020 (2020/6797). The project was reviewed by the regional ethical committee (REK) (2017/1084/REK nord), which found that the project was exempt from ethical approval.

When the pandemic occurred and digital solutions were being rapidly implemented, PACT requested an evaluation study to assess the consequences for the person-centered care approach when videoconferencing was used for follow-up of their patients. The first and second authors had not previously been involved in the research collaboration and were commissioned to conduct the interviews and data analysis to provide an “outsider” view in the evaluation.

All the participants were provided with a consent form for participation, which they approved in writing and sent to us electronically. Participation in the interviews was voluntary.

Results

Based on the response in the four focus group interviews, we evaluated how PACT health care personnel and managers experienced the use of videoconferencing to provide person-centered care for persons with complex long-term needs during the initial phase of the COVID-19 pandemic. In accordance with the analysis, the results are presented as condensed text with quotations and are organized according to the following five overall themes: the workflow in PACT before and after the pandemic; technical training and support as prerequisites for effective digital care; new means of collaboration and coordination; divided opinions on using videoconferencing for person-centered patient care; and how saving time on patient travel contributes to person-centered care.

The Workflow in PACT Before and After the Pandemic

In the traditional way of organizing the work, PACT—as represented by 2-4 health care professionals—would make an initial assessment of the patient’s situation by visiting the patient either while the patient was still in hospital or at the patient’s home. Health care personnel from PACT made 3-4 patient visits per day, in addition to attending collaboration meetings at hospitals or in municipalities. All collaborators, including GPs, homecare services, and health care personnel from the hospital, were required to attend collaborative meetings with the patient and PACT with the aim of coordinating patient services. As a result of the pandemic, however, PACT was required to limit face-to-face meetings with patients and collaborators as much as possible:

> *COVID-19 was like a sudden wake-up call; we just had to adjust quickly and improvise along the way. We had no choice but to be creative and rethink our work practice.* [Participant 7]

Normally, the health care personnel in PACT were organized into different teams, although they frequently collaborated with each other and shared their knowledge and experience of treatment and care. Due to the COVID-19 situation, it was necessary to physically separate the PACT team into different cohorts at the hospital, and the physician and pharmacist were required to work from home so that if one team was quarantined, the other team could keep working. There was an overall agreement among the PACT members that this separation hampered collaboration between the teams in terms of the informal sharing of knowledge: "I missed the small talk and information exchange when passing each other in the hallway" (Participant 1).

The team members agreed that the most significant changes to PACT's workflow when using only videoconferencing for patient communication and care can be summarized as follows: (1) there was less need for ambulation, which saved an extensive amount of time for PACT; (2) the number of health care professionals visiting the patients either at home or in the hospital ward was reduced from 2-4 to 1, with the other necessary team members participating in the meeting by videoconferencing, which saved a lot of resources; (3) there was a reduced need for PACT to visit patients at home, as they collaborated through videoconferencing with assistance from homecare services; and (4) using videoconferencing for collaboration meetings reduced the time spent on travelling to meetings. In addition, collaborators such as GPs were able to attend these meetings more regularly. Finally, PACT found itself performing a support role for hospital ward and homecare services that lacked experience in videoconferencing. These issues are elaborated upon further in the next sections.

Technical Training and Support Are Prerequisites for Effective Digital Care

Transforming the provision of care from in-person meetings to virtual care demanded extensive training and testing of the equipment, as multiple solutions and setups were being used simultaneously. The goal for PACT was for all members to be able to access and use the different videoconferencing solutions, including Skype, Join, Teams, and Easymeeting, with respect to setting up and running meetings, creating links, choosing the right browsers for the different solutions, and connecting sound and picture. All PACT members had the opportunity to practice and test the different digital solutions during the first month of the COVID-19 pandemic because there were fewer referrals from the municipalities and the hospital. However, although some of the PACT members were comfortable using digital solutions from previous videoconferencing projects, others were not. All team members agreed that the videoconferencing technology itself was the largest barrier to adoption. PACT decided that it was necessary to be comfortable with the technical equipment to integrate videoconferencing into clinical work:

> *We arranged morning meetings with 8 to 9 different units. At first it took ages before everybody was connected. We were disconnected from the meetings, the sound did not work, the picture froze, you name it… We were given an informal support role, and we used an hour to get everybody online.* [Participant 5]

All PACT members gained much useful experience from testing and trying out the equipment: "We have learnt that it is smart to test the equipment before a meeting starts to make sure the sound and picture are OK and that the solutions communicate with each other" (Participant 15).

In addition, PACT set out rules for good behavior in videoconferencing, for example, with respect to noise reduction, muting, and taking turns to speak.

The majority of the PACT members had a steep learning curve regarding the use of videoconferencing for clinical purposes, as the technology was challenging. However, the overall perception of the new way of interacting with patients and collaborators was positive: "We all had to think differently about patient care" (Participant 7). This generated innovative initiatives to continue providing person-centered care. When the use of videoconferencing rapidly increased, it led to an extensive need for support for team collaborators as well. PACT reported that health care personnel from other wards and organizations were not as privileged and had less time to familiarize themselves with and test videoconferencing equipment: "In a busy hospital ward, it is unlikely that health care personnel will be allocated time to test technical solutions the way we did" (Participant 9). Moreover, in the hospital, most of the wards had videoconferencing equipment in meeting rooms or offices that could not be used at a patient's bedside. The PACT members found that when setting up videoconferencing with patients at home, the patients often needed a relative or a health care professional to assist them and provide technical support: "It is important that videoconferencing become available for all patients and not only the ones that are lucky enough to have relatives to assist them" (Participant 14).

Multidisciplinary interaction is key to patient-centered care, and all PACT members stated that they often had to take on a support role to help their collaborators: "In this transitional period, we have been flexible and facilitated the increased use of videoconferencing" (Participant 8). Because videoconferencing was a new way of working for most health care personnel, it was important to make the experience positive for both patients and health care workers: "If the first meeting went well, it was more likely that the patients and collaborators would try videoconferencing for further meetings as well" (Participant 8). PACT had no authority to request that others use videoconferencing; however, they were able to persuade many collaborators and patients to try it, while assisting them along the way. Nevertheless, none of the PACT members considered this support role to be their responsibility in the future use of videoconferencing: "Our focus needs to be on the core competency of patient care. We must be careful not to end up as a videoconferencing helpdesk" (Participant 15).

New Means of Collaboration and Coordination

Collaborating Directly With Patients

There was a general understanding in PACT that introducing videoconferencing for interaction with patients would require close collaboration with homecare services, GPs, and hospital

wards. The participants explained that homecare services already had equipment for digital care before the pandemic, namely tablets for documenting treatment and care. During the pandemic, homecare services started using the tablets for videoconferencing meetings in collaboration with PACT. In addition, the physiotherapist from PACT, who was located at the hospital, found that it was possible to assess a patient's mobility and gait function using videoconferencing with assistance from homecare services. Moreover, PACT's physician and pharmacist, located in their respective home offices, said that they could collaborate with homecare services through videoconferencing to conduct a review of a patient's medication by "looking into" the medicine cabinets in the patient's home. Some of the PACT members said that homecare services were also involved in facilitating videoconferencing for the assessment of housing conditions while a patient was still hospitalized, with the aim of preparing the home before the patient's discharge: "You had the home presented from the patient's perspective, and they had a more active role in preparing their homecoming while they were still in hospital" (Participant 9).

The participants also said that videoconferencing was used when homecare services needed clinical support from PACT when patients' conditions acutely worsened. All the PACT units observed that this occurred more frequently than normal during the COVID-19 pandemic. The PACT members assumed that this was related to the patients not being able to attend day care centers and having less contact with GPs and homecare services, whereby changes in patient condition would normally be recognized. During the pandemic, there were times when a patient's condition would worsen and the homecare services would arrange videoconferencing from the patient's home with the PACT physician, who would assess the patient. In one case, a patient was admitted to hospital directly, bypassing the normal procedure of first physically visiting a GP. The PACT physician said: "Apart from a clinical examination, I get just as much information about the patient in a videoconferencing meeting as in a face-to-face meeting."

Videoconferencing was used for several meetings for which a patient normally would have travelled to a hospital. One of the PACT members reported on a stoma follow-up that was conducted at a medical center in a municipality setting through videoconferencing with a stoma nurse at the hospital. PACT members said that some patients also worried about travelling from a municipality with no COVID-19 cases to hospitals in cities with many cases. Over half of the PACT members were satisfied with the patient follow-up on videoconferencing: "We made things happen that we never thought would be possible before we started. It was exciting, it worked very well and people were surprisingly positive" (Participant 10). They also discussed the importance of coming up with good ideas and smart solutions for following up with a patient without physically being in the same place.

Approximately half of the professionals highlighted, from a person-centered care perspective, that using videoconferencing placed the patient in a central position in collaborative meetings involving several actors:

> *When the patient is at home holding an iPad, the patient is really the one being focused on rather than being just one of many in a face-to-face meeting. In videoconferencing, the patient can easily address everybody directly and thus has a more active and central role. I really liked the idea of having the patient as the center of attention.* [Participant 5]

Compared with the telephone, most of the PACT members felt they achieved a more personal follow-up using videoconferencing.

Coordination Among Health Care Personnel

In the beginning, PACT needed to convince collaborating health care personnel, including both hospital and homecare service personnel and GPs, to start using videoconferencing instead of the telephone: "The health care personnel were just as skeptical, if not even more so, than the patients about using videoconferencing for communication and collaboration" (Participant 3). However, PACT observed that after the collaborating health care personnel used videoconferencing for a while, they felt more comfortable, and it was easier to continue:

> *We had to work a lot with the wards to get them started. First, they said that their patients were not capable of attending videoconferencing meetings because of their condition and that it was necessary to meet the patients in person, to physically examine them. Nevertheless, they started to rethink when we argued that even physiotherapists made videoconferencing consultations with patients work.* [Participant 16]

During the first month of the pandemic, synchronous video communication became a valuable tool for interaction between health care personnel in different organizations and geographical locations. The PACT members all agreed that videoconferencing made it easier for collaborators such as GPs, homecare services, and service administrators (*forvaltingskontor*/*tildelingskontor* in Norwegian) to attend on a regular basis because they saved time on transportation and parking: "We have been even more multi-disciplinary than before the pandemic at coordination meetings" (Participant 14). There was general agreement in PACT that GPs in particular attended videoconferencing coordination meetings more regularly than they did when physical meetings were the only option. In addition, some PACT members said that it was easier to arrange meetings at short notice when participants could attend from their offices. There was general agreement in PACT that collaborative meetings on video were shorter than in-person meetings. Some of the PACT members elaborated, saying that collaborating health care personnel seemed to be better prepared for videoconferencing meetings and more focused on the problem to be addressed, and there was less small talk: "I think video meetings lasted only half as long as in-person ones; they were both time- and cost-saving" (Participant 19). Despite this, all PACT members found that videoconferencing was more demanding and intense than physical meetings, and some of them recommended limiting the meetings to 1 hour.

All PACT members agreed that introducing videoconferencing for collaboration across organizations improved the interorganizational relationship between all the actors involved in treating a patient. PACT had extensive collaboration with homecare services, as mentioned previously, as homecare services were the only professionals still making regular face-to-face patient visits. However, several PACT members reported that the personnel in hospital wards often seemed to find the technology challenging and preferred for PACT to configure the tablet used for videoconferencing at the patient's bedside. They further stated that health care personnel often assisted and supported patients in the wards in digital meetings. This was highly successful, as the ward professionals could add supplemental information to the patient's story and any ambiguities could be clarified. All PACT members emphasized that the hospital wards should have their own tablets for reasons of infection control, and that in the long run, the staff should have proper training so that they could manage videoconferencing equipment themselves. Although all PACT members were satisfied with being able to reach out to their patients using videoconferencing, there were pros and cons regarding its extensive use.

Different Opinions on Using Videoconferencing for Person-Centered Patient Care

The PACT members had different opinions on how well the person-centered focus was preserved when videoconferencing replaced face-to-face follow-up. Some found it difficult to work digitally when assessing patients. For example, if a patient was confined to bed in the hospital, it was difficult to obtain an overall impression of the patient's physical status through a screen:

> *I think physical meetings present a much broader view of the patient's status. My impression is that many patients "straighten up" for a few minutes during a digital visit. When you are in the same room, it is easier to assess the patient and evaluate his condition.* [Participant 2]

On the other hand, another team member said, "We had an extensive number of videoconferencing meetings over the last few months, and I think they worked extremely well, both for collaboration and for patient meetings" (Participant 9). PACT ran several multidisciplinary meetings to follow up with patients: "It was exciting and worked very well, and people were surprisingly positive" (Participant 11).

Persons with complex long-term needs are a heterogeneous group, and there was general agreement in PACT that videoconferencing was not a good solution for all of them. The health care personnel in PACT had a range of opinions on the potential of using videoconferencing to preserve person-centered care for this patient group. Some of the PACT members found it demanding to communicate by video with patients who had vision or hearing impairment. They reported that patients with hearing impairment had problems with verbal communication using a tablet. Other PACT members recognized that digital meetings could be challenging for patients with reduced cognitive capacities. One patient, for example, had difficulty understanding that the health care personnel participating by videoconferencing were actually "live" in the same room and not just a picture on the wall: "Oh yes, I remember her, when are we going to see her?" (Participant 1). All PACT members emphasized that patients with cognitive impairments often needed health care personnel or relatives to support them in videoconferencing. Hence, PACT reported that if health care personnel did not find that patients were competent to participate in video meetings, the patients were sometimes excluded from the meetings. However, other PACT members reported the opposite, namely that videoconferencing made it possible to include even the frailest patients from home or from the hospital:

> *We, the health care personnel, often have the most prejudice about which patients to include in videoconferencing. However, we had meetings with several patients with cognitive impairments and other disabilities that worked surprisingly well.* [Participant 7]

Another complicating issue brought up by some of the PACT members was the difficulty of preserving privacy in videoconferencing meetings with bedridden patients admitted to multibed patient rooms in wards: "It is a challenge that hospitals lack rooms with videoconferencing that are adapted for patients confined to bed" (Participant 15). However, there was common agreement among the PACT members that videoconferencing was a good way of keeping in contact with patients and maintaining collaboration amongst health care personnel in the demanding situation of the COVID-19 pandemic: "It is much more complex to work digitally; still, it is better than not reaching patients at all" (Participant 3).

Saving Time on Patient Travel Contributes to Person-Centered Care

All PACT members emphasized that providing person-centered care for persons with complex long-term needs is also about limiting the health care disruption to patients' daily lives as much as possible. PACT highlighted the importance of seeing the patient as an individual person and not just as an accumulation of different diseases: "For us health care workers, the disease often defines the patient. For the patient, the disease is just a part of their life, not all of it" (Participant 15). All PACT members conducted their work from a person-centered care focus, in terms of starting the collaboration by asking the patient "What matters to you?" and translating the answers into health-related goals.

Some PACT members mentioned that efficiency and cost savings have been important focus areas in recent years; however, saving patients' time should be just as valuable. They added that several persons with complex long-term needs struggle with mobility challenges with their diseases; accordingly, travelling takes much time and energy. PACT reported that many appointments and meeting cancellations might have been avoided through the use of virtual communication: "If the patient is allowed to exchange the long, exhausting trip to the hospital with a 15-minute videoconferencing check-up, this is an important person-centered gain" (Participant 15). PACT elaborated that for some patients, it takes days to recover from travelling all day: "We had so many positive results from using

videoconferencing that we believe many patients can reduce the number of trips they make to the hospital for follow-up" (Participant 9).

Discussion

Principal Findings

In the Results section, we present the empirical findings from the focus group interviews on how PACT experienced the change in their work practice from using videoconferencing for interaction with persons with complex long-term needs and with their relatives and health care collaborators when providing person-centered care. In summary, some PACT members found it very challenging to use videoconferencing for patient communication and wanted to return to the traditional way of working. Others thought that videoconferencing worked very well for all patient groups and found it inspiring to use videoconferencing as a means of continuing a person-centered approach for treatment and care under challenging distancing measures. The largest challenge with using video meetings was the technology itself and the unpredictability of making the technology work. PACT emphasized the need for extensive training and support, as well as for just one video solution for collaboration between different organizations. The greatest advantages of videoconferencing were the new opportunities for collaboration between patients and PACT, the related significant time saving on travelling for both parties and the better use made of the limited resources of health care personnel. The improved collaboration with other health care services was also highlighted. On the basis of the empirical findings presented, we will discuss the role of videoconferencing in preserving a person-centered care approach related first, to improving personalized pathways; second, to reorganizing work practice; and third, to addressing new support roles. As a frame for the discussion, we used the previous research presented in the Introduction and the experience of the three authors (health care personnel and researchers).

Video Meetings Improve the Personalized Patient Pathway

Health care organizations are increasingly shifting from a disease-centered to a person-centered care focus. A person-centered care approach is believed to enhance both the technical and patient-experienced quality of care and to better achieve the quadruple aim of improved care experience, health and function, cost-benefit ratios, and improved work life for those who deliver care [34]. There is no agreed-upon definition or method of measurement for person-centered care [28]. It is an approach which includes quality dimensions best assessed by the patient. However, this approach addresses how a fragmented care system can create a seamless, personalized pathway that addresses a person's needs, values, and preferences as they develop over time [24]. The results show that PACT found that videoconferencing worked well as a means of improving collaboration and communication in person-centered care: "For most patients, the person-centered care follow-up has been at least as good as normal when using videoconferencing" (Participant 7).

Efficiency and cost savings have been important focus areas in recent years [11,27]; however, saving the patient's time should be just as valuable. Previous research demonstrates that telemedicine can be used to replace referrals to an outpatient clinic, thus reducing travel and unnecessary hospital visits, especially for those living in remote areas [10,11]. Using video meetings eliminates the risk of costly cancellations, such as those due to bad weather such as snowstorms leading to closed roads and cancelled flights; this is a constant risk in northern Norway, where the winter is approximately 6 months long. The patient can avoid unnecessary travel time [2], which is beneficial for this fragile patient group. Technological solutions such as videoconferencing can support persons with complex long-term needs by allowing improved access to different parts of the fragmented care system, which can be adjusted to the person's needs and preferences as these develop over time [11,19,24,27]. A number of hospital consultations and follow-ups can be changed to video meetings. Hence, several PACT members stated that videoconferencing can help facilitate living with complex disorders and reduce the disruption of the patient's everyday life due to the disease as much as possible, which we believe is an important principle in further developing the concept of person-centered health care and which is in line with Bower et al [27,35].

As elaborated upon in the results section, persons with complex long-term needs are a heterogeneous group, and the use of videoconferencing was challenging for patients with vision and hearing impairments as well as for those with cognitive impairments. Greenhalgh et al [10] state that not all clinical situations are appropriate for video consultations. This is in line with the feedback from PACT and raises the need for guidelines on how to tailor the use of videoconferencing to different patient capabilities and disabilities [10].

Video Meetings Require Reorganization of Work Practices

Digital care can improve the use of fragmented care system resources [11,19,27]. PACT includes a combination of disciplines, such as physicians, nurses, and physiotherapists. When PACT visits patients at home, normally, they require 3-4 people to be present. During the pandemic, they reduced the number of team members visiting patients at first to 0 and then to 1 or 2. Using videoconferencing as a supplement reduced the need for outreach resources, as they could contact additional team members virtually when needed. They found this approach to be a more efficient and resource-saving way of organizing their multidisciplinary work. In addition, it was easier for other collaborating health care personnel in the municipality, hospital, and GP offices to participate in collaboration meetings by using videoconferencing from their offices, as this saved them from spending time traveling to other locations. Furthermore, as PACT members have stated, videoconferencing meetings were more "to the point" in terms of focusing on patient needs, all participants came prepared, and meetings were completed within the allocated time: "We did much more than expected and saved a great deal of time, even though we were not physically in the same place" (Participant 8).

Patient pathways for persons with complex long-term needs include all levels of health care services. Therefore, using videoconferencing to improve the fragmented care system requires that all actors be engaged, for which reason using the potential of videoconferencing requires an overall rearrangement of health care services so that videoconferencing is included as part of daily services [11,27]. This is in line with previous studies (eg, Alami et al [3]), which found that using digital care calls for organizational changes in processes, practices, cultures, communication, and the division of work. As stated in the introduction, previous research has identified three major themes that affect the implementation and use of telemedicine. The results from this study also support the importance of these themes when scaling its implementation in an organization. Using the three themes to understand the empirical results of this study is helpful for focusing on the premises for using videoconferencing, which are to not only preserve but also improve the person-centered approach for persons with complex long-term needs. The three themes are discussed below.

Theme 1: Governance and Strategy

The fragmented portfolio of videoconferencing solutions made it challenging for health care personnel to connect to meetings. All the PACT members stated they would like health care organizations to standardize the process to use only one video-meeting solution, as compared to the many solutions currently used. The PACT members said that the complex technology created a feeling of incompetence and of being unprofessional as health care personnel, which created a barrier to using videoconferencing.

Theme 2: Organizational and Professional Dimensions

Conducting videoconferencing meetings led to additional work for PACT. For example, PACT members were required to set up tablet computers for hospital wards and take on an information technology support role for other health care units. Moreover, in this empirical case, PACT were the ones coordinating the collaboration around the patients; hence, they continued this coordination in virtual meetings. Nevertheless, it is important to discuss and clarify who has *legal* responsibility when assessments, treatment, and care decided upon in virtual meetings must be conducted on behalf of an actor from a different organization (eg, when homecare services monitor blood pressure or an electrocardiogram on behalf of specialist health care).

Theme 3: Economic and Financial Dimensions

Videoconferencing shows the need for clarification of the regulations for reimbursement. In Norway, during the early phase of the COVID-19 pandemic, the reimbursement for videoconferencing was changed so that it was equal to that for face-to-face meetings in view of the increasing use of digital care. This covered the use of videoconferencing for outpatient consultations and appointments with GPs. Nevertheless, as described above, when actors from one organization conducted assessments and procedures on behalf of an actor from another organization, this could raise the question of reimbursement for the executing party. In addition, the question arose of who was responsible for financing the technical support when actors from different organizations collaborated on caring for the same patient. These issues were barely mentioned by PACT; however, in line with previous research, they are still important to highlight [3]. In addition, the potential of telemedicine in general to improve care and reduce costs is limited by a lack of rigorous evidence of actual impact [27].

Scaling Up Digital Care Creates a Need for New Support Roles

The results of this study show that new roles that support the technical use of telemedicine for health care personnel must be established as part of using videoconferencing to improve a person-centered care approach. This technical support must work in close collaboration with personnel at all organizational levels of health care:

> *It is important to organize high-quality support. If you have five doctors in the meeting and it takes 15 minutes to connect everyone, you have wasted more than an hour in total. You don't have to spend many of these 15-minute periods before you have spent the equivalent of paying a technician to handle the practical and technical issues related to videoconferencing.* [Participant 6]

The support must be readily available in terms of a short response time: "When you have video meetings, there is a risk that you will focus more on the technology and how to make everything work than on patient issues and the other actors in the meeting" (Participant 16). PACT was concerned that it would be locked into a future support role and would be in danger of compromising its focus on patients. PACT cannot maintain this support role; however, the support organization must have the clinical competence to understand clinician needs.

As described in the results, patients often needed assistance from homecare services or relatives to connect to videoconferencing. Not all patients are familiar with synchronous video communication, and not everyone has relatives to support them. One suggested solution was to extend the collaboration with homecare services to facilitate and support such meetings. However, this would create new work tasks for health care personnel in a care organization, which raises the aforementioned question of reimbursement and training of homecare personnel. Accordingly, scaling up the use of digital care has wide-ranging implications of a clinical, organizational, and economic character [11].

In sum, the COVID-19 crisis has shown that the health care service needs an alternative approach to face-to-face treatment and care. Based on the clinical perspectives in this study, the use of videoconferencing requires a long process of learning and adaptation, both individually and collectively, to ensure its successful integration into an organization [11]. The use of telemedicine must be recognized as an organizational development issue [3], in which the scaling up of videoconferencing creates challenges at different stages of its deployment [8].

Future Research

For the continued integration and sustainable use of videoconferencing to improve person-centered care, further research must be conducted that takes into consideration the complexities of individual patients, the need for tailored technical support, organizational factors, and economic issues. The first step is to interview PACT about its use of videoconferencing for more than the half-year following its rapid implementation and to investigate the use of videoconferencing as the COVID-19 pandemic subsides in this part of the world. It is also important to interview patients and PACT collaborators, such as homecare services and GPs, to evaluate their experience.

Limitations

In this study, we interviewed a specific group of health care personnel, namely PACT. PACT's experience of using digital care relied extensively on the adoption of videoconferencing by both its collaborators and patients. Although different health care professions are represented in PACT, a more generic view of the experience of preserving a person-centered care approach during the rapid digitalization and use of videoconferencing would have been gained from its collaborating partners in other organizations, such as homecare services, and from GPs and specialists in hospitals. In addition, we did not interview any patients; therefore, all the patient-related responses are secondhand knowledge as reported by PACT. To describe the person-centered care focus, we would need to interview patients directly as well.

Conclusions

Scaling up the use of digital care has been a goal in Norwegian health care for years. The COVID-19 pandemic and the rapid adoption of digital care have created a unique opportunity to continue developing a health service for not only preserving but also improving the person-centered care approach for persons with complex long-term needs.

Using digital care has provided an opportunity for more time-saving and efficient follow-up and coordination of this patient group in terms of reduced traveling time for patients and the more efficient use of health care personnel in PACT, as well as for improving the information exchange between health care levels.

In this study, we found that digital care cannot replace face-to-face interaction between patients and health care personnel. All the PACT members underscored the importance of face-to-face patient interaction as the foundation of care and that videoconferencing was a supplement to enhance follow-up and coordination. However, scaling up the use of digital care requires overall agreement and collaboration among the different health care levels and actors in terms of training, guidelines for when to use digital care, procedures for its use, technical support, etc, to ensure the success of digital health care services. Videoconferencing can then contribute to better preparing the health care services for future scenarios, such as new waves of COVID-19 or similar outbreaks, where there may be a need for social distancing in public and health care services.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOC File , 46 KB - formative_v5i3e25220_app1.doc]

References

1. Rush KL, Howlett L, Munro A, Burton L. Videoconference compared to telephone in healthcare delivery: a systematic review. Int J Med Inform 2018 Oct;118:44-53. [doi: 10.1016/j.ijmedinf.2018.07.007] [Medline: 30153920]
2. Zanaboni P, Wootton R. Adoption of routine telemedicine in Norwegian hospitals: progress over 5 years. BMC Health Serv Res 2016 Dec 20;16:496 [FREE Full text] [doi: 10.1186/s12913-016-1743-5] [Medline: 27644324]
3. Alami H, Gagnon MP, Wootton R, Fortin JP, Zanaboni P. Exploring factors associated with the uneven utilization of telemedicine in Norway: a mixed methods study. BMC Med Inform Decis Mak 2017 Dec 28;17(1):180 [FREE Full text] [doi: 10.1186/s12911-017-0576-4] [Medline: 29282048]
4. Hong Z, Li N, Li D, Li J, Li B, Xiong W, et al. Telemedicine during the COVID-19 pandemic: experiences From western China. J Med Internet Res 2020 May 08;22(5):e19577 [FREE Full text] [doi: 10.2196/19577] [Medline: 32349962]
5. Klonoff DC. Using telemedicine to improve outcomes in diabetes—an emerging technology. J Diabetes Sci Technol 2009 Jul 01;3(4):624-628 [FREE Full text] [doi: 10.1177/193229680900300401] [Medline: 20144303]
6. Bashshur RL. Telemedicine effects: cost, quality, and access. J Med Syst 1995 Apr;19(2):81-91. [doi: 10.1007/bf02257059]
7. Stowe S, Harding S. Telecare, telehealth and telemedicine. Eur Geriatr Med 2010 Jun;1(3):193-197. [doi: 10.1016/j.eurger.2010.04.002]
8. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
9. Korona: Slik kommer du i gang med videokonsultasjon. Direktoratet for e-helse. URL: https://ehelse.no/aktuelt/korona-slik-kommer-du-i-gang-med-videokonsultasjon [accessed 2020-10-07]

10. Greenhalgh T, Wherton J, Shaw S, Morrison C. Video consultations for covid-19. BMJ 2020 Mar 12;368:m998. [doi: 10.1136/bmj.m998] [Medline: 32165352]
11. Davis SM, Jones A, Jaynes ME, Woodrum KN, Canaday M, Allen L, et al. Designing a multifaceted telehealth intervention for a rural population using a model for developing complex interventions in nursing. BMC Nurs 2020 Feb 04;19(1):9 [FREE Full text] [doi: 10.1186/s12912-020-0400-9] [Medline: 32042264]
12. Fagherazzi G, Goetzinger C, Rashid MA, Aguayo GA, Huiart L. Digital health strategies to fight COVID-19 worldwide: challenges, recommendations, and a call for papers. J Med Internet Res 2020 Jun 16;22(6):e19284 [FREE Full text] [doi: 10.2196/19284] [Medline: 32501804]
13. Vogeli C, Shields AE, Lee TA, Gibson TB, Marder WD, Weiss KB, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. J Gen Intern Med 2007 Dec;22 Suppl 3:391-395 [FREE Full text] [doi: 10.1007/s11606-007-0322-1] [Medline: 18026807]
14. Bayliss EA, Edwards AE, Steiner JF, Main DS. Processes of care desired by elderly patients with multimorbidities. Fam Pract 2008 Aug;25(4):287-293 [FREE Full text] [doi: 10.1093/fampra/cmn040] [Medline: 18628243]
15. Tinetti ME, Fried TR, Boyd CM. Designing health care for the most common chronic condition—multimorbidity. JAMA 2012 Jun 20;307(23):2493-2494 [FREE Full text] [doi: 10.1001/jama.2012.5265] [Medline: 22797447]
16. Wallace E, Salisbury C, Guthrie B, Lewis C, Fahey T, Smith SM. Managing patients with multimorbidity in primary care. BMJ 2015 Jan 20;350:h176. [doi: 10.1136/bmj.h176] [Medline: 25646760]
17. Zulman DM, Pal Chee C, Wagner TH, Yoon J, Cohen DM, Holmes TH, et al. Multimorbidity and healthcare utilisation among high-cost patients in the US Veterans Affairs Health Care System. BMJ Open 2015 Apr 16;5(4):e007771-e007771 [FREE Full text] [doi: 10.1136/bmjopen-2015-007771] [Medline: 25882486]
18. Wang L, Si L, Cocker F, Palmer AJ, Sanderson K. A systematic review of cost-of-illness studies of multimorbidity. Appl Health Econ Health Policy 2018 Feb;16(1):15-29. [doi: 10.1007/s40258-017-0346-6] [Medline: 28856585]
19. Berntsen GKR, Dalbakk M, Hurley JS, Bergmo T, Solbakken B, Spansvoll L, et al. Person-centred, integrated and pro-active care for multi-morbid elderly with advanced care needs: a propensity score-matched controlled trial. BMC Health Serv Res 2019 Oct 03;19(1):682 [FREE Full text] [doi: 10.1186/s12913-019-4397-2] [Medline: 31581947]
20. Valentijn PP, Vrijhoef HJM, Ruwaard D, Boesveld I, Arends RY, Bruijnzeels MA. Towards an international taxonomy of integrated primary care: a Delphi consensus approach. BMC Fam Pract 2015 May 22;16:64 [FREE Full text] [doi: 10.1186/s12875-015-0278-x] [Medline: 25998142]
21. Nolte E, McKee M. Caring for People with Chronic Conditions: A Health System Perspective. Berkshire, UK: Open University Press; 2008.
22. Muth C, van den Akker M, Blom JW, Mallen CD, Rochon J, Schellevis FG, et al. The Ariadne principles: how to handle multimorbidity in primary care consultations. BMC Med 2014 Dec 08;12(1):223 [FREE Full text] [doi: 10.1186/s12916-014-0223-1] [Medline: 25484244]
23. Smith SM, Cousins G, Clyne B, Allwright S, O'Dowd T. Shared care across the interface between primary and specialty care in management of long term conditions. Cochrane Database Syst Rev 2017 Feb 23;2:CD004910 [FREE Full text] [doi: 10.1002/14651858.CD004910.pub3] [Medline: 28230899]
24. Berntsen G, Høyem A, Lettrem I, Ruland C, Rumpsfeld M, Gammon D. A person-centered integrated care quality framework, based on a qualitative study of patients' evaluation of care in light of chronic care ideals. BMC Health Serv Res 2018 Jun 20;18(1):479 [FREE Full text] [doi: 10.1186/s12913-018-3246-z] [Medline: 29925357]
25. Bokhour BG, Fix GM, Mueller NM, Barker AM, Lavela SL, Hill JN, et al. How can healthcare organizations implement patient-centered care? Examining a large-scale cultural transformation. BMC Health Serv Res 2018 Mar 07;18(1):168 [FREE Full text] [doi: 10.1186/s12913-018-2949-5] [Medline: 29514631]
26. Luxford K, Safran DG, Delbanco T. Promoting patient-centered care: a qualitative study of facilitators and barriers in healthcare organizations with a reputation for improving the patient experience. Int J Qual Health Care 2011 Oct;23(5):510-515. [doi: 10.1093/intqhc/mzr024] [Medline: 21586433]
27. Bower P, Cartwright M, Hirani SP, Barlow J, Hendy J, Knapp M, et al. A comprehensive evaluation of the impact of telemonitoring in patients with long-term conditions and social care needs: protocol for the whole systems demonstrator cluster randomised trial. BMC Health Serv Res 2011 Aug 05;11:184 [FREE Full text] [doi: 10.1186/1472-6963-11-184] [Medline: 21819569]
28. Person-centred care made simple: what everyone should know about person-centred care. The Health Foundation. 2014 Oct. URL: https://www.health.org.uk/publications/person-centred-care-made-simple [accessed 2020-10-12]
29. Reuben DB, Tinetti ME. Goal-oriented patient care—an alternative health outcomes paradigm. N Engl J Med 2012 Mar 01;366(9):777-779. [doi: 10.1056/NEJMp1113631] [Medline: 22375966]
30. Mathers N, Paynton D. Rhetoric and reality in person-centred care: introducing the House of Care framework. Br J Gen Pract 2015 Dec 30;66(642):12-13. [doi: 10.3399/bjgp16x683077]
31. Braut G, Åbotsvik B, Dalbakk M, Spansvoll L, Berntsen G. «Sløyfa» ? en metode for analyse av risiko i klinisk arbeid i grenseflaten mellom spesialisthelsetjenesten og kommunehelsetjenesten. Tidsskrift for omsorgsforskning 2020;6?208 2020 Nov 20;6(2). [doi: 10.18261/issn.2387-5984-2020-02-16]

32. Dalbakk M, Spansvoll L. Årsrapport 2017 Prosjekt Pasientsentrert helsetjenesteteam (PSHT). 2019 Jan 18. URL: https://docplayer.me/143638297-Arsrapport-2017-prosjekt-pasientsentrert-helsetjenesteteam-psht-v-monika-dalbakk-og-lisbeth-spansvoll.html [accessed 2020-10-07]
33. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health 2012 Dec;40(8):795-805. [doi: 10.1177/1403494812465030] [Medline: 23221918]
34. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014;12(6):573-576 [FREE Full text] [doi: 10.1370/afm.1713] [Medline: 25384822]
35. May C, Montori VM, Mair FS. We need minimally disruptive medicine. BMJ 2009 Aug 11;339(aug11 2):b2803-b2803. [doi: 10.1136/bmj.b2803] [Medline: 19671932]

Abbreviations

3P: Patients and Professionals in Partnership
GP: general practitioner
NSE: Norwegian Centre for E-health Research
PACT: Patient-Centered Team
UNN: University Hospital of North Norway

Edited by G Eysenbach; submitted 23.10.20; peer-reviewed by E Børøsund, C Varsi, E van der Velde; comments to author 08.11.20; revised version received 23.11.20; accepted 27.02.21; published 05.03.21.

Please cite as:
Silsand L, Severinsen GH, Berntsen G
Preservation of Person-Centered Care Through Videoconferencing for Patient Follow-up During the COVID-19 Pandemic: Case Study of a Multidisciplinary Care Team
JMIR Form Res 2021;5(3):e25220
URL: https://formative.jmir.org/2021/3/e25220
doi:10.2196/25220
PMID:33646965

Original Paper

A Patient-Initiated Digital COVID-19 Contact Notification Tool (TellYourContacts): Evaluation Study

Kelechi S Okpara[1,2], PharmD; Jennifer Hecht[3,4], MPH; Dan Wohlfeiler[3], MJ, MPH; Matthew Prior[5], MPH; Jeffrey D Klausner[6], MD, MPH

[1]David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States
[2]Charles R Drew University of Medicine and Science, Los Angeles, CA, United States
[3]Building Healthy Online Communities, San Francisco, CA, United States
[4]San Francisco AIDS Foundation, San Francisco, CA, United States
[5]National Coalition of STD Directors, Washington, DC, United States
[6]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

Corresponding Author:
Kelechi S Okpara, PharmD
David Geffen School of Medicine
University of California, Los Angeles
10833 Le Conte Ave
Los Angeles, CA
United States
Phone: 1 (310) 825 6373
Email: kokpara@mednet.ucla.edu

Abstract

Background: Contact notification is a method used to control the spread of infectious disease. In this process, a patient who tests positive for an infectious disease and public health officials work to identify the patient's close contacts, notify them of their risk of possible exposure to the disease, and provide resources to facilitate the decreased spreading of disease. Contact notification can be done physically in person, via phone call, or digitally through the use of media such as SMS text messages and email. When alerts are made through the latter, it is called digital contact notification.

Objective: For this study, we aim to perform a preliminary evaluation of the use of the TellYourContacts website, a digital contact notification tool for COVID-19 that can be used confidentially and anonymously. We will gather information about the number of website users and message senders, the types of messages sent, and the geographic distribution of senders.

Methods: Patients who chose to get tested for COVID-19 and subsequently tested positive for the disease were alerted of their positive results through Curative Inc (a COVID-19 testing laboratory) and Healthvana (a results disclosure app). Included in the notification was a link to the TellYourContacts website and a message encouraging the person who tested positive for COVID-19 to use the website to alert their close contacts of exposure risk. Over the course of three months, from May 18, 2020, to August 17, 2020, we used Google Analytics and Microsoft Excel to record data on the number of website users and message senders, types of messages sent, and geographic distribution of the senders.

Results: Over the course of three months, 9130 users accessed the website and 1474 unique senders sent a total of 1957 messages, which included 1820 (93%) SMS text messages and 137 (7%) emails. Users sent messages from 40 US states, with the majority of US senders residing in California (49%).

Conclusions: We set out to determine if individuals who test positive for COVID-19 will use the TellYourContacts website to notify their close contacts of COVID-19 exposure risk. Our findings reveal that, during the observation period, each unique sender sent an average of 1.33 messages. The TellYourContacts website offers an additional method that individuals can and will use to notify their close contacts about a recent COVID-19 diagnosis.

(JMIR Form Res 2021;5(3):e23843) doi:10.2196/23843

KEYWORDS

patient-led digital contact notification; COVID-19; digital contact tracing; contact notification website

Introduction

The process of contact tracing for COVID-19 varies, depending on the facility. However, the core concept includes first identifying those with an infection. This is done either when an individual tests positive for an infectious disease or if there is belief that there is a high probability that an individual has an infection. Then begins the process of determining if there are any close contacts of the person with COVID-19 that are at risk for contracting the disease. With the assistance of the health care team, patients can be educated on how to identify their close contacts and notify their contacts themselves. Conversely, working with the patient, public health officials and members of the health care team can work together to determine who is a close contact and notify them. The notification of close contacts of their exposure risk can be done in person or via phone call, SMS text message, email, or any additional form of communication. Lastly, the goal of contact tracing and contact notification is to control and decrease disease spread. As a result, when alerting close contacts, resources such as testing, treatments, and educational counselling (including next steps and how to approach the suggested guidelines) can be supplied to facilitate the goal of avoiding the spread of disease [1,2].

Digital contact notification is a form of contact notification where technological media such as emails, SMS text messages, and smartphone apps are used to notify the close contacts of people who tested positive for an infectious disease, such as COVID-19, of their exposure risk [3]. It can allow for greater flexibility within the realm of contact tracing, as opposed to the use of solely in-person or telephone-based contact notification. The use of emails and text messages to notify contacts has been used most frequently in partner notification for HIV and other sexually transmitted diseases (STDs), such as gonorrhea, syphilis, and chlamydia [4-11]. Previous studies on both patient-initiated (where the individual positive for the disease alerts their contacts) and health care provider–initiated digital contact or partner notification have concluded that the use of emails and/or text messages to notify contacts increases the total number of contacts alerted when used in addition to traditional in-person, telephone-based, or written partner notification [5-9]. This method allows for the ability to reach populations that are inaccessible or hard to reach with traditional partner notification [5-7].

However, not all studies have demonstrated a high level of utilization of digital means of notification. In evaluations of inSPOT, a patient-initiated STD partner notification website where those who test positive for an STD can send electronic postcards to their partners, investigators of the project found that <10% of the observed participants sent messages to their partners through the website [12-14]. This could be due, in part, to patients' preference of notifying their contacts in person [12]. However, more recently, TellYourPartner, an updated version of inSPOT, recorded a high proportion of visitors to messages sent.

We performed a preliminary evaluation of TellYourContacts [15], a patient-initiated COVID-19 contact notification website where users who have tested positive for COVID-19 have the option of sending either text messages or emails to notify their close contacts, to observe the use of the website over a period of three months. As defined by the Centers for Disease Control and Prevention (CDC), a close contact is one who was within 6 feet of the infected person for a cumulative total of at least 15 minutes over a 24-hour period, starting two days before the positive case started experiencing symptoms or two days prior to specimen collection [16,17]. We determined the frequency of website use and the frequency, types, and geographic distribution of messages sent.

Methods

TellYourContacts Website

Building Healthy Online Communities (BHOC) and Team Klausner Saving Lives (TKSL) collaborated to create and build the TellYourContacts website, which officially launched on May 13, 2020. TellYourContacts is a rapid adaptation of TellYourPartner, a website on which those positive for an STD can anonymously notify their sex partners [18]. Given the rapid spread of COVID-19 and the serious nature of some symptoms, a number of BHOC partners requested the adaptation of TellYourPartner for COVID-19. Combined with a review of CDC information as well as input from on-the-ground providers, we developed the TellYourContacts website, following quickly evolving best practices on contact notification as they were known at the time.

Many individuals are first informed about the TellYourContacts website when they are notified of their positive COVID-19 test result by Curative Inc or Healthvana. Within that notification detailing the positive result, a link to the TellYourContacts website is also provided, encouraging patients to notify their close contacts of their risk of contracting COVID-19. Curative Inc is a COVID-19 testing laboratory that supports COVID-19 testing in states throughout the United States and sends patients an email about their COVID-19 test results. If positive, the email contains a link to the TellYourContacts website (Figure 1).

Figure 1. Curative Inc email sent to those with a positive COVID-19 test result, with a "notify" link to the TellYourContacts website (Florida).

Hi {{patient_name}},

This is an automatic notification about your recent COVID-19 test result from Curative. Your COVID-19 test was **POSITIVE**, meaning that the virus that causes COVID-19 was detected in the specimen you provided.

You can expect to be contacted by a member of the Florida Department of Health in your county in the coming days.

Many patients experience mild to moderate symptoms similar to the flu. These can be managed with rest, hydration, and appropriate doses of fever-reducing medication. If at any time you experience chest pain, shortness of breath, difficulty breathing, confusion, vomiting that will not stop, or if you feel that your symptoms are worsening in any way, please call 911 or go to your nearest Emergency Room.

Please isolate yourself and minimize any contact with others for at least 10 days. Please notify anyone with whom you may have been in close contact (> 15 minutes and < 6 feet) starting two days before your symptom onset, or if you asymptomatic, two days before you were tested. You may call, text or email them to advise them that they may have been exposed and should quarantine for 14 days since they last had close contact with you. They should consider getting tested if they develop symptoms. Please visit this link for isolation guidance and more information from the Florida Department of Health.

Here is information about your test:

Appointment: {{reference_number}}
Patient: {{patient_name}}
Collected at: {{collection_time}}
Result: **POSITIVE**
Released at: {{release_time}}

Click here to find additional COVID-19 resources for Floridians.

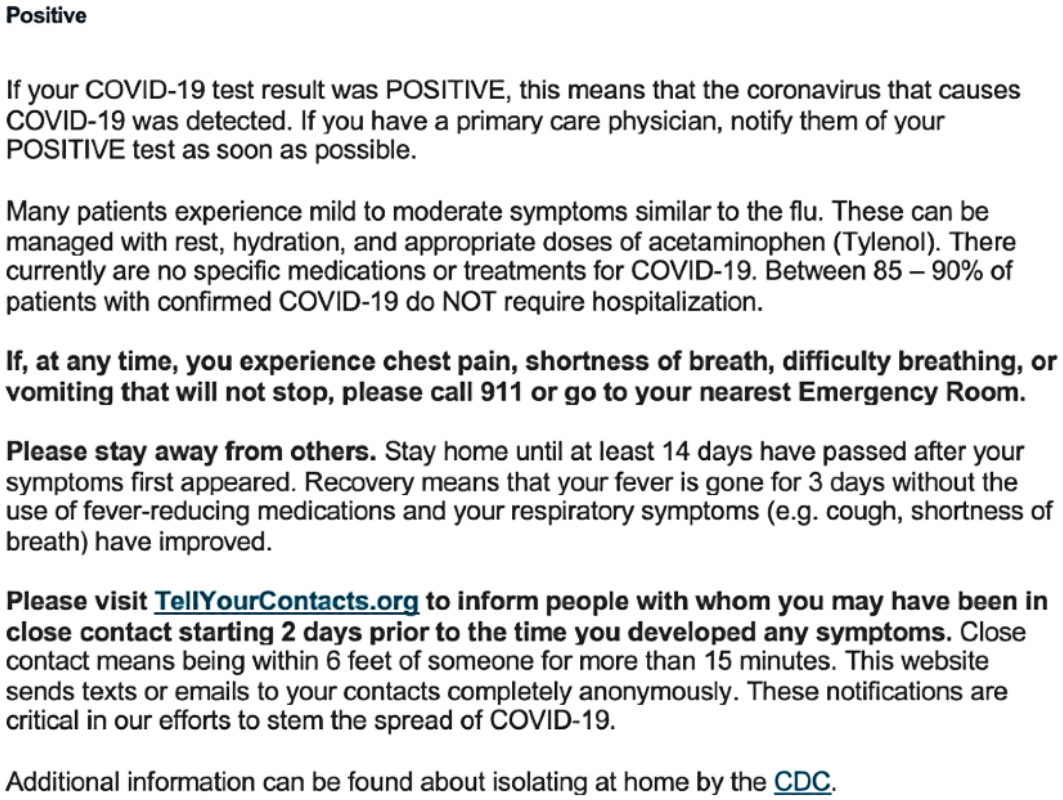
Be well,
Curative
http://curativeinc.com

Healthvana, a results disclosure app, delivers COVID-19 test results to patients located in California. Within the positive result message, Healthvana sends patients with COVID-19 the TellYourContacts website link [19] (Figure 2).

In addition, the TellYourContacts website was also advertised to members of CDC-funded STD organizations as well as through a press release, social media, and word of mouth.

Once on the TellYourContacts website, users have the option to either send their notification(s) via email or text message to up to 10 close contacts at a time. Additionally, they have the option to send their message(s) anonymously or confidentially, where the user has the option to enter their name. Users also have the option to create their own custom message or to send a prewritten message. The TellYourContacts message alerts the recipient that they may have been exposed to COVID-19 and it is recommended for them to isolate at home as well as seek testing. They are also directed to the CDC's coronavirus information webpage (Figures 3 and 4).

Figure 2. Healthvana message sent to those with a positive COVID-19 test result, with a link to the TellYourContacts website.



Positive

If your COVID-19 test result was POSITIVE, this means that the coronavirus that causes COVID-19 was detected. If you have a primary care physician, notify them of your POSITIVE test as soon as possible.

Many patients experience mild to moderate symptoms similar to the flu. These can be managed with rest, hydration, and appropriate doses of acetaminophen (Tylenol). There currently are no specific medications or treatments for COVID-19. Between 85 – 90% of patients with confirmed COVID-19 do NOT require hospitalization.

If, at any time, you experience chest pain, shortness of breath, difficulty breathing, or vomiting that will not stop, please call 911 or go to your nearest Emergency Room.

Please stay away from others. Stay home until at least 14 days have passed after your symptoms first appeared. Recovery means that your fever is gone for 3 days without the use of fever-reducing medications and your respiratory symptoms (e.g. cough, shortness of breath) have improved.

Please visit TellYourContacts.org to inform people with whom you may have been in close contact starting 2 days prior to the time you developed any symptoms. Close contact means being within 6 feet of someone for more than 15 minutes. This website sends texts or emails to your contacts completely anonymously. These notifications are critical in our efforts to stem the spread of COVID-19.

Additional information can be found about isolating at home by the CDC.

Figure 3. The TellYourContacts website offers prewritten and custom messages for users to send via SMS text message.

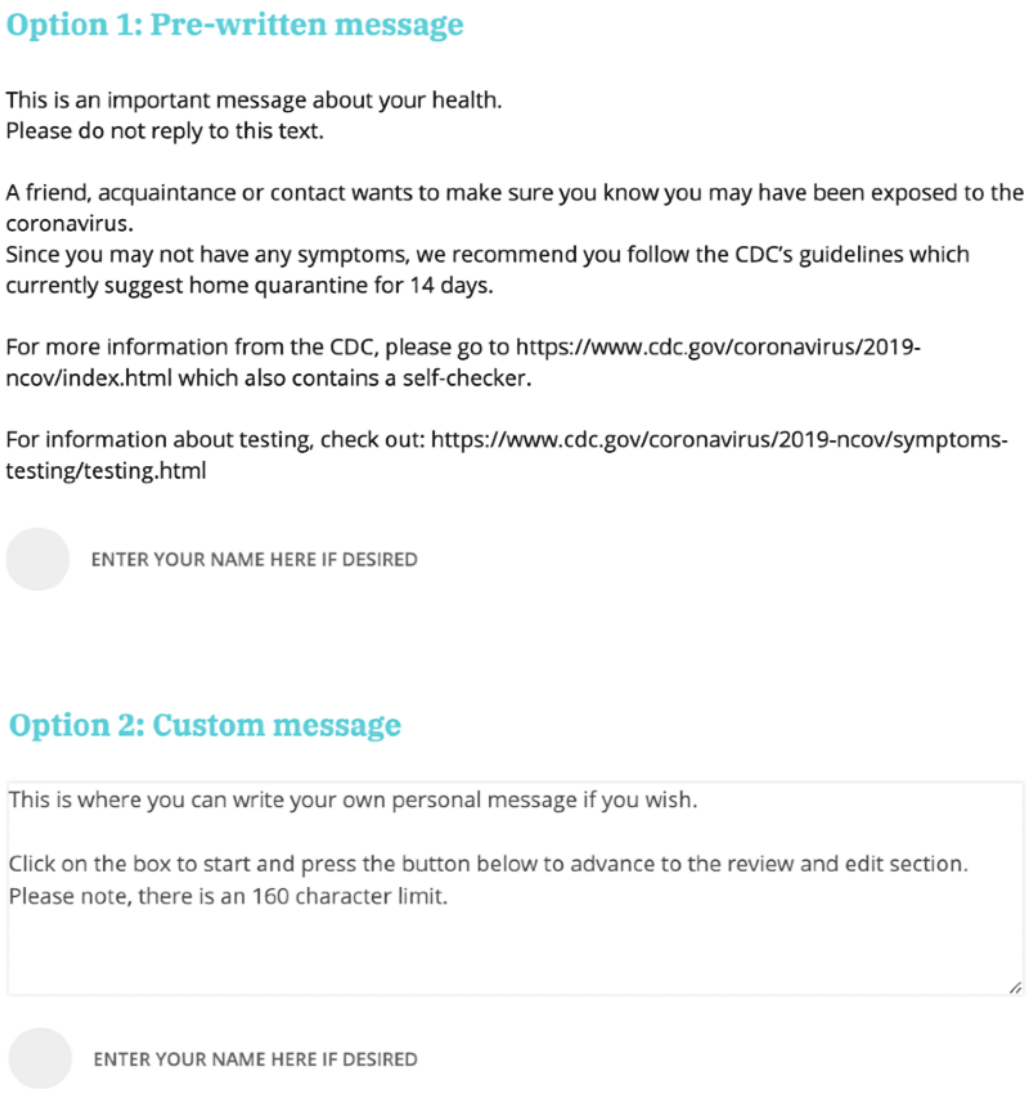
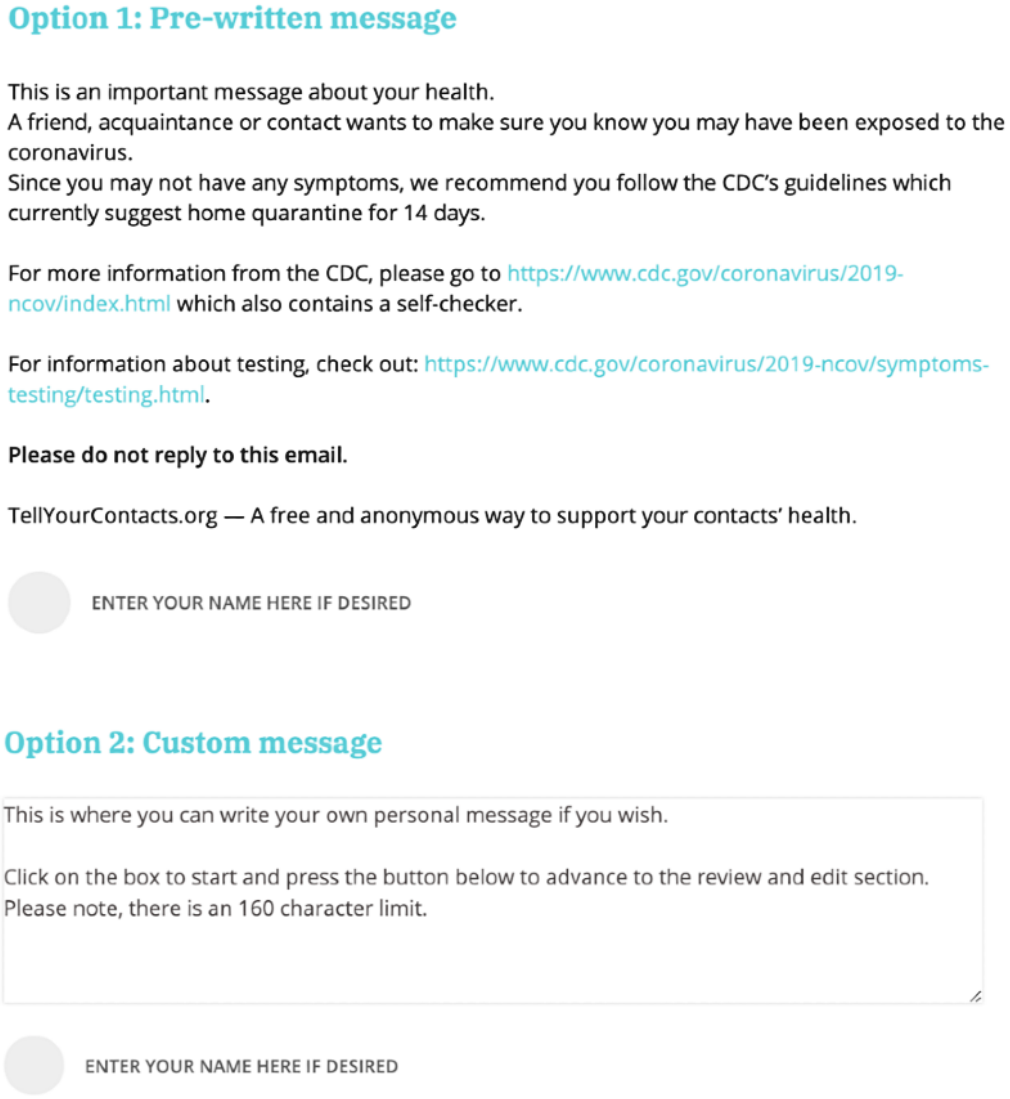


Figure 4. The TellYourContacts website offers prewritten and custom messages for users to send via email.



Lastly, the website provides information about which contacts should be notified and provides direct links to the CDC website for users to learn more about COVID-19 [15,20].

To maintain confidentiality and privacy, TellYourContacts does not record the user's contact information unless voluntarily provided. As users access the website and input information such as their contacts' email addresses and phone numbers, their connection and data are encrypted and secured via hypertext transfer protocol by the third-party service provider GoDaddy (GoDaddy Inc). The TellYourContacts website does not use cookies, though its third-party providers may. Additionally, the website neither stores the email addresses or phone numbers of the user's contacts nor does it link information to an individual. The website also uses services from Mailjet (Mailjet Inc), Twilio (Twilio Inc), and Google Analytics (Google) for email transmission, SMS text message transmission, and collection of aggregate data that includes the total number of users who access the website, the total number and types of messages sent, and the general location of users as determined by internet protocol (IP) address [21]. These providers have their own respective measures of security, which include ISO 27001 Information Security Management compliance and General Data Protection Regulation–compliant email service provider certification [22-24].

Data Collection and Analysis

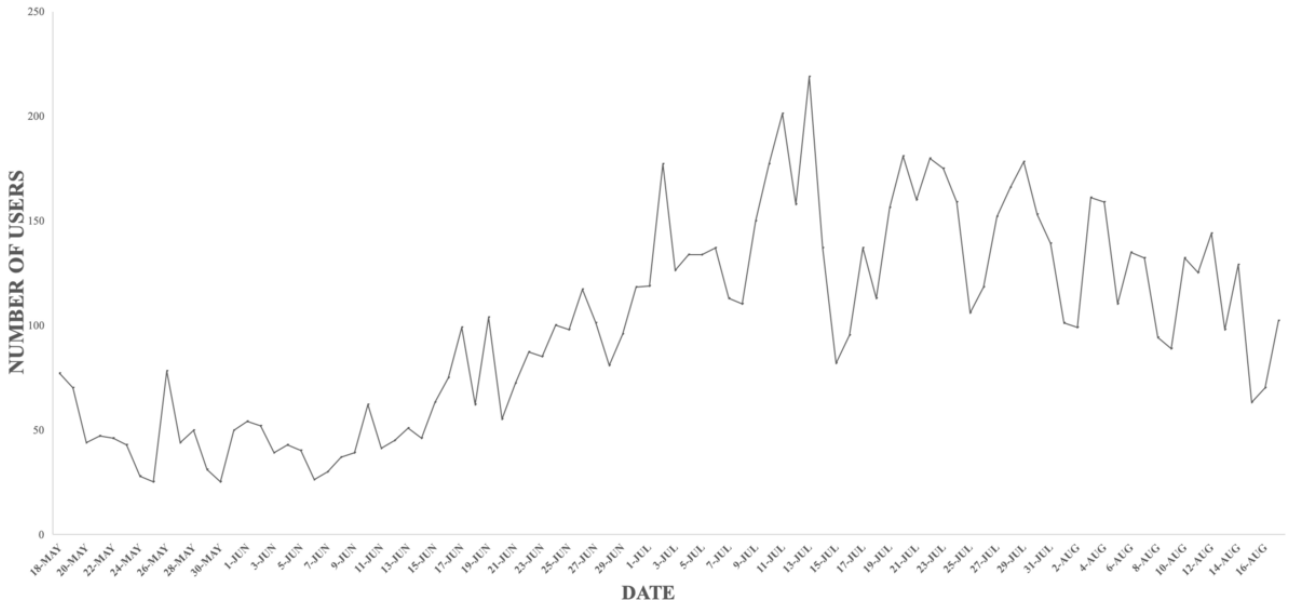
We collected data about the total number of emails and messages that were sent by both Curative Inc and Healthvana to notify patients of their positive COVID-19 results. From that, we estimated the number of messages sent to patients observed in our study. Additionally, we collected data regarding TellYourContacts website use including message senders, types of messages sent, and approximate geographic location of senders from Google Analytics. We input the data into Microsoft Excel software (version 16.16.12; Microsoft Corp) to sum the total quantities of the aggregate website data and used descriptive statistics to determine the mean number of messages sent per unique user within the observation period.

Results

From May 18, 2020, to August 17, 2020, Healthvana reported 125,652 positive COVID-19 test results to patients in California. Additionally, from May 18, 2020, to August 17, 2020, Curative Inc notified 71,866 patients of their positive COVID-19 test

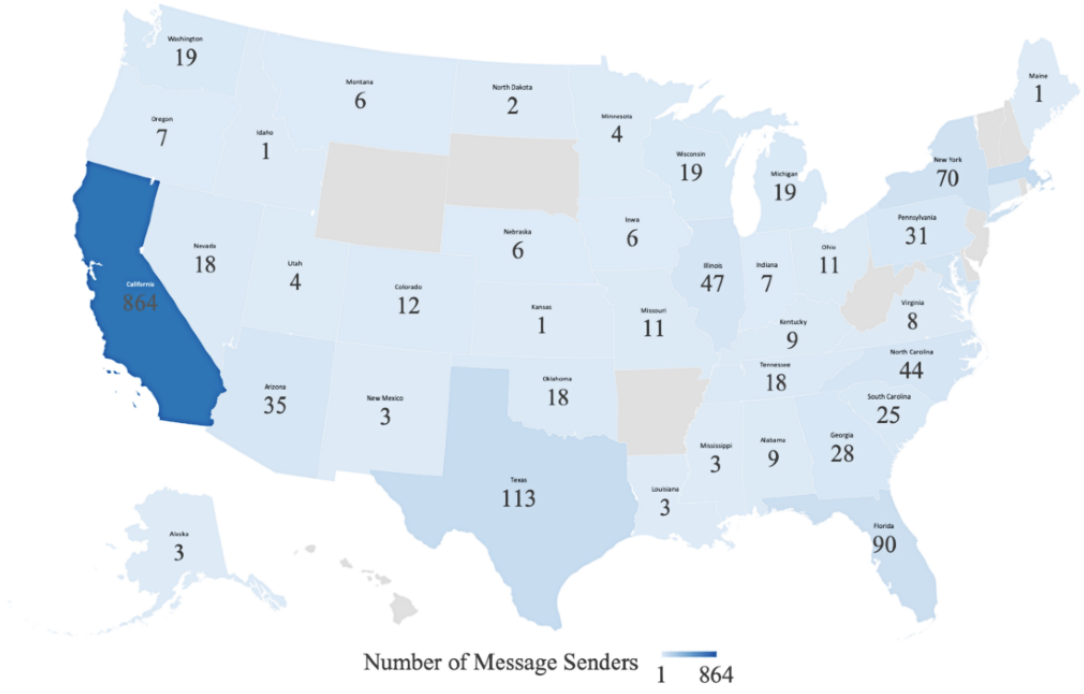
results across the United States. During the observation period, from May 18, 2020, until August 17, 2020, a total of 9161 users accessed the TellYourContacts website (Figure 5).

Figure 5. Number of TellYourContacts users per date, from May 18, 2020, to August 17, 2020.



Of these users, a total of 1957 sent messages. We found that users chose to send SMS text messages over emails in a 13:1 ratio (1820 [93%] via SMS text messaging and 137 [7%] via email). Messages were sent by users from the following 13 countries: the United States, Canada, Dominican Republic, Mexico, the United Kingdom, South Africa, Japan, Bulgaria, Puerto Rico, French Guiana, the United Arab Emirates, Sweden, Australia, and Morocco. Senders in the United States made up the majority of total senders (97%), representing 40 states. Within the United States, California accounted for the majority of senders (864/1754, 49%). Texas and Massachusetts accounted for the second-most senders, both with 113 (6%) each (Figure 6).

Figure 6. Number of TellYourContacts message senders by state, from May 18, 2020, to August 17, 2020.



Discussion

We described the use of a newly created digital notification platform, TellYourContacts, and found that over the course of three months, 9161 users accessed the website and 1474 unique senders sent a total of 1957 messages notifying their close contacts of possible COVID-19 exposure risk; the messages were sent from the majority of states in the United States. Our

results demonstrate that individuals will use a digital notification platform to notify their contacts.

Currently, there are various modalities being used for digital COVID-19 contact notification globally, such as smartphone-based contact tracing apps. These apps use either Bluetooth, GPS, or quick response (QR) codes to inform individuals that they may have been exposed to a case. Bluetooth-based contact notification apps measure the signal strength of nearby smartphones to determine the user's proximity to and duration of an encounter with those infected. GPS-based contact notification apps geolocate users to determine their proximity to those infected. Lastly, QR-based contact notification apps use QR barcodes that are scanned upon one's entrance to public spaces to monitor the user's visited locations [3]. In contrast to other forms of COVID-19 contact notification, such as the Apple and Google Exposure Notification System that uses Bluetooth to automatically notify close contacts that have opted in, TellYourContacts assists users in the process of contact notification while maintaining user agency [25]. It gives users the ability to control how they send their messages by providing them with the option to personalize their message or use preformed messages and anonymize themselves. This tool can be especially helpful for users who might not feel comfortable directly telling their contacts or who might not know how to directly tell their contacts about their exposure risk, all while providing their contacts with website links to COVID-19 educational resources. Additionally, TellYourContacts can be used without the need to download and install an app—which would require authorizing additional permissions and continuous updates—on a phone or other electronic device.

Users of TellYourContacts chose to send the majority of messages to their close contacts via text message. This result might possibly be due to the convenience and increasing popularity of text messaging over the past two decades. In 2019, 81% of Americans owned a smartphone, an increase of over 20% compared to five years prior [26]. Geographically, California accounted for the majority of the locations from where messages were sent. This may be because most positive cases observed in our study were tested in California.

This study is not without its limitations. First, our evaluation may have underestimated the number of people who may have received information about the TellYourContacts website because those who were tested at institutions such as nursing homes and homeless shelters were manually excluded from the total observed count. Second, the TellYourContacts website has the potential to be misused, as messages can be sent for unintended purposes. IP addresses are collected by Google Analytics, a third-party provider, to geolocate and provide relevant COVID-19–related information by region. Currently, an anti–IP address spoofing mechanism is in use to filter out false IP addresses associated with bots. Lastly, due to the nature and privacy policies of the website, data regarding the demographic characteristics of each person tested for COVID-19, the percentage of those who sent anonymous messages, and amount of contacts who read the message and took precautions such as isolation, symptom self-checking, and testing, cannot be obtained and analyzed.

With innovations in technology, 90% of Americans use the internet and 81% own a smartphone [26,27]. As a result, digital means such as emailing and text messaging are increasingly used to communicate. TellYourContacts gives individuals who are positive for COVID-19 the ability to initiate contact notification with tools they already use, such as smartphones and computers. After the evaluation of the TellYourContacts website, we found supportive evidence that TellYourContacts is a viable, low-threshold method that people will use to aid in contact notification for COVID-19. Continued social marketing research is needed to further increase its use.

Acknowledgments

The authors wish to thank Isaac Turner of Curative, Inc., Ramin Bastani of Healthvana, Gene Peterson of Bitrageous, LLC, Erin Fratto of National Coalition of STD Directors (NCSD), and Annabel Mangold of Mangold Design for their work and assistance with TellYourContacts. Funding for TellYourContacts.org was provided by Team Klausner Saving Lives (TKSL), Building Healthy Online Communities (BHOC), and NCSD.

Conflicts of Interest

None declared.

References

1. Centers for Disease Control and Prevention. Contact Tracing for COVID-19. URL: https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/contact-tracing.html [accessed 2020-08-07]
2. Cowan FM, French R, Johnson AM. The role and effectiveness of partner notification in STD control: a review. Genitourin Med 1996 Aug;72(4):247-252 [FREE Full text] [doi: 10.1136/sti.72.4.247] [Medline: 8976827]
3. Kleinman RA, Merkel C. Digital contact tracing for COVID-19. CMAJ 2020 Jun 15;192(24):E653-E656 [FREE Full text] [doi: 10.1503/cmaj.200922] [Medline: 32461324]
4. Levine D, Woodruff AJ, Mocello AR, Lebrija J, Klausner JD. inSPOT: the first online STD partner notification system using electronic postcards. PLoS Med 2008 Oct 21;5(10):e213 [FREE Full text] [doi: 10.1371/journal.pmed.0050213] [Medline: 18942887]

5. Hightow-Weidman L, Beagle S, Pike E, Kuruc J, Leone P, Mobley V, et al. 'No one's at home and they won't pick up the phone': using the Internet and text messaging to enhance partner services in North Carolina. Sex Transm Dis 2014 Feb;41(2):143-148. [doi: 10.1097/OLQ.0000000000000087] [Medline: 24413497]
6. Udeagu CN, Bocour A, Shah S, Ramos Y, Gutierrez R, Shepard CW. Bringing HIV partner services into the age of social media and mobile connectivity. Sex Transm Dis 2014 Oct;41(10):631-636. [doi: 10.1097/OLQ.0000000000000181] [Medline: 25211262]
7. Vest JR, Valadez AM, Hanner A, Lee JH, Harris PB. Using e-mail to notify pseudonymous e-mail sexual partners. Sex Transm Dis 2007 Nov;34(11):840-845. [doi: 10.1097/OLQ.0b013e318073bd5d] [Medline: 17621245]
8. Ehlman DC, Jackson M, Saenz G, Novak DS, Kachur R, Heath JT, et al. Evaluation of an innovative internet-based partner notification program for early syphilis case management, Washington, DC, January 2007-June 2008. Sex Transm Dis 2010 Aug;37(8):478-485. [doi: 10.1097/OLQ.0b013e3181e212cb] [Medline: 20539261]
9. Clark JL, Segura ER, Oldenburg CE, Salvatierra HJ, Rios J, Perez-Brumer AG, et al. Traditional and Web-Based Technologies to Improve Partner Notification Following Syphilis Diagnosis Among Men Who Have Sex With Men in Lima, Peru: Pilot Randomized Controlled Trial. J Med Internet Res 2018 Jul 03;20(7):e232 [FREE Full text] [doi: 10.2196/jmir.9821] [Medline: 29970355]
10. Hopkins CA, Temple-Smith MJ, Fairley CK, Pavlin NL, Tomnay JE, Parker RM, et al. Telling partners about chlamydia: how acceptable are the new technologies? BMC Infect Dis 2010 Mar 09;10:58 [FREE Full text] [doi: 10.1186/1471-2334-10-58] [Medline: 20211029]
11. Gilbart VL, Town K, Lowndes CM. A survey of the use of text messaging for communication with partners in the process of provider-led partner notification. Sex Transm Infect 2015 Mar;91(2):97-99. [doi: 10.1136/sextrans-2014-051701] [Medline: 25352690]
12. Rietmeijer CA, Westergaard B, Mickiewicz TA, Richardson D, Ling S, Sapp T, et al. Evaluation of an online partner notification program. Sex Transm Dis 2011 May;38(5):359-364. [doi: 10.1097/OLQ.0b013e31820ef796] [Medline: 21343844]
13. Kerani RP, Fleming M, DeYoung B, Golden MR. A randomized, controlled trial of inSPOT and patient-delivered partner therapy for gonorrhea and chlamydial infection among men who have sex with men. Sex Transm Dis 2011 Oct;38(10):941-946. [doi: 10.1097/OLQ.0b013e318223fcbc] [Medline: 21934569]
14. Plant A, Rotblatt H, Montoya JA, Rudy ET, Kerndt PR. Evaluation of inSPOTLA.org: an Internet partner notification service. Sex Transm Dis 2012 May;39(5):341-345. [doi: 10.1097/OLQ.0b013e31824e5150] [Medline: 22504593]
15. TellYourContacts.org. URL: https://tellyourcontacts.org/ [accessed 2020-07-24]
16. Why use this site. TellYourContacts.org. URL: https://tellyourcontacts.org/why-use-this-site/ [accessed 2020-08-13]
17. Contact tracing. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/appendix.html#contact [accessed 2021-01-29]
18. Tell Your Partner. Accessed August 13, 2020. URL: https://tellyourpartner.org/about/ [accessed 2020-08-13]
19. Healthvana. URL: https://healthvana.com/faq/ [accessed 2020-07-24]
20. NCSD Launches TellYourContacts.org, A Cutting-Edge Tool Allowing COVID-19 Cases To Notify Contacts. NCSD. 2020 May 13. URL: https://www.ncsddc.org/ncsd-launches-tellyourcontacts-org-allowing-covid-19-cases-to-notify-contacts-in-cutting-edge-effort-to-contain-the-virus/ [accessed 2020-07-24]
21. Website Privacy Policy. TellYourContacts.org. URL: https://tellyourcontacts.org/wp-content/uploads/2020/04/TYC-privacy-policy_4.22.20.pdf [accessed 2021-01-20]
22. Twilio Cloud Communications Security. URL: https://s3.amazonaws.com/ahoy-assets.twilio.com/Whitepapers/Twilio_Whitepaper_Security-Architecture.pdf [accessed 2021-01-20]
23. Mailjet Becomes the First ISO 27001 Certified and GDPR-Compliant Email Service Provider. Mailjet. URL: https://www.mailjet.com/gdpr/mailjet-first-esp-iso-27001-and-gdpr-certified/ [accessed 2021-01-20]
24. Safeguarding Your Data - Analytics Help. Google Analytics. URL: https://support.google.com/analytics/answer/6004245#zippy=%2Cdata-privacy-and-security [accessed 2021-01-20]
25. Exposure Notifications: Helping fight COVID-19. Google. 2020. URL: https://www.google.com/covid19/exposurenotifications/ [accessed 2020-08-21]
26. Demographics of Mobile Device Ownership and Adoption in the United States. Pew Research Center: Internet & Technology. 2020 Jun 05. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2020-07-24]
27. Demographics of Internet and Home Broadband Usage in the United States. Pew Research Center: Internet & Technology. 2020 Jun 05. URL: https://www.pewresearch.org/internet/fact-sheet/internet-broadband/ [accessed 2020-07-21]

Abbreviations

BHOC: Building Healthy Online Communities
CDC: Centers for Disease Control and Prevention
IP: internet protocol

NCSD: National Coalition of STD Directors
QR: quick response
STD: sexually transmitted disease
TKSL: Team Klausner Saving Lives

Edited by G Fagherazzi; submitted 25.08.20; peer-reviewed by B Suffoletto, B Bente; comments to author 09.12.20; revised version received 29.01.21; accepted 10.02.21; published 05.03.21.

Please cite as:
Okpara KS, Hecht J, Wohlfeiler D, Prior M, Klausner JD
A Patient-Initiated Digital COVID-19 Contact Notification Tool (TellYourContacts): Evaluation Study
JMIR Form Res 2021;5(3):e23843
URL: https://formative.jmir.org/2021/3/e23843
doi:10.2196/23843
PMID:33621189

Original Paper

Internet Access and Usage Among Stroke Survivors and Their Informal Caregivers: Cross-sectional Study

Imama Ali Naqvi[1], MD; Tahani Casameni Montiel[2], BBA; Yazan Bittar[3], BA; Norma Hunter[4], MA; Munachi Okpala[5], MBA, MSN; Constance Johnson[6], RN, PhD; Mark G Weiner[7], MD; Sean Savitz[5], MD; Anjail Sharrief[5], MPH, MD; Jennifer Elizabeth Sanner Beauchamp[8], RN, PhD

[1]Division of Stroke and Cerebrovascular Diseases, Department of Neurology, Columbia University Irving Medical Center, New York, NY, United States

[2]Department of Nursing Research, Cizik School of Nursing, The University of Texas Health Science Center at Houston, Houston, TX, United States

[3]Lewis Katz School of Medicine, Temple University, Philadelphia, PA, United States

[4]Department of Neurology, Institute for Stroke and Cerebrovascular Disease, The University of Texas Health Science Center at Houston, Houston, TX, United States

[5]Department of Neurology, Institute for Stroke and Cerebrovascular Disease, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX, United States

[6]Department of Nursing Research, Cizik School of Nursing and School of Bioinformatics, The University of Texas Health Science Center at Houston, Houston, TX, United States

[7]Department of Population Health Sciences, Weill Cornell Medicine, New York, NY, United States

[8]Department of Nursing Research, Cizik School of Nursing, Institute for Stroke and Cerebrovascular Disease, The University of Texas Health Science Center at Houston, Houston, TX, United States

Corresponding Author:
Imama Ali Naqvi, MD
Division of Stroke and Cerebrovascular Diseases
Department of Neurology
Columbia University Irving Medical Center
710 W 168th St
6th Fl, Ste 650
New York, NY, 10032
United States
Phone: 1 212 305 8253
Email: ian2108@cumc.columbia.edu

Abstract

Background: Web-based interventions have shown promise for chronic disease management but have not been widely applied to populations with stroke. Existing barriers may inhibit the adoption of web-based interventions among stroke survivors and necessitate the involvement of informal caregivers. However, limited information is available on internet accessibility and usability among stroke survivors and their caregivers.

Objective: This study aims to investigate internet access and usage in a cohort of stroke survivors and their caregivers.

Methods: A cross-sectional survey was conducted with 375 participants (248 stroke survivors and 127 caregivers). Descriptive statistics were generated using cross-tabulation. Comparisons with categorical data were conducted using the chi-square test, whereas the Mann-Whitney *U* test was used for comparisons involving ordinal variables.

Results: Overall, 86.1% (323/375) of the participants reported having internet access. Caregivers were more likely than stroke survivors to access the internet (N=375, χ^2_1=18.5, *P*<.001) and used text messaging (n=321, χ^2_1=14.7, *P*<.001). Stroke survivors and caregivers with internet access were younger than stroke survivors and caregivers without internet access. The highest number of participants who reported internet access were non-Hispanic White. Smartphones were the most common devices used to access the internet. Email was the most common type of internet usage reported. Patients who survived for >12 months after a stroke reported higher internet access than those who survived <3 months (*P*<.001). The number of hours per week spent using the internet was higher for caregivers than for stroke survivors (*P*<.001).

Conclusions: Future feasibility and acceptability studies should consider the role of the informal caregiver, participant age, race and ethnicity, the use of smartphone apps, email and text correspondence, and the amount of time elapsed since the stroke event in the design and implementation of web-based interventions for populations with stroke.

(JMIR Form Res 2021;5(3):e25123) doi:10.2196/25123

KEYWORDS

internet access; stroke; caregivers; surveys; questionnaires; mobile phone

Introduction

Background

In the United States, 795,000 people experience stroke annually [1]. Although improvements in the acute management of stroke have led to a decline in associated mortality, stroke-related morbidity leads to chronic disability in approximately half of all stroke survivors [2]. Comprehensive poststroke interventions should consider common stroke sequelae, including functional disabilities (eg, limb paralysis and sensory disturbances), speech disabilities (eg, aphasia), emotional disturbances (eg, poststroke anxiety and depression), and cognitive impairments (eg, impaired memory) [1]. Mobility is reduced in more than half of all stroke survivors aged ≥65 years [3]. Difficulty in producing and understanding speech, known as aphasia, occurs in an estimated 25%-40% of stroke survivors [4]. The prevalence of poststroke anxiety and depression ranges from 21% to 29% and from 29% to 31%, respectively [5-7]. Stroke is the second most common cause of cognitive impairment and dementia, with approximately 30% of stroke survivors experiencing cognitive impairment [8] or dementia [9]. Despite the complex needs of stroke survivors, poststroke care systems are inadequate. Stroke survivors may have limited access to outpatient care because of impaired mobility, limited access to transportation, and lack of support [10]. Most stroke survivors are discharged home from the hospital and receive care provided primarily by unprepared informal caregivers (eg, spouses and family members) [11,12]. Measures that expand access to poststroke care and comprehensively address the challenges that stroke survivors and their caregivers encounter are needed [13].

Web-based telehealth interventions have been found to be effective for acute stroke care, potentially beneficial for extended neurology care in rural areas [14], and cost-effective in stroke and dementia populations [15,16].

Purpose

Although a number of studies have examined internet access and usage among populations other than those with stroke (eg, patients with diabetes) [17], studies are needed to determine the feasibility and acceptability of web-based interventions across diverse stroke populations with complex disabilities and various levels of ability [18]. Uncertainties remain regarding the implementation of web-based interventions, including suitable stroke survivors and necessary stroke survivor support structures [19], as well as the role of the caregiver. Only 67% of US adults aged ≥65 years reported internet access compared with 44% of adults aged ≥80 years [20]. The lower rates of internet access with advancing age present an additional challenge, as an estimated three-fourths of all strokes occur in adults aged ≥65 years [21] and 17% of all strokes occur in adults aged >85 years [22]. Non-Hispanic Black and Hispanic US adults are less likely than non-Hispanic White adults to have access to the internet in their home environment [23]. The highest increase in stroke prevalence (29%) is estimated to be reported among Hispanic men [1], and known racial and ethnic disparities occur in almost every aspect of stroke care [24]. To develop appropriate and accessible web-based interventions for stroke survivors and caregivers, we must fully understand these digital disparities within the context of stroke survivorship. Therefore, the purpose of our study was to investigate internet access and usage in a cohort of stroke survivors and their caregivers.

Methods

Design and Sampling

This was an observational study of cross-sectional survey data collected from a convenience sample of stroke survivors and their caregivers from 2 large metropolitan areas: Houston, Texas, and Philadelphia, Pennsylvania. Institutional review board (IRB) approval and authorization for data sharing were obtained from both participating universities. As the only record linking the participant and the research was the informed consent document, the IRBs waived the requirement for written informed consent. Each participant received a letter of information outlining the study, and completion of the survey was taken as the form of consent to participate.

Recruitment

Recruitment sites included outpatient clinics, inpatient units, and community support groups. Participants were recruited during a routine visit to an outpatient stroke clinic within the Texas Medical Center (Houston, TX). The outpatient stroke clinic manages care of racially and ethnically diverse stroke survivors and treated approximately 1000 new stroke survivors in 2019. Participants were also recruited at an annual Houston community stroke festival and 2 Houston-area stroke support groups to supplement recruitment. In Philadelphia, Pennsylvania, participants were recruited from a comprehensive stroke center (CSC) and affiliated outpatient stroke clinic. The CSC serves a diverse minority and medically underserved population and treated approximately 600 patients with acute stroke in 2019. Screening measures to determine eligibility criteria included accessing electronic health records and participant self-reports. Identifiers (eg, names) collected for approaching potential participants were not retained.

Eligibility criteria for patients from both participating universities were as follows: patients who (1) were aged ≥18 years, (2) spoke and read English or Spanish, and (3) had a

history of stroke or self-identified as an informal caregiver of a stroke survivor. In Houston, surveys were collected in person using an Apple iPad and Research Electronic Data Capture (REDCap) [25,26] survey links or paper surveys and manually entered into REDCap [25,26] by a trained research member. Surveys with the survey links were also emailed or completed by telephonic interviews by a trained research member who manually entered the data into REDCap [25,26]. Surveys were carried out from September 2018 to July 2019. In Philadelphia, in-person surveys were collected directly via REDCap [25,26] survey links on an iPad. Surveys were carried out from March 2019 to July 2019. All surveys were assigned a study identifier without personal identifiers. The mode of data collection was dependent on user comfort and iPad availability. All data were collected when trained surveyors were available, except for the Houston site, which also emailed REDCap [25,26] survey links for participants to complete. REDCap [25,26] is a secure, web-based app that supports data capture and export procedures at both universities.

Variables

The survey was not intended to collect psychometric data and thus did not rely on a validated psychometric instrument. However, contributions from experts in neurology, nursing, and bioinformatics were used to create a 14-question survey in English and Spanish languages (Multimedia Appendix 1). In total, 8 demographic questions included variables such as gender, race, ethnicity, and health insurance status. The duration in months from stroke events for stroke survivors and relationship (eg, spouse) between stroke survivors and caregivers were obtained. A total of 6 internet access and usage questions included an inquiry into the form of internet access at home, including cellular phone data as well as types of devices used to access the internet. Modes of communication, including email, text messages, web browsing, and gaming interactions, were included as discrete queries. Time spent and the language predominately used while on a device used to access the internet were separated for choice.

Statistical Analysis

Data were analyzed using SPSS 25.0. Participants who selected being both stroke survivors and caregivers were counted as stroke survivors for the analysis. Descriptive data were generated to explore trends in internet access and usage through cross-tabulation. Comparisons with categorical data were done using the chi-square test, whereas the Mann-Whitney U test was used for comparisons involving ordinal variables. On the basis of a sample size of 127, a one-sample chi-square test had 80% power when the hypothesized proportion of internet users was 0.9, and the proportion of internet users in the sampled population was 0.8 [27]. As we planned to test this hypothesis in both stroke survivors and caregivers, we recruited 248 stroke survivors and 127 caregivers.

Results

Sample Characteristics of Stroke Survivors and Informal Caregivers

Of the 397 surveys collected, 375 were analyzed (Table 1). Overall, 22 surveys were excluded because 19 participants did not indicate stroke survivors or caregiver status, and 3 participants did not complete the internet access and usage questions. Of the 375 surveys, 248 (66.1%) were from stroke survivors and 127 (33.9%) were from caregivers. Overall, 45.3% (169/373) of the participants were male, and 89.4% (329/368) reported having health insurance. Most participants (54/127, 42.5%) reported a spousal caregiver relationship to the care recipient, followed by a child caregiver relationship (39/127, 30.7%) to the parent care recipient. Participants (318/375, 84.8%) were primarily recruited from outpatient sites. A total of 107 of the 375 (28.5%) participants surveyed were from Philadelphia sites, including 58 (54.2%) from inpatient units and 49 (45.8%) from an outpatient clinic.

Table 1. Comparison of demographic characteristics of stroke survivors and informal caregivers (N=375) by internet access.

Characteristics	Total participants[a]	With internet access[b] (n=323, 86.1%)		Without internet access[b] (n=52, 14%)	
		Informal caregivers (n=123)	Stroke survivors (n=200)	Informal caregivers (n=4)	Stroke survivors (n=48)
Age (years), mean (SD)	58	51 (14)	59 (14)	60 (19)	69 (12)
Sex (male), n (%)	169 (45.3)	33 (26.8)	108 (54.3)	1 (25)	27 (57.4)
Health insurance, n (%)	329 (89.4)	100 (82)	185 (94.9)	4 (100)	40 (85)
Race and ethnicity, n (%)					
Non-Hispanic Black	132 (35.3)	35 (29)	71 (36)	1 (33)	25 (52)
Non-Hispanic White	114 (30.4)	48 (39)	63 (32)	—[c]	3 (6)
Hispanic	91 (24)	31 (25)	44 (22)	2 (67)	14 (29)
Other	37 (10)	9 (7)	22 (11)	—	6 (13)
Time since stroke (months), n (%)					
<3	73 (30)	—	48 (24)	—	25 (57)
3-6	33 (14)	—	28 (14)	—	5 (11)
6-12	33 (14)	—	29 (15)	—	4 (9)
>12	102 (42.3)	—	92 (47)	—	10 (23)
Relationship[d], n (%)					
Parent	11 (9)	9 (7)	—	2 (50)	—
Spouse	54 (43)	52 (42)	—	2 (50)	—
Son or daughter	39 (31)	39 (32)	—	—	—
Sibling	6 (5)	6 (5)	—	—	—
Grandchild	4 (3)	4 (3)	—	—	—
Friend	13 (11)	13 (11)	—	—	—

[a]Values in this column represent the number of participants who answered the specific demographic survey item.

[b]Counts may not add up to the n value indicated in the column header because of missing data. Percentages may not add up to 100% because of rounding off.

[c]Not available.

[d]Caregiver's relationship with the care recipient.

Stroke survivors with internet access were younger (mean 59, SD 14 years) than stroke survivors without internet access (mean 69, SD 12 years). Similarly, caregivers with internet access were younger (51, SD 14 years) than those without internet access (mean 60, SD 19 years). Overall, 36.4% (71/197) non-Hispanic Black, 31.9% (63/197) non-Hispanic White, and 22.3% (44/197) Hispanic stroke survivors reported internet access. Fifty two percent (25/48) non-Hispanic Black, 6.2% (3/48) non-Hispanic White, and 29.1% (14/48) Hispanic stroke survivors reported no internet access. Of those that reported internet access, 28.5% (35/123) non-Hispanic Black, 39% (48/123) non-Hispanic White, and 25.2% (31/123) Hispanic caregivers reported internet access. One (1/3, 33%) non-Hispanic Black and two (2/3, 67%) Hispanic caregivers reported no internet access. More stroke survivors with stroke events >12 months ago (n=197) had internet access as compared with those with stroke events <3 months ago (n=44). The difference in internet access between these 2 groups of stroke survivors was statistically significant (U=2782.0, z=−3.93, P<.001).

Characteristics of Stroke Survivor and Informal Caregiver Internet Users

Overall, 86.1% (323/375) of participants reported internet access (Table 2). Compared with an estimate that 89% of American adults have internet access [27,28], 80.6% (200/248) of stroke survivors (P<.001) and 96.8% (123/127) of caregivers (P=.001) had access. Caregivers were more likely than stroke survivors to access the internet (N=375, χ^2_1=18.5, P<.001). Smartphones were the most common type of device used to access the internet. Of the stroke survivors, 82.5% (165/200) reported using a smartphone, 59.5% (119/200) reported using a computer, and 40.5% (81/200) reported using an iPad to access the internet. Similarly, 91.1% (112/123) caregivers reported using a smartphone, 78% (96/123) reported using a computer, and 57.7% (71/123) reported using an iPad to access the internet. Email was the most common type of internet usage reported among stroke survivors (143/200, 71.2%) and caregivers (110/123, 89.4%), followed by browsing the web (stroke survivors=122/200, 61%; caregivers=104/123, 84.6%) and video games (stroke survivors=40/200, 20%; caregivers=42/123,

34.1%). The number of hours per week spent using the internet by caregivers (n=122) was higher than that of stroke survivors (n=194), and the difference was statistically significant (U=8922.00, z=−3.81, P<.001). The majority of stroke survivors (177/200, 88.5%) and caregivers (116/123, 94.3%) reported English as the primary language used in their devices. Caregivers were more likely to use text messaging than stroke survivors (n=321, χ^2_1=14.74, P<.001).

Table 2. Internet usage characteristics of stroke survivors and informal caregivers.

Internet usage characteristics	Total participants with internet access[a] (n=323)	
	Informal caregivers (n=123)	Stroke survivors (n=200)
Type of device, n (%)		
Smartphone	112 (91.1)	165 (82.5)
Computer	96 (78)	119 (59.5)
iPad or tablet	71 (58)	81 (40.5)
Other	8 (7)	8 (4)
Do not access the internet[b]	—[c]	4 (2)
Internet usage, n (%)		
Email	110 (89.4)	143 (71.2)
Web pages	104 (84.6)	122 (61)
Video games	42 (34)	40 (20)
Other	30 (24)	40 (20)
Do not access the internet[b]	1 (1)	23 (12)
Internet hours per week (hours), n (%)		
0-5	27 (22)	87 (45)
6-10	28 (23)	38 (20)
11-15	25 (20)	25 (13)
16-20	14 (12)	13 (7)
21-25	6 (5)	5 (3)
>25	22 (18)	26 (13)
Device language used, n (%)		
English	116 (94.3)	177 (88.5)
Spanish	14 (11)	25 (12.5)
Other	4 (3)	6 (3)
Text messaging		
Yes, n (%)	104 (84.6)	121 (60.5)

[a]Counts may not add up to the n values indicated in the column header because of missing data. Percentages may not add up to 100% because of rounding off.

[b]Home internet access is available but does not personally access the internet from home.

[c]Not available.

Discussion

Principal Findings

Overall, most of the sampled stroke population reported having internet access. Compared with a national estimate of adults with internet access, fewer stroke survivors reported internet access, whereas a greater number of caregivers reported access. Caregivers were more likely to access the internet and spend more time per week using the internet than stroke survivors. Stroke survivors and caregivers with internet access were younger than those without internet access. Internet access was significantly higher in stroke survivors more than 12 months after stroke than in stroke survivors less than 3 months after stroke. Smartphones were the most common devices used to access the internet.

Comparison With Prior Work

Overall, 85% (323/375) of the participants reported internet access, which is lower than that of the general public [27,29], but higher than the 72% for US adult internet users living within the confines of a chronic condition [30]. Similar to our findings,

mobile devices and computers are commonly used platforms [31] for accessing the internet and may be feasible and acceptable platforms for providing web-based stroke recovery interventions. Optimizing web-based interventions to be accessible by smartphones and computers may increase accessibility, given the predominance of smartphone and computer access among this sample. Lesser internet technology use among aging patients and their caregivers compared with younger adults [30] may indicate that these platforms may be challenging for the older stroke population. However, stroke is no longer a chronic condition in older individuals alone. The increasing rates (men 41.5% and women 30%, aged 35-44 years) of acute ischemic stroke from 2003 to 2012 in young adults coexist with the increasing prevalence of traditional risk factors [32,33] and emphasize the importance of focusing on web-based stroke recovery efforts in younger adults. Ischemic stroke events have increased significantly in adults aged 18-54 years [32]. Furthermore, the perception of the *digital divide* based on advancing age is rapidly changing, as internet use becomes more pervasive in the United States. Adults in the United States are reporting internet usage during the COVID-19 pandemic, with 84% of individuals aged ≥50 years, 98% of individuals aged 30-49 years, and 100% of individuals aged 18-29 years of age reporting internet usage or owning a smartphone device [34]. It is likely that web-based approaches will become more feasible and acceptable with the changing exposures and needs of technologically diverse stroke populations.

Chronic stroke survivors (stroke event>12 months ago) reported the highest rates of internet access, whereas those less than 3 months poststroke had the least access. This suggests that access and ease of internet use may be more robust in chronic poststroke care. Web-based stroke recovery interventions may require personalized strategies for stroke survivors' recovery status and individual preferences [15]. Stroke recovery care may also be limited by physical and geographical barriers. Although not among stroke survivors, web-based telehealth visits have been successfully used in the general population and have been shown to improve blood pressure control in hypertensive patients [35]. These web-based visits may be used as a platform to capture poststroke patient-reported outcomes and implement interventions centered on secondary stroke prevention, such as risk factor control, medication adherence, and lifestyle modifications. The COVID-19 pandemic has highlighted the utility of web-based telehealth visits. For example, the rapid transition to telehealth visits for outpatient care among patients with neurological diseases has been implemented, allowing patients to communicate with health care providers via smartphones and other devices [36]. As web-based telehealth services expand, health care professionals are likely to learn more about the feasibility and accessibility of these web-based services across different populations with incentivization to assemble infrastructures for effective implementation of web-based poststroke care.

Because most stroke survivors are discharged home from the hospital, many only receive care from unprepared caregivers [37]. The amount of care provided by the caregiver to the stroke survivor appears to increase significantly immediately after hospital discharge and remains high throughout the first 12 months after the stroke [38]. Web-based interventions that actively engage caregivers may improve postacute stroke care. Caregivers were more likely to access the internet and spend more time using on it than stroke survivors. Caregivers of individuals with chronic conditions appear to use the internet for general purposes, to access health-related information and track health-related indicators (eg, weight), support, and services [30,31]. Caring for individuals with a chronic condition is considered a major life stressor that negatively affects the health and well-being of caregivers [39-42]. However, stroke survivor-informal caregiver dyad interventions have predominately focused on the health and well-being of stroke survivors rather than caring for oneself as a caregiver [40]. Although more studies are needed to determine the effectiveness of web-based interventions aimed at meeting the needs of caregivers [40,43], the American Heart Association or American Stroke Association recommends web-based stroke recovery interventions that meet the evolving needs of technologically advanced caregivers of stroke survivors [40]. Optimal stroke recovery requires web-based strategies that target the health and well-being of both stroke survivors and caregivers, with both being active participants [40].

The highest number of participants who reported internet access were non-Hispanic Whites. Racial differences in internet access and technology usage emphasize the need to address known disparities [28]. Racial and ethnic disparities also exist in stroke recovery care. Non-Hispanic Black and Hispanic participants receive less intensive stroke rehabilitation, education, and counseling than non-Hispanic White participants [44,45]. Minority groups, including non-Hispanic Black and Hispanic participants, have a higher risk of stroke [3]. The highest increase in stroke prevalence was observed in Hispanic men [3]. Non-Hispanic Black participants have a higher prevalence of uncontrolled blood pressure, which is the most important risk factor for stroke [46]. Notably, efforts to reduce racial and ethnic disparities in blood pressure control among stroke survivors have not been effective [47]. Web-based interventions to reduce stroke recurrence and improve risk factor control could address this gap in care by providing internet services and devices to individuals from the highest risk groups. Considerations for language barriers should also be given in web-based poststroke care.

Limitations

One study limitation is that the sample may not be representative of all US stroke survivors and their informal caregivers. Therefore, the results should be interpreted with caution.

Detailed information regarding socioeconomic status and urban or rural location was not obtained. Future studies of internet access and use should consider targeted oversampling of economically disadvantaged stroke populations and stroke survivors and caregivers living in remote areas with limited broadband connectivity. Socioeconomic details can be gauged by collecting individual- and neighborhood-level social determinants of health data. Geo-mapping using ZIP codes for areas with the highest socioeconomic disparities and geographic barriers to classify the type and range of web-based services rendered will be useful before system-wide implementation of

web-based stroke recovery interventions. The sample size was adequate, and the sample was racially and ethnically diverse; however, the nonprobability convenience sampling strategy may have resulted in selection bias. Stroke survivors and caregivers not surveyed were likely missed at random because efforts were limited to surveyor availability. Furthermore, the generalizability of the results is limited by the lack of nonparticipant data (eg, demographics and reasons). However, the sample came from 2 large urban areas providing data from considerably underinvestigated minority stroke survivor and caregiver populations. Although the survey was developed and edited by a multidisciplinary team, the internal consistency and content validity of the survey were not tested.

Conclusions

Web-based interventions following stroke should consider the role of the caregiver, participant age, race, ethnicity, the use of smartphone apps, email and text correspondence, and the amount of time since the stroke event. The results suggest that web-based interventions may be feasible and acceptable for certain stroke survivors and caregivers. Future feasibility and acceptability studies should consider these findings when designing and implementing web-based stroke recovery interventions to minimize barriers to access, tailor the intervention to maximize adherence, and target those most likely to use web-based resources.

Acknowledgments

The authors acknowledge and thoughtfully thank the University of Texas Health Science Center at Houston (UTHealth) Neurology Clinic, UTHealth Cizik School of Nursing, especially Tina Varughese, RN, and Andrea Ancer Leal, RN-BSN, LMSW, West Houston Stroke Warriors, Houston Aphasia Recovery Center, UTHealth Stomp Out Stroke Festival, Temple University Hospital Neurology Clinic staff, and all participating stroke survivors and informal caregivers. This work was supported by NIH/NCATS grants UL1 TR000445 and UL1 TR001105. UTHealth study data were collected and managed using REDCap hosted at the University of Texas School of Biomedical Informatics at Houston.

Authors' Contributions

IAN contributed to data curation; formal analysis; investigation; project administration; resources; software; validation; and writing, reviewing, and editing the manuscript. TCM contributed to software; writing, reviewing, and editing the manuscript; visualization; and project administration. YB contributed to data curation, software, and visualization. NH and MO contributed to data curation. CJ reviewed and edited the manuscript and contributed to survey development. MGW contributed to supervision and reviewed and edited the manuscript. SS contributed to supervision, reviewing and editing the article, and survey development. AS contributed to investigation, validation, supervision, reviewing and editing the manuscript, and survey development. JESB contributed to conceptualization, methodology, investigation, validation, supervision, reviewing and editing the manuscript, and survey development.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Internet Usage Survey Questions.
[DOCX File , 19 KB - formative_v5i3e25123_app1.docx]

References

1. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, et al. Heart Disease and Stroke Statistics—2020 Update: A Report From the American Heart Association. Circulation 2020 Mar 03;141(9) [FREE Full text] [doi: 10.1161/CIR.0000000000000757]
2. Published January 9, 2019. Stroke Fact Sheet|Data & Statistics|DHDSP|CDC. URL: https://www.cdc.gov/dhdsp/data_statistics/fact_sheets/fs_stroke.htm [accessed 2019-10-30]
3. Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, et al. Heart Disease and Stroke Statistics—2019 Update: A Report From the American Heart Association. Circulation 2019 Mar 05;139(10). [doi: 10.1161/CIR.0000000000000659]
4. National Aphasia Association. Aphasia FAQs. URL: https://www.aphasia.org/aphasia-faqs/ [accessed 2020-10-12]
5. Hackett M, Pickles K. Part I: frequency of depression after stroke: an updated systematic review and meta-analysis of observational studies. Int J Stroke 2014 Dec;9(8):1017-1025. [doi: 10.1111/ijs.12357] [Medline: 25117911]
6. Rafsten L, Danielsson A, Sunnerhagen K. Anxiety after stroke: A systematic review and meta-analysis. J Rehabil Med 2018 Sep 28;50(9):769-778 [FREE Full text] [doi: 10.2340/16501977-2384] [Medline: 30184240]
7. Sanner Beauchamp JE, Casameni Montiel T, Cai C, Tallavajhula S, Hinojosa E, Okpala MN, et al. A Retrospective Study to Identify Novel Factors Associated with Post-stroke Anxiety. Journal of Stroke and Cerebrovascular Diseases 2020 Feb;29(2):104582. [doi: 10.1016/j.jstrokecerebrovasdis.2019.104582]

8. Cumming TB, Marshall RS, Lazar RM. Stroke, Cognitive Deficits, and Rehabilitation: Still an Incomplete Picture. International Journal of Stroke 2012 Dec 24;8(1):38-45. [doi: 10.1111/j.1747-4949.2012.00972.x]
9. Mijajlović MD, Pavlović A, Brainin M, Heiss W, Quinn TJ, Ihle-Hansen HB, et al. Post-stroke dementia – a comprehensive review. BMC Med 2017 Jan 18;15(1). [doi: 10.1186/s12916-017-0779-7]
10. Padberg I, Hotter B, Liebenau A, Knispel P, Lehnerer S, Heel S, et al. Unmet Need for Social and Emotional Support and Lack of Recalled Screening Is Associated with Depression in the Long-Term Course After Stroke . RMHP 2020 Apr;Volume 13:285-293. [doi: 10.2147/rmhp.s228265]
11. BergstrÃ¶m A, Koch L, Andersson M, Tham K, Eriksson G. Participation in everyday life and life satisfaction in persons with stroke and their caregivers 3â€“6 months after onset. J Rehabil Med 2015 Jun;47(6):508-515. [doi: 10.2340/16501977-1964]
12. LA Health. 2019. URL: http://publichealth.lacounty.gov/ha/reports/habriefs/2007/Caregiving_Final.pdf [accessed 2019-10-25]
13. Redfern J, McKevitt C, Wolfe CD. Development of Complex Interventions in Stroke Care. Stroke 2006 Sep;37(9):2410-2419. [doi: 10.1161/01.str.0000237097.00342.a9]
14. Almallouhi E, Al Kasab S, Harvey J, Simpson KN, Turner N, Debenham E, et al. Teleneurology Network to Improve Access to Neurologists for Patients in Rural Areas: A Real-World Experience. Telemed J E Health 2020 Jan;26(1):110-113. [doi: 10.1089/tmj.2018.0290] [Medline: 30762494]
15. Pugliese M, Ramsay T, Johnson D, Dowlatshahi D. Mobile tablet-based therapies following stroke: A systematic scoping review of administrative methods and patient experiences. PLoS One 2018;13(1):e0191566 [FREE Full text] [doi: 10.1371/journal.pone.0191566] [Medline: 29360872]
16. Goldberg LR, Bell E, King C, O'Mara C, McInerney F, Robinson A, et al. Relationship between participants' level of education and engagement in their completion of the Understanding Dementia Massive Open Online Course. BMC Med Educ 2015 Mar 26;15(1). [doi: 10.1186/s12909-015-0344-z]
17. Grant RW, Cagliero E, Chueh HC, Meigs JB. Internet use among primary care patients with type 2 diabetes. J Gen Intern Med 2005 May;20(5):470-473. [doi: 10.1111/j.1525-1497.2005.04239.x]
18. Hatcher-Martin JM, Adams JL, Anderson ER, Bove R, Burrus TM, Chehrenama M, et al. Telemedicine in neurology. Neurology 2019 Dec 04;94(1):30-38. [doi: 10.1212/wnl.0000000000008708]
19. Kim DY, Kwon H, Nam K, Lee Y, Kwon H, Chung YS. Remote Management of Poststroke Patients With a Smartphone-Based Management System Integrated in Clinical Care: Prospective, Nonrandomized, Interventional Study. J Med Internet Res 2020 Feb 27;22(2):e15377. [doi: 10.2196/15377]
20. Anderson M, Perrin A. Technology use among seniors. Pew Research Center; Internet & Technology. URL: https://www.pewresearch.org/internet/2017/05/17/technology-use-among-seniors/ [accessed 2020-09-25]
21. Yousufuddin M, Young N. Aging and ischemic stroke. Aging 2019 May 01;11(9):2542-2544. [doi: 10.18632/aging.101931]
22. Russo T, Felzani G, Marini C. Stroke in the very old: a systematic review of studies on incidence, outcome, and resource use. J Aging Res 2011;2011:108785 [FREE Full text] [doi: 10.4061/2011/108785] [Medline: 21876804]
23. Perrin A, Turner E. Smartphones help blacks, Hispanics bridge some but not all digital gaps with whites. Pew Research Center. URL: https://www.pewresearch.org/fact-tank/2019/08/20/smartphones-help-blacks-hispanics-bridge-some-but-not-all-digital-gaps-with-whites/ [accessed 2020-10-06]
24. Cruz-Flores S, Rabinstein A, Biller J, Elkind MSV, Griffith P, Gorelick PB, American Heart Association Stroke Council, Council on Cardiovascular Nursing, Council on Epidemiology Prevention, Council on Quality of CareOutcomes Research. Racial-ethnic disparities in stroke care: the American experience: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2011 Jul;42(7):2091-2116. [doi: 10.1161/STR.0b013e3182213e24] [Medline: 21617147]
25. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
26. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. Journal of Biomedical Informatics 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208]
27. Anderson M, Perrin A, Jiang J. Some Americans don't use the internet. Who are they?. URL: http://www.pewresearch.org/fact-tank/2018/03/05/some-americans-dont-use-the-internet-who-are-they/ [accessed 2018-09-17]
28. Demographics of internet and home broadband usage in the United States. URL: https://www.pewresearch.org/internet/fact-sheet/internet-broadband/ [accessed 2020-04-27]
29. The Diagnosis Difference. URL: https://www.pewresearch.org/science/2013/11/26/the-diagnosis-difference/ [accessed 2020-10-16]
30. Shaffer K, Chow P, Cohn W, Ingersoll KS, Ritterband LM. Informal Caregivers' Use of Internet-Based Health Resources: An Analysis of the Health Information National Trends Survey. JMIR Aging 2018 Dec 18;1(2):e11051 [FREE Full text] [doi: 10.2196/11051] [Medline: 31518244]
31. Fox S, Duggan M, Purcell K. Family caregivers are wired for health. Pew Research Center: Internet, Science & Tech. URL: https://www.pewresearch.org/internet/2013/06/20/family-caregivers-are-wired-for-health/ [accessed 2020-10-16]

32. George MG, Tong X, Bowman BA. Prevalence of Cardiovascular Risk Factors and Strokes in Younger Adults. JAMA Neurol 2017 Jun 01;74(6):695. [doi: 10.1001/jamaneurol.2017.0020]
33. George MG. Risk Factors for Ischemic Stroke in Younger Adults. Stroke 2020 Mar;51(3):729-735. [doi: 10.1161/strokeaha.119.024156]
34. Schumacher S, Kent N. 2020. 8 charts on internet use around the world as countries grapple with COVID-19. URL: https://www.pewresearch.org/fact-tank/2020/04/02/8-charts-on-internet-use-around-the-world-as-countries-grapple-with-covid-19/ [accessed 2020-10-16]
35. Chen M, Chen K, Chiang S, Daimon M, Lee J, Yu EW, et al. A telehealth service model for the treatment of hypertension. J Telemed Telecare 2013 Jul;19(5):238-241. [doi: 10.1177/1357633X13495476] [Medline: 24163232]
36. Al Kasab S, Almallouhi E, Holmstedt CA. Optimizing the Use Of Teleneurology During the COVID-19 Pandemic. Telemedicine and e-Health 2020 Oct 01;26(10):1197-1198. [doi: 10.1089/tmj.2020.0109]
37. Joo H, Dunet DO, Fang J, Wang G. Cost of informal caregiving associated with stroke among the elderly in the United States. Neurology 2014 Oct 10;83(20):1831-1837. [doi: 10.1212/wnl.0000000000000986]
38. Olai L, Borgquist L, Svärdsudd K. Life situations and the care burden for stroke patients and their informal caregivers in a prospective cohort study. Upsala Journal of Medical Sciences 2015 Jun 14;120(4):290-298. [doi: 10.3109/03009734.2015.1049388]
39. Schulz R, Tompkins C. Informal caregivers in the United States: prevalence, caregiver characteristics, and ability to provide care. https://www.ncbi.nlm.nih.gov/books/NBK210048/: National Research Council (US) Committee on the Role of Human Factors in Home Health Care; 2010.
40. Bakas T, McCarthy M, Miller ET. Update on the State of the Evidence for Stroke Family Caregiver and Dyad Interventions. Stroke 2017 May;48(5). [doi: 10.1161/strokeaha.117.016052]
41. Cameron JI, Naglie G, Green TL, Gignac MA, Bayley M, Huijbregts M, et al. A feasibility and pilot randomized controlled trial of the “Timing it Right Stroke Family Support Program”. Clin Rehabil 2014 Dec 31;29(11):1129-1140. [doi: 10.1177/0269215514564897]
42. Achilike S, Beauchamp J, Cron S, Okpala M, Payen SS, Baldridge L, et al. Caregiver Burden and Associated Factors Among Informal Caregivers of Stroke Survivors. J Neurosci Nurs 2020 Dec;52(6):277-283. [doi: 10.1097/JNN.0000000000000552] [Medline: 33156591]
43. Hopwood J, Walker N, McDonagh L, Rait G, Walters K, Iliffe S, et al. Internet-Based Interventions Aimed at Supporting Family Caregivers of People With Dementia: Systematic Review. J Med Internet Res 2018 Jun 12;20(6):e216. [doi: 10.2196/jmir.9548]
44. Morgenstern LB, Sais E, Fuentes M, Ifejika NL, Jiang X, Horn SD, et al. Mexican Americans receive less intensive stroke rehabilitation than non-hispanic whites. Stroke 2017 Jun;48(6):1685-1687. [doi: 10.1161/strokeaha.117.016931]
45. Jacobs BS, Birbeck G, Mullard AJ, Hickenbottom S, Kothari R, Roberts S, et al. Quality of hospital care in African American and white patients with ischemic stroke and TIA. Neurology 2006 Mar 27;66(6):809-814. [doi: 10.1212/01.wnl.0000203335.45804.72]
46. Sacco RL, Boden-Albala B, Gan R, Chen X, Kargman DE, Shea S, et al. Stroke incidence among white, black, and hispanic residents of an urban community: the northern manhattan stroke study. Am J Epidemol 1998 Feb 1;147(3):259-268. [doi: 10.1093/oxfordjournals.aje.a009445]
47. Bridgwood B, Lager K, Mistri A. Interventions for improving modifiable risk factor control in the secondary prevention of stroke. Cochrane Database Syst Rev 2018;5. [doi: 10.1002/14651858.CD009103.pub3]

Abbreviations

CSC: comprehensive stroke center
IRB: institutional review board
REDCap: Research Electronic Data Capture

Edited by G Eysenbach; submitted 19.10.20; peer-reviewed by R Watson, P Tremoulet; comments to author 26.12.20; revised version received 05.01.21; accepted 17.01.21; published 08.03.21.

Please cite as:
Naqvi IA, Montiel TC, Bittar Y, Hunter N, Okpala M, Johnson C, Weiner MG, Savitz S, Sharrief A, Beauchamp JES
Internet Access and Usage Among Stroke Survivors and Their Informal Caregivers: Cross-sectional Study
JMIR Form Res 2021;5(3):e25123
URL: https://formative.jmir.org/2021/3/e25123
doi:10.2196/25123
PMID:33683206

Original Paper

Remote Measurement in Rheumatoid Arthritis: Qualitative Analysis of Patient Perspectives

Katie M White[1*], BSc; Alina Ivan[1*], MSc; Ruth Williams[2], MSc; James B Galloway[3], PhD; Sam Norton[3,4], PhD; Faith Matcham[1], PhD

[1]Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom
[2]Department of Academic Rheumatology, Faculty of Life Sciences and Medicine, King's College London, London, United Kingdom
[3]The Centre for Rheumatic Diseases, Faculty of Life Sciences and Medicine, King's College London, London, United Kingdom
[4]Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom
[*]these authors contributed equally

Corresponding Author:
Katie M White, BSc
Department of Psychological Medicine
Institute of Psychiatry, Psychology and Neuroscience
King's College London
16 De Crespigny Park
London, SE5 8AB
United Kingdom
Phone: 44 7850684847
Email: katie.white@kcl.ac.uk

Abstract

Background: Rheumatoid arthritis (RA) is characterized by recurrent fluctuations in symptoms such as joint pain, swelling, and stiffness. Remote measurement technologies (RMTs) offer the opportunity to track symptoms continuously and in real time; therefore, they may provide a more accurate picture of RA disease activity as a complement to prescheduled general practitioner appointments. Previous research has shown patient interest in remote symptom tracking in RA and has provided evidence for its clinical validity. However, there is a lack of co-design in the current development of systems, and the features of RMTs that best promote optimal engagement remain unclear.

Objective: This study represents the first in a series of work that aims to develop a multiparametric RMT system for symptom tracking in RA. The objective of this study is to determine the important outcomes for disease management in patients with RA and how these can be best captured via remote measurement.

Methods: A total of 9 patients (aged 23-77 years; mean 55.78, SD 17.54) with RA were recruited from King's College Hospital to participate in two semistructured focus groups. Both focus group discussions were conducted by a facilitator and a lived-experience researcher. The sessions were recorded, transcribed, independently coded, and analyzed for themes.

Results: Thematic analysis identified a total of four overarching themes: important symptoms and outcomes in RA, management of RA symptoms, views on the current health care system, and views on the use of RMTs in RA. Mobility and pain were key symptoms to consider for symptom tracking as well as symptom triggers. There is a general consensus that the ability to track fluctuations and transmit such data to clinicians would aid in individual symptom management and the effectiveness of clinical care. Suggestions for visually capturing symptom fluctuations in an app were proposed.

Conclusions: The findings support previous work on the acceptability of RMT with RA disease management and address key outcomes for integration into a remote monitoring system for RA self-management and clinical care. Clear recommendations for RMT design are proposed. Future work will aim to take these recommendations into a user testing phase.

(JMIR Form Res 2021;5(3):e22473) doi:10.2196/22473

KEYWORDS

rheumatoid arthritis; remote measurement technologies; symptom assessment; disease management; smartphone; qualitative research; mobile phone

Introduction

Background

Rheumatoid arthritis (RA) is an autoimmune disease that affects approximately 1% of the adult population in the United Kingdom [1]. Its primary symptoms are recurrent joint pain, swelling, stiffness, and deformities, contributing to fatigue, reduced ability to function [2], increased prevalence of depression [3], reduced quality of life [4], and premature mortality [5]. Symptoms vary from day to day, and disease progression is unpredictable [6]. Currently, clinical status is monitored by regular clinical appointments at fixed intervals, typically every 6 months. As symptom severity fluctuates between visits, appointments may not capture the critical time points of symptom exacerbation.

Remote measurement technologies (RMTs), including smartphone apps and wearable devices, have recently emerged as useful tools for supporting health management [7-9]. Health tracking apps enable patients to actively log changes in symptoms as well as to collect passive data from built-in smartphone sensors or wearable devices. RMT offers opportunities to track symptom severity continuously and in real time, allowing the collection of rich amounts of data in naturalistic settings and overcoming difficulties in exploring symptoms during time-limited appointments [10]. There is growing evidence to support the cost-effectiveness of mobile health interventions. A number of studies have shown positive outcomes, including improved attendance rates at health promotion centers and medication adherence, as well as positive costing outcomes on economic evaluators (eg, a score of 79.6% on the Consolidated Health Economic Evaluation Reporting Standards checklist) [11].

There is a huge appetite for the integration of RMTs into clinical care for RA. A recent systematic review reported the availability of 19 Android or iOS apps for symptom measurement in RA, representing a range of self-reported and passively collected features [12]. Along with this ambition is a growing body of evidence examining engagement with RMT for symptom tracking; the concept of RMT has a good level of face validity, with an estimated 86% of patients with RA reporting a willingness and interest in using apps for symptom monitoring [13].

Despite this ambition, evidence suggests huge variability in engagement with RMT, with adherence levels ranging between 11% and 65%, depending on the requirements for the patients, burden of questionnaire completion, and length of follow-up [14,15]. Barriers for engagement are extensive and specific to the individual; systematic review evidence suggests that perceived clinical value, symptom severity, and convenience are key drivers of uptake [16].

When developed in close collaboration with patients and clinicians, the use of apps in clinical care may positively affect the health outcomes of patients with RA. Co-design involves target end users working with researchers through development, pilot testing, and dissemination [17]. The REMORA (Remote Monitoring of Rheumatoid Arthritis) project, which carefully co-designed a symptom measurement app with patients, has demonstrated excellent levels of patient engagement and has identified temporal changes that might have been previously missed by consultants [18]. In another study, patients with RA who tracked their disease through validated questionnaires and digitally recorded joint counts better adhered to medication, better managed activities of daily living, and reported less worry about the future [19]. Incorporating accelerometer and objective gait balance, alongside symptom reporting, can also accurately predict in-clinic RA activity [20,21].

However, this level of patient input for RMT development and testing is uncommon. Multiple systematic reviews have highlighted the lack of patient involvement in the majority of the apps available, calling for more user experience research to feed into the design and development of app-based symptom measurement systems [12,22]. There is an even greater dearth of patient involvement in the development of platforms to merge subjective symptom reporting with passive data collection [23].

Objectives

Accordingly, the aim of this body of research is to develop an app-based symptom measurement system that fully meets the needs of the patient and clinician users. This study represents the first step in this program, which is to elicit the views of patients to start developing the requirements for a system that captures both subjective and objective symptoms through patient-reported assessments and data collected via passive sensors. The objectives of this project are (1) to identify the symptoms prioritized by patients with RA for inclusion in a multiparametric RMT data collection platform and (2) to identify the key requirements that the platform would need to have for maximized utility, uptake, and long-term engagement with symptom management.

Methods

Design and Ethics

This study used qualitative methodology to understand patient views on the aspects of RA health and clinical care which may be most amenable to measurement via digital technologies. Semistructured focus groups explored the key symptoms and key requirements of an RMT platform. A topic guide (Multimedia Appendix 1) was developed based on recent systematic review evidence [16], which provided a loose framework for discussion. The focus groups were comoderated by a service-user researcher, experienced in qualitative research methods (RW), and the lead investigator (FM). Neither of the moderators were involved in patient care.

The study protocol and topic guide were approved by the Office for Research Ethics Committees Northern Ireland (ORECNI, REC number: 17/NI/0179).

Study Participants and Recruitment

The eligibility criteria for inclusion in this study were as follows: (1) clinically verified RA, (2) aged 18 years or older, (3) able to speak English fluently, and (4) able to give informed consent. Patients were considered ineligible if they were unable to

physically attend a focus group discussion at King's College Hospital or had major cognitive impairment or dementia.

Eligible patients attending outpatient appointments at the King's College Hospital National Health Service (NHS) Foundation Trust rheumatology service were consecutively approached and invited to participate. Patients were initially approached by a clinical trials practitioner, who was not directly involved in the patients' clinical care. Patients were provided with an information sheet and provided verbal information about the nature of the project. With the patient's consent, their contact details were provided to the lead researcher (FM), who then approached participants separately to discuss the study in detail.

All participants provided written informed consent to participate and were informed that their data would be anonymized and that they were free to withdraw at any time with no consequences to their clinical care.

Data Collection

Two semistructured focus groups were conducted in November 2017 and January 2018. Both were conducted by two facilitators: facilitator 1 (RW), a lived-experience service user, and facilitator 2 (FM), a postdoctoral researcher and health psychologist at King's College London. Each focus group lasted approximately 60 min, and participants also completed brief assessments to establish key demographics such as age, gender, comorbidities, and RA disease duration.

Data Analysis

The focus groups were recorded and transcribed verbatim. Each participant was anonymized by assigning a unique number throughout the transcripts. Data-driven thematic analysis was used by two researchers (KW and AI), who independently analyzed both transcripts using NVivo 12 software (QSR International). Codes were discussed between the two researchers and grouped into overall themes. Owing to the data-driven approach used, it was anticipated that topics would emerge that deviated from the original interview guide. These were also included in the thematic analysis to highlight the concepts that the research team may not have considered in advance.

Results

Participant Characteristics

Figure 1 shows the flow of participants from identifying eligible participants to the final sample group of 9. Table 1 shows final participant demographics. Participants were all females (N=9), aged 23 years to 77 years (mean 55.78, SD 17.54) who had lived with a diagnosis of RA for 1-47 years (mean 20.22, SD 14.33; Table 1). Ethnicities included White (n=6), White and Black African (n=1), Pakistani (n=1), and Caribbean (n=1).

Figure 1. Participant flowchart. Percentages were calculated using the eligible participants identified (N=58) as the denominator.

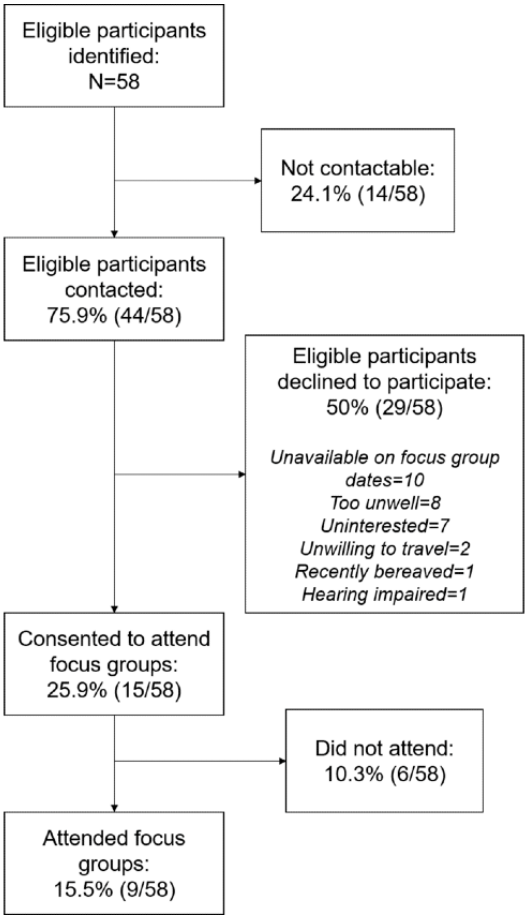


Table 1. Participant demographics.

Participant number	Gender	Age (years), range	Ethnicity
Focus group 1 (n=6)			
P1	Female	75-84	White other
P2	Female	65-74	Pakistani
P3	Female	35-44	White and Black African
P4	Female	65-74	White British
P5	Female	55-64	White British
P6	Female	18-24	White British
Focus group 2 (n=3)			
P7	Female	35-44	Caribbean
P8	Female	65-74	White British
P9	Female	45-54	White British

Themes

A total of four overarching themes emerged from these data. These include the following: (1) important symptoms and outcomes in RA, (2) ways in which patients manage RA symptoms, (3) views on the current RA health care system, and (4) views on remote measurement in RA.

Symptoms Experienced and Important Outcomes

Patients with RA experience a host of symptoms related to both their physical and mental health. Both groups revealed pain and mobility to be the most important outcomes to assess and target during treatment:

> *You say you were crawling. That's pain and stiffness…* [Facilitator 1]
>
> *Oh total pain, total pain.* [P4]
>
> *Well I agree with that. Well I always sort of say what I'm most scared of is losing my functioning, not the pain.* [P8]

Following this, mood was also seen as an important outcome. Patients explained how their physical symptoms impacted their mental health:

> *For me, to be as mobile as possible, so it's not impinging on their lives really, and to control the pain, so I don't get so irritable [chuckles]* [P9]
>
> *Physically, you're also at a low ebb emotionally…because your body hurts. When you're in pain, you try and keep your pecker up, but actually…* [P4]

Further discussions showed that the onset of RA had an impact not only on people's health but also on their lifestyle. Important outcomes here related to maintaining social life and the ability to work and/or be a parent:

> *I gained my teaching assistant qualification July 2015 and then my health deteriorated the September, just after I got offered a job, so I had to turn the job down... I'm sitting there in agony with my feet, with my wrists, not being - having my wrists splintered, not being able to write at the board because my wrists are hurting so much.* [P6]

The inability to maintain a professional work life was also linked with the outcome of mood, with the same patient stating:

> *Yeah, even when I was diagnosed, first diagnosed, I was off work for four months. Initially, it's like, yeah, I can relax, but then it's, I don't know what to do with myself. Then I know, personally, I'll go down. I'm under the psychologist here as well.* [P6]

Although there was a general acknowledgment of the importance of these health and lifestyle outcomes in RA management, there were some individual differences between patients. They explained that these personalized impacts of RA are often the outcomes that are overlooked in the medical profession:

> *You know the outcomes about one's life, is what matters. And that's not always raised.* [P8]

Symptom Management

Symptom management emerged as the second key point of discussion. Ways of dealing with symptoms were first discussed in terms of medication and then in relation to the use of alternative symptom management strategies.

There was a general sense of uncertainty surrounding the effectiveness of medication in both focus groups. Some patients expressed positive experiences of using RA medication to improve outcomes, whereas others described medication as having had no effect on their symptoms. A couple of patients expressed reluctance to use traditional disease-modifying antirheumatic drugs (DMARDs) and steroid medications because of their effectiveness and side effects, respectively. They also felt that there was a lack of support in dealing with the confusion or fear surrounding taking different medications:

> *There's nobody you can talk to, and you can't talk to the consultants about it because it's take it or leave it to some extent. But, there's no experience to draw on, you know, have people got cancer from it, have they got septicaemia? What are the symptoms?* [P9]

Complementary treatments emerged as more acceptable than just taking medication alone. Patients reported improved outcomes from dietary changes, meditation, physiotherapy, and exercise. There was agreement that integration of medication and complementary treatments was important, with suggestions for a more holistic approach to RA management. Several barriers to this approach were highlighted, one being the perceived reluctance of health professionals to prescribe alternatives:

> *There seems to be a slight irritability if you're reluctant to take the drugs from the doctors. And to explain that, you know, I've got 2 children, I don't want to get cancer, I don't want to have septicemia.* [P9]

> *they [consultant] said that there was no evidence to suggest that anything I was taking was going to help.... I was literally just told that my alternatives were not going to do anything for me, but I still went ahead.* [P7]

Other barriers to incorporating complementary treatments into RA management include the allocation of services based on location and restricted access to information. One patient suggested the creation of a website that personalizes exercise class availability based on location. That said, patients also tended to agree that pain or feelings of self-consciousness presented the biggest barrier to using exercise:

> *I think it's that sort of encouragement because it's a hard thing to do, particularly if you're in pain. And also if you feel like your body isn't right it makes you feel less good about your body or the idea of doing exercise.* [P8]

Current Views on the System and Ideas for Improvement

Key aspects in effective disease management include the importance of continuity of care, patient-centered care, and the impact of symptom fluctuations on appointment effectiveness.

Continuity of care was highlighted several times as a major area of importance. Perceived lack of coordination, both between departments and external health care providers, was cited as a barrier to continuity of care. Patients discussed their differing experiences in facilitating the link between their consultant, pharmacist, and prescription delivery service:

> *…I wasn't able to do [DMARDs] for two weeks because I didn't have any medication... I had to try and explain that to them. I said, I've been off my [medication name] since middle of September. You only delivered four [DMARDs]… So missing medication through no fault of my own was not really- so in terms of outcomes, I'd quite like to get my medication, please [laughs].* [P3]

Following this, patient-centered care was also discussed as being of equal importance. Patients reported not seeing the same consultants at each stage of their treatment, meaning that information is often unnecessarily repeated. One individual explained how their experience with an unfamiliar registrar made them feel as though their longstanding *story* and experience of RA was devalued:

> *Yeah, sort of treating the patient as an individual as well isn't is, there's a sort of party line and a protocol, but who is this person sitting in front of you, what is their story? She wouldn't have had to read all the notes just to see I had been diagnosed a long time ago and have been coming to the clinic for many years.* [P8]

More positive experiences tended to include times when medical professionals acknowledged the outcomes of importance that were specific to the patient:

> *He was brilliant, because I love- I like working and I like travelling. I was going to see my daughter in Australia and I took my mum. He came- he notified my GP practice, who came round, pulled out [a pint] in each knee and injected me... he would do it in anticipation of- because I was going to be at my daughter's wedding or whatever it was... I was a person. I was not a number.* [P4]

Patients went on to suggest how patient-centered care could be improved. For some, this involved building trust with medical professionals; for others, being able to provide some background information before appointments would suffice:

> *I know that sometimes you feel if only they had a card to give to all the persons before, if the person could report and that they could even tick, maybe, things so that, yes, that person has that or that or that- that would make it much easier for them.* [P1]

The third key issue questioned the effectiveness of current appointments. Both focus groups consistently noted that symptom fluctuations had a major impact on whether their appointments were effective. In general, it was agreed that symptom fluctuations, especially in relation to pain, are often not captured in appointments:

> *... when they say, how are you? Very often [laughs], if I go, sometimes I have no pain at that time... I forget that I had the pain because I tend to forget I've got pain, but then at times it's, ooh [laughs], excruciating.* [P1]

One patient gave an example of the time that their appointment coincided with a period of heightened pain, which resulted in hospitalization. However, this was described as an exception, rather than the rule:

> *P4: I was very lucky. I had a - I didn't know I had, but I had - I felt I had massive pains everywhere. I happened to be seeing Mr. X or Dr X the next day. He just - he took - I had a blood test done and my CRP was 300 and something and it should be below two... I was found a bed that - there and I was in until the Friday, when it was coming down...*

> *P5: But that was luck more...*

> *P4: Oh totally, totally.*

Discussions surrounding how this could be improved were two-pronged. On the one hand, patients thought it a good idea to be able to conduct appointments or blood tests on days immediately after their pain symptoms. However, there was

consensus that being able to log symptom flares for the purpose of reporting at the time of an appointment would be beneficial as well. It was even suggested that this could be in the form of a digitalized message, for example, a text or an email:

> *I wouldn't mind getting an email saying, how have you been feeling this past week or month or even- that's- I'd actually feel really quite positive about that.* [P3]

Remote Measurement in RA

Patients with RA might benefit both from providing personal information to new clinicians before appointments and keeping a log of symptom fluctuations to review during appointments. As such, each of the focus groups was asked about the potential of an RMT platform, such as a wearable and smartphone app, to track changes in their health.

Patients had varied views on which specific aspects of RA management should be tracked remotely. These generally culminate in two main symptoms: pain and mobility. With pain, patients were keen to highlight that pain can fluctuate by body part, as well as by time of day. With mobility, it was noted that this can vary both within and between days. These were all perceived as important aspects to capture in remote measurements:

> *I suppose I would imagine being able to sort of put in particular parts of the body which tend to have pain: joints, or in my case tendons as well. And then monitor maybe daily how they were doing. Some things that come and go, would be quite interesting.* [P8]

> *Maybe just the, maybe the periods of stiffness, when they are, is it different in the morning.* [P9]

One patient went as far as to describe a way in which their pain fluctuations could be visualized:

> *But if we had a - and we just ticked, for instance, with a timetable and ticked such a time that was crucial. You could even do it in color, I suppose. If it's very bad, you could have it in red or black or green or whatever.* [P1]

It became apparent that it might be of interest to track not only the main symptoms but also the triggers of these symptoms. It was thought that this measurement might act towards preventing symptom flares. There was much deliberation over the symptom triggers that patients were already aware of, and these tended to differ by individual. In general, these included sleep, tiredness, diet, exercise, stress, mood, and other psychological factors. Often, these factors are present comorbidly, and it is difficult to determine the direction of causality:

> *Yeah I think maybe sleep would be another one. But it would be quite interesting to see how those things, sort of mood and also feelings of stress, and tension, how wound up you are about things, might see how that fluctuates, in parallel or not, with joint symptoms. And yes so, it works both ways. How it affects your sleep that's actually quite a big one.* [P8]

Patients also thought it useful to have a list of common measures to track, alongside an *open box*, to note additional symptoms that are relevant to the individual:

> *Facilitator 2: Or if everyone gets to select before they even start what they want to measure, so if I could provide them with a long list of things that are available (voices saying yes) and before you even start you could say "I'm, I identify with this, this, this, and this". And prioritize your top 5 things that you want to be able keep track of....*
>
> *P8: Yes, that would make it more realistic, and even you know, you could review that after 6 months and change it, or look at different things for a period.*

Although there seems to be a general sense that remote measurement in RA might be beneficial in some formats, there were a few differing views on whether encompassing this into an app would be helpful. This seemed to be dependent on the individual's established coping mechanisms:

> *I'm an app freak, I've literally got an app for everything, anything around technology. So I thought, if there's a way I can manage my arthritis through using an app then that would just be brilliant!* [P7]

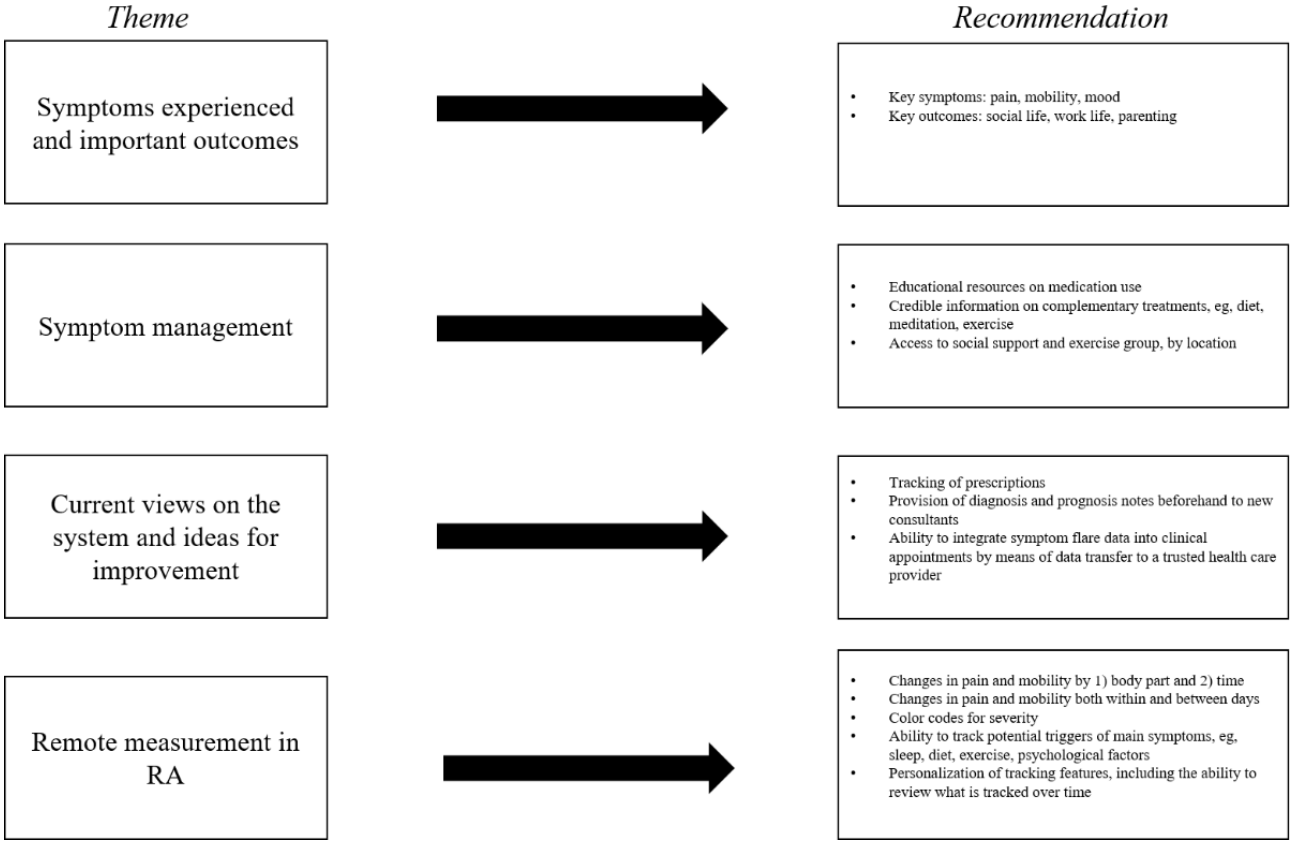
> *The problem with the app, as we're talking about it I think, is that the more you put in to sort of cover everything the more daunting it's going to look like [laughs] the less everyone's going to [interrupted]... it's appealing but it all goes against my main coping strategy which is not thinking about it.* [P8]

Discussion

Principal Findings

This study aimed to provide the first step in developing a multiparametric RMT system for RA symptom tracking by eliciting views on (1) key symptoms prioritized by patients for inclusion in the system and (2) key requirements for maximized utility, uptake, and long-term engagement with RMT symptom management. In addition to these two original objectives, useful information about how RMT systems may be incorporated into existing health care services was elicited. A total of four key themes emerged: (1) key symptoms experienced and important outcomes to consider; (2) clinical and self-management strategies for RA symptoms; (3) ways in which the current health care system can be improved; and (4) facilitating RMT use in patients with RA. Figure 2 breaks down each of these themes into recommendations for consideration when designing a multiparametric RMT system for RA symptom tracking.

Figure 2. Key recommendations for consideration in a remote measurement technology system for rheumatoid arthritis symptom tracking, split by key theme. RA: rheumatoid arthritis.



Comparison With Previous Work

The results of this study align with previous findings supporting the acceptability of RMT apps for use in RA disease management [13-15,24]. Patients approved the development of an app for the purpose of individual symptom tracking. In terms of specific symptoms to include, results reinforce work led by Crouthamel et al [14] and El Miedany et al [19] in the admission that pain and mobility fluctuations are key symptoms to track. In particular, in the study by Crouthamel et al [14], use of a novel joint pain map for the interactive recording of the number and severity of painful joints seems particularly relevant. Our results suggest that patients might find it useful to be able to record these data at multiple periods throughout the day, alongside at weekly intervals. Interestingly, our results show that patients are also interested in tracking symptom triggers, such as sleep, diet, exercise, and mood. Reade et al [15] go some way toward capturing this in their study app, which collects self-reported measures across 10 variables, but crucially this work suggests that patients might prefer to personalize the symptom triggers they are reporting on, as well as to adapt these over time. Much of this work has focused on symptom tracking to improve health outcomes in patients with RA. Although our results suggest these to be of utmost importance, patients also placed value on lifestyle-related outcomes, such as work, parenthood, and socialization. It is less clear at this stage how these aspects of RA management would be incorporated into a symptom-tracking app.

In addition to tracking RA symptoms for self-management, these results suggest that an RMT system should also include a means of transferring data to a health care provider. Logistical barriers exist with respect to sharing data between third-party apps and secure NHS servers, so it is important to consider the value of linking patient symptom trackers with electronic health records (EHRs). Patients proposed the benefits of providing information before appointments, both for the purpose of aiding understanding of their condition and to accurately recall symptom fluctuations. This sentiment resonates well with qualitative work in this area [24]; participants were keen to use RMT tools to communicate with their physician, provided they felt that this would be incorporated into their care. In a similar sense, our findings suggest that provision of information to consultants seems dependent on the trust held for that professional. Previous work demonstrated that integrating remotely captured symptom fluctuations into EHRs presents an effective way for consultants to identify temporal changes and provide tailored disease management [18].

Our findings suggest that an app for patients with RA should include both symptom tracking and data transmission components. At present, the most popular apps available to download for RA management are those that combine symptom tracking and educational content [22]. Educational content, defined as information on disease pathology, diagnosis, or explanation of inputted symptom data, was not explicitly mentioned in our focus groups as an important feature of an app. However, discussions regarding disease management revealed uncertainty surrounding the use of RA medication and complementary treatments. An app that could provide dedicated information and recommendations for a range of treatment

options might be the best fit. Interestingly, our results show that perceived reluctance to consider medication alternatives by health care providers is a barrier to treatment adherence. It might also be useful to consider combining educational content with data transmission features, such that patients could research their personal preferences and report back to consultants for review in upcoming appointments. In addition, patients reported motivation and access as key barriers to accessing complementary treatments. Inclusion of a social support feature in an app would go some way toward combating them. These results are in line with recommendations of Luo et al [22] to implement data transmission and access to social communities, alongside symptom tracking and education, into an app for RA management.

Limitations

Several evaluation points must be considered. First, the sample size was small; therefore, additional themes related to this topic might have been missed, and there is limited generalizability to the wider RA population. Second, all patients across the two focus groups, although presenting a variation in age and disease experience, were female. Both facilitators were also women. Male participants were originally recruited but did not attend on the day. Including a female service user as a facilitator might have been more relatable to patients, yet results may still be biased toward a female experience of RA. It is not yet clear in the literature whether the perception of remote symptom measurement varies by gender. This is a common limitation, and the gender ratio is similar to other studies in the field [24]. Third, sampling for qualitative work inevitably results in a selection effect, whereby certain individuals might be more motivated to participate. Similarly, consecutive sampling through a single center excluded patients who were not currently attending outpatient appointments or could not physically attend a focus group. Given the fluctuations in pain and mobility that are apparent in RA, this might have skewed attendance toward those who were feeling well on the day. Had the groups included those presenting with symptom flares, certain themes may have been exacerbated. Fourth, the topic guide (Multimedia Appendix 1) was intentionally vague. This has elicited some important areas for discussion that were not previously considered by researchers, but it has also resulted in the underdevelopment of intended topics. For example, the concept of passive data monitoring has not been widely discussed in the groups, perhaps owing to the relative novelty of the technologies.

Applications for Future Research

This work provides the foundation for developing a multiparametric RMT system for symptom management in RA. Having explored patient views on key symptoms and concepts for consideration, there are some clear applications for app design (Figure 2). An RMT system should include, at the very least, options to track changes in pain and mobility. Symptom severity may be best tracked visually, via a color chart, with the option of tracking changes by body part and over time. Ideally, users should also be able to add additional personalized symptoms and symptom triggers that they feel are pertinent to their experience of RA. There should also be some informational content available through the app regarding the use of medication or the local availability of complementary therapies. Developers should also consider the concept of data transmission of such information to relevant health care providers.

The next steps for future work in this body of research are to develop an RMT system that can undergo subsequent user testing with a similar group of patients with RA. Service-user workshops offer opportunities to facilitate the co-design of aspects such as user interface and usability [18]. Alongside app development, this should also include the provision of a passive, wearable device that can complement active symptom tracking. Given that our discussions found mobility to be a key outcome of importance, passive monitoring is likely to offer additional, unobtrusive insight into symptoms. These sessions could run over several days and include the option of attendance via video call to allow the inclusion of participants experiencing symptom flares or not attending clinic on the day. The purpose of such user testing should be to assess the usability and feasibility of the system and to understand how to maximize the utility of the data collected while minimizing the burden on patients. In turn, this would provide further insight into the barriers of uptake and long-term engagement with using RMT for symptom tracking.

A key requirement for implementing such technologies in clinical care is to assess the viewpoints of all stakeholders [10]. Parallel work should, of course, look to incorporate the views of rheumatology professionals into these discussions. This is especially relevant given that our findings highlight a clear desire to use an RMT system to send information to clinicians ahead of appointments, in the form of personalized details and symptom fluctuations. It is of upmost importance to assess the feasibility of such data transmission. This work could encompass a combination of both patient and clinician stakeholder views in single focus group sessions, discussing how the data would be incorporated into appointments.

Conclusions

This paper has provided an in-depth exploration of the clinical outcomes valued by a group of patients with RA and, as a result, the key areas of consideration for inclusion in a disease management system. Static time point assessments miss important information for patients with fluctuating disease symptoms. Patients are interested in symptom tracking, and there is a clear and consistent message from patients that remote monitoring via an RMT system has a place in RA self-management and clinical care. This work has helped pave the way for the initial design of such an app, the success of which will be contingent on further co-design with patients. It should capture relevant and personalized outcomes with the possibility of integration with EHRs. Future work in this program aims to combine this app with passive symptom monitoring to create an optimal RMT system for RA symptom tracking. In the current climate of the COVID-19 pandemic, health services are witnessing a rapid shift toward remote management of disorders through telemedicine [25]. This work represents a step toward creating an acceptable and engaging remote system for use as an interface between self-management and clinical care during unprecedented times and beyond.

Acknowledgments

The authors would like to thank the participants who provided the data for this study. They would also like to acknowledge the work of Ella Foncel in approaching the participants for eligibility screening.

This paper represents independent research funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College, London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

This work was supported by the KHP R&D Challenge Fund (Penelope and Eugene Rosenberg Awards) provided by Guys and St Thomas Charity (R160601 Rosenberg Funding).

Authors' Contributions

FM and RW gathered data by conducting the recorded focus groups. KW and AI analyzed the data. All authors contributed to the drafts, read, and approved the final manuscript.

Conflicts of Interest

FM received honoraria or speaking fees from Pfizer (less than $10,000) for contributing to work unrelated to this piece of research.

Multimedia Appendix 1
Semistructured focus group topic guide.
[DOCX File , 21 KB - formative_v5i3e22473_app1.docx]

References

1. Symmons DPM. Epidemiology of rheumatoid arthritis: determinants of onset, persistence and outcome. Best Pract Res Clin Rheumatol 2002 Dec;16(5):707-722. [doi: 10.1053/berh.2002.0257] [Medline: 12473269]
2. Aletaha D, Neogi T, Silman AJ, Funovits JA, Felson DT, Bingham CO, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League against rheumatism collaborative initiative. Arthritis Rheum 2010 Sep 10;62(9):2569-2581 [FREE Full text] [doi: 10.1002/art.27584] [Medline: 20872595]
3. Matcham F, Rayner L, Steer S, Hotopf M. The prevalence of depression in rheumatoid arthritis: a systematic review and meta-analysis. Rheumatology (Oxford) 2013 Dec 03;52(12):2136-2148 [FREE Full text] [doi: 10.1093/rheumatology/ket169] [Medline: 24003249]
4. Matcham F, Scott IC, Rayner L, Hotopf M, Kingsley GH, Norton S, et al. The impact of rheumatoid arthritis on quality-of-life assessed using the SF-36: a systematic review and meta-analysis. Semin Arthritis Rheum 2014 Oct;44(2):123-130 [FREE Full text] [doi: 10.1016/j.semarthrit.2014.05.001] [Medline: 24973898]
5. Pincus T, Sokka T, Wolfe F. Premature mortality in patients with rheumatoid arthritis: evolving concepts. Arthritis & Rheumatism 2001 Jun;44(6):1234-1236. [doi: 10.1002/1529-0131(200106)44:6<1234::aid-art213>3.0.co;2-r]
6. McInnes IB, Schett G. Pathogenetic insights from the treatment of rheumatoid arthritis. The Lancet 2017 Jun;389(10086):2328-2337. [doi: 10.1016/s0140-6736(17)31472-1]
7. Naslund JA, Marsch LA, McHugo GJ, Bartels SJ. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. J Ment Health 2015;24(5):321-332 [FREE Full text] [doi: 10.3109/09638237.2015.1019054] [Medline: 26017625]
8. Malhi GS, Hamilton A, Morris G, Mannie Z, Das P, Outhred T. The promise of digital mood tracking technologies: are we heading on the right track? Evid Based Ment Health 2017 Nov 30;20(4):102-107. [doi: 10.1136/eb-2017-102757] [Medline: 28855245]
9. Sebri SV, Savioni L. An introduction to personalized eHealth. In: P5 eHealth: An Agenda for the Health Technologies of the Future. Switzerland: Springer; 2019:53-70.
10. Matcham F, Hotopf M, Galloway J. Mobile apps, wearables and the future of technology in rheumatic disease care. Rheumatology (Oxford) 2019 Jul 01;58(7):1126-1127. [doi: 10.1093/rheumatology/key391] [Medline: 30535022]
11. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLoS One 2017;12(2) [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
12. Grainger R, Townsley H, White B, Langlotz T, Taylor WJ. Apps for people with rheumatoid arthritis to monitor their disease activity: a review of apps for best practice and quality. JMIR Mhealth Uhealth 2017 Feb 21;5(2):e7 [FREE Full text] [doi: 10.2196/mhealth.6956] [Medline: 28223263]
13. Azevedo R, Bernardes M, Fonseca J, Lima A. Smartphone application for rheumatoid arthritis self-management: cross-sectional study revealed the usefulness, willingness to use and patients' needs. Rheumatol Int 2015 Oct;35(10):1675-1685. [doi: 10.1007/s00296-015-3270-9] [Medline: 25903352]

14. Crouthamel M, Quattrocchi E, Watts S, Wang S, Berry P, Garcia-Gancedo L, et al. Using a researchkit smartphone App to collect rheumatoid arthritis symptoms from real-world participants: feasibility study. JMIR Mhealth Uhealth 2018 Sep 13;6(9):e177 [FREE Full text] [doi: 10.2196/mhealth.9656] [Medline: 30213779]
15. Reade S, Spencer K, Sergeant JC, Sperrin M, Schultz DM, Ainsworth J, et al. Cloudy with a chance of pain: engagement and subsequent attrition of daily data entry in a smartphone pilot study tracking weather, disease severity, and physical activity in patients with rheumatoid arthritis. JMIR Mhealth Uhealth 2017 Mar 24;5(3):e37 [FREE Full text] [doi: 10.2196/mhealth.6496] [Medline: 28341616]
16. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. J Med Internet Res 2018 Jul 12;20(7) [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
17. Boyd H, McKernon S, Mullin B, Old A. Improving healthcare through the use of co-design. N Z Med J 2012 Jun 29;125(1357):76-87. [Medline: 22854362]
18. Austin L, Sharp CA, van der Veer SN, Machin M, Humphreys J, Mellor P, et al. Providing 'the bigger picture': benefits and feasibility of integrating remote monitoring from smartphones into the electronic health record. Rheumatology (Oxford) 2020 Feb 01;59(2):367-378 [FREE Full text] [doi: 10.1093/rheumatology/kez207] [Medline: 31335942]
19. El Miedany Y, El Gaafary M, Youssef S, Bahlas S, Almedany S, Ahmed I, et al. Toward electronic health recording: evaluation of electronic patient-reported outcome measures system for remote monitoring of early rheumatoid arthritis. J Rheumatol 2016 Dec;43(12):2106-2112. [doi: 10.3899/jrheum.151421] [Medline: 27633823]
20. Nishiguchi S, Ito H, Yamada M, Yoshitomi H, Furu M, Ito T, et al. Self-assessment tool of disease activity of rheumatoid arthritis by using a smartphone application. Telemed J E Health 2014 Mar;20(3):235-240. [doi: 10.1089/tmj.2013.0162] [Medline: 24404820]
21. Nishiguchi S, Ito H, Yamada M, Yoshitomi H, Furu M, Ito T, et al. Self-assessment of rheumatoid arthritis disease activity using a smartphone application. Development and 3-month feasibility study. Methods Inf Med 2016;55(1):65-69. [doi: 10.3414/ME14-01-0106] [Medline: 26391694]
22. Luo D, Wang P, Lu F, Elias J, Sparks JA, Lee YC. Mobile apps for individuals with rheumatoid arthritis: a systematic review. J Clin Rheumatol 2019 Apr;25(3):133-141. [doi: 10.1097/RHU.0000000000000800] [Medline: 29933327]
23. Morcillo LF, Morer-Camo P, Rodriguez Ferradas MI, Cazón Martín A. The wearable co-design domino: a user-centered methodology to co-design and co-evaluate wearables. Sensors (Basel) 2020 May 21;20(10):2934 [FREE Full text] [doi: 10.3390/s20102934] [Medline: 32455756]
24. Navarro-Millán I, Zinski A, Shurbaji S, Johnson B, Fraenkel L, Willig J, et al. Perspectives of rheumatoid arthritis patients on electronic communication and patient-reported outcome data collection: a qualitative study. Arthritis Care Res (Hoboken) 2019 Jan 27;71(1):80-87 [FREE Full text] [doi: 10.1002/acr.23580] [Medline: 29669191]
25. Mann DM, Chen J, Chunara R, Testa PA, Nov O. COVID-19 transforms health care through telemedicine: evidence from the field. J Am Med Inform Assoc 2020 Jul 01;27(7):1132-1135 [FREE Full text] [doi: 10.1093/jamia/ocaa072] [Medline: 32324855]

Abbreviations

DMARD: disease-modifying antirheumatic drug
EHR: electronic health record
NHS: National Health Service
NIHR: National Institute for Health Research
RA: rheumatoid arthritis
REMORA: Remote Monitoring of Rheumatoid Arthritis
RMT: remote measurement technology

Edited by G Eysenbach; submitted 13.07.20; peer-reviewed by L Garcia-Gancedo, W Dixon; comments to author 30.09.20; revised version received 18.11.20; accepted 20.12.20; published 09.03.21.

Please cite as:
White KM, Ivan A, Williams R, Galloway JB, Norton S, Matcham F
Remote Measurement in Rheumatoid Arthritis: Qualitative Analysis of Patient Perspectives
JMIR Form Res 2021;5(3):e22473
URL: https://formative.jmir.org/2021/3/e22473
doi:10.2196/22473
PMID:33687333

Original Paper

Cognitive Bias Modification Training to Improve Implicit Vitality in Patients With Breast Cancer: App Design Using a Cocreation Approach

Roos Wolbers[1,2], MSc; Christina Bode[1], PhD; Ester Siemerink[3], MD, PhD; Sabine Siesling[4,5], PhD; Marcel Pieterse[1], PhD

[1]Centre for eHealth and Wellbeing Research, University of Twente, Enschede, Netherlands

[2]Saxion University of Applied Science, Enschede, Netherlands

[3]Department of Internal Medicine, Ziekenhuis Groep Twente (ZGT), Hengelo, Netherlands

[4]Department of Research and Development, Netherlands Comprehensive Cancer Organization (IKNL), Utrecht, Netherlands

[5]Department of Health Technology and Services Research, Technical Medical Centre, University of Twente, Enschede, Netherlands

Corresponding Author:
Christina Bode, PhD
Centre for eHealth and Wellbeing Research
University of Twente
P.O. Box 217
Enschede, 7500 AE
Netherlands
Phone: 31 534896044
Email: c.bode@utwente.nl

Abstract

Background: More than 50% of all patients with breast cancer experience fatigue symptoms during and after their treatment course. Recent evidence has shown that fatigue is partly driven by cognitive biases such as the self-as-fatigued identity bias, which may be corrected with computer-based cognitive bias modification (CBM) techniques.

Objective: The aim of this study was to design a CBM-training app by adopting a cocreation approach.

Methods: Semistructured interviews were conducted with 7 health care professionals, 3 patients with breast cancer, and 2 patient advocates. The aim of the interviews was to collect input for the design of the CBM training, taking the values and preferences of the stakeholders into account, and to determine the timing and implementation of the training in the treatment course.

Results: Overall, the interviews showed that the concept of CBM was accepted among all stakeholders. Important requirements were revealed such as the training needs to be simple and undemanding, yet engaging and persuasive. Based on the results, an eHealth app IVY (Implicit VitalitY) was created. The findings from the interviews suggested that IVY should be offered early in the breast cancer treatment course and should be carefully aligned with clinical treatment.

Conclusions: The findings of this study show that using CBM as a preventive approach to target cancer-related fatigue is an innovative technique, and this approach was embraced by breast cancer stakeholders. Our study suggests that CBM training has several benefits such as being easy to use and potentially increasing perceived self-control in patients.

(JMIR Form Res 2021;5(3):e18325) doi:10.2196/18325

KEYWORDS

breast cancer; cognitive bias modification; eHealth; fatigue; oncology; psychology; vitality

Introduction

Cancer-related fatigue is a highly prevalent and pervasive symptom that has a major impact on the quality of life [1]. More than 50% of all patients with breast cancer experience fatigue complaints during their treatment course, with a substantial proportion retaining serious fatigue symptoms after treatment [2]. Patients report fatigue as more distressing than other cancer-related symptoms such as nausea and pain [1,3]. Furthermore, fatigue is directly related to limitations in occupational participation [4], social participation, and even in performing the daily routine [5].

Cancer-related fatigue is a complex symptom, in which the etiology is multifactorial and not entirely clarified. It is characterized as "an unusual, persistent, subjective sense of tiredness related to cancer or cancer treatment that interferes with usual functioning" [3]. Daily limitations can be related to physical functioning, mood, motivation, cognitions, and social functioning, including interindividual variations. The start, development, and course of cancer-related fatigue are rarely studied. Regarding fully developed serious fatigue symptoms, one of the underlying mechanisms seems to be the psychological distortion in perception, also known as cognitive biases. Previous research studies have suggested that breast-cancer survivors have an attentional bias for cancer-related words compared to healthy individuals [6]. In a similar way, patients with breast cancer can develop a distorted fatigue self-concept ("I am a person with low energy"). This may implicitly consolidate and aggravate fatigue-related behavior and consequently, fatigue as a symptom. These cognitive biases are largely implicit and occur outside awareness. The adverse effect of cognitive biases can be both direct (due to the higher sensitivity to signals of fatigue) and indirect (fatigue-related behavior is unconsciously avoided, resulting in self-imposed limitations and a decrease in physiological capacity) as has been found for pain by Crombez et al [7] and for fatigue by Hughes and colleagues [8]. In order to prevent the full development of serious fatigue symptoms and related behavior, interventions are needed that can be applied early in the process of development of cancer-related fatigue and accompanying cognitive biases. The existing interventions for cancer-related fatigue are reactive and focus on managing existing fatigue symptoms. Our study, in contrast, focuses not only on reducing the existing fatigue but also on pre-emptively enhancing vitality in patients, to make them more resilient during treatment against developing fatigue.

An emerging technique to directly correct or prevent maladaptive biases with simple repetitive association tasks is cognitive bias modification (CBM). CBM programs are aimed at retraining learnt thought patterns (biases). The underlying theories (dual process models) describe human behavior, thoughts, and feelings as the integration of conscious reflexive processes on the one hand and automated and unconscious processes on the other hand [9]. Existing interventions aimed at self-management or rehabilitation primarily target reflexive processes in patients. The method proposed here (CBM) targets automated unconscious processes and can therefore be considered as complementary to current fatigue treatments.

CBM has already shown beneficial effects in reducing anxiety, depression, chronic pain, and addiction [10-12]. The most robust evidence for cognitive bias in the context of fatigue is *attentional bias*. In a systematic review, Hughes et al [8] concluded that patients with chronic fatigue syndrome consistently show an attentional bias toward health-threatening cues compared to healthy controls. This is supported by evidence from research on insomnia [13], a condition closely related to fatigue. While attentional bias represents a perceptual orientation toward fatigue cues, s*elf-identity bias* reflects a more elaborate interpretative cognitive process. As has been shown in pain research, automatically activated associations between pain and "self" play a role in the perception of pain severity [14]. Self-identity is conceptually closely related to a self-schema [15], defined as a person's integrated set of memories, beliefs, and generalizations about his behavior in a certain domain. When patients experience fatigue symptoms over an extended period, either during or following their treatment, they are at risk of developing a schema of the self as a "tired person." Based on evidence from various domains [16,17], it is evident that such self-schemas are strongly rooted within the implicit system and should therefore also be treated at the implicit level. CBM interventions do exactly this. We explore the assumed treatment mechanism (self-identity bias modification to improve fatigue in patients with breast cancer) and work on a theoretical substantiation of the working mechanism for fatigue in this patient group. A theoretical model that informs our conceptual work is the Schema Enmeshment Model of pain proposed by Pincus and Morley [17]. This theory describes the onset and mechanisms of information-processing biases as the intersection of 3 psychological schemas representing the symptom, the illness, and the self. Repeated simultaneous activation of these schemas is hypothesized to result in gradual enmeshment of these schemas, where activation of an element within one schema automatically spreads to the other schemas. Assuming that the Pincus Schema Enmeshment model is equally applicable to the psychosomatic symptom of fatigue, the self-as-fatigued identity bias was chosen as the target in this study.

Based on the hypothesized relationships and working mechanisms, cognitive biases seem to play a role in developing and maintaining fatigue. Therefore, it seems warranted to study whether CBM can be beneficial in reducing and preventing cancer-related fatigue. Furthermore, it might strengthen vitality-related cognitive bias to further buffer the fatigue-stimulating processes. Therefore, the aim of this project was to develop a CBM-training app to increase vitality bias and reduce fatigue bias in patients with breast cancer. The focus of this training will be on retraining fatigue and vitality self-concept associations by letting patients repeatedly pair vitality-related words with "self" and fatigue-related words with "others." The words that we used for CBM were general words describing fatigue and vitality. We explicitly do not use words that are linked to only 1 dimension of fatigue in order to maximize the ability for identification. Offering this training at an early stage of treatment or even pre-emptively by strengthening a vitality-rich self-concept, thereby creating a "buffer" against fatigue symptoms could help making patients more resilient during and after treatment.

Although CBM is a promising technique, acceptance by patients and health care professionals has not yet been systematically studied. Beard et al [18] studied attitudes toward CBM in patients with anxiety disorders. The perceived advantages of CBM were that the tasks were easy and straightforward. However, the barriers to CBM were that the tasks were repetitive and boring. Furthermore, patients lacked understanding about the purpose of the tasks and thought it was strange and not credible. In particular, in the domain of somatic diseases, psychosocial treatments offered to alleviate psychosomatic symptoms such as fatigue are known to evoke feelings of self-blame in patients [19]. Taking into account the possible

resistance in stakeholders with regard to CBM, it seems essential to involve patients and health care professionals throughout the design process [20]. This cocreation approach, based on user-centered design principles, has been shown to lead to more acceptance of systems, better adherence, greater user satisfaction, and better implementation chances of the technology [21].

In this cocreation process, it is first important to obtain insight into the perspective of patients with breast cancer and health care professionals on fatigue and how they perceive the course of fatigue, to determine at which stage of the patient's journey a preventive CBM-vitality training can best be offered to patients. One possibility is that the training should be offered as early in the treatment course as possible in order to obtain the best preventive results. However, in this early phase, patients have just been diagnosed and are preparing for their first treatment, which makes it less likely that patients will be open to use an intervention. Furthermore, the motivation to receive training will probably be higher when the patients experience the first fatigue signals. Therefore, the first aim of this study was to explore the window of opportunity for introducing the CBM intervention in the early stage. The second aim of this study was to examine the attitude of the stakeholders with regard to the concept of CBM to understand the main drivers and barriers for accepting this technique. The third aim was to gather information about their needs and values to formulate concrete design requirements. Finally, it was important to gain more insight into how the CBM training can be successfully implemented in regular clinical care. This leads to the following 4 research questions:

1. What is the preferred moment in the patient's journey to introduce the CBM treatment pre-emptively?

2. What is the attitude among patients and health care professionals toward the concept of CBM aimed at preventing chronic fatigue in patients?

3. What are the requirements that need to be taken into account in the design of the CBM training?

4. How can the training be fitted into the regular clinical treatment regimen in order to facilitate sustainable implementation?

Methods

Participants in This Study

Twelve participants were recruited by purposive sampling [22]. Participants were patients with breast cancer (n=3), patient advocates from BVN (Dutch breast cancer patient association) (n=2), and health care professionals involved in the treatment of patients with breast cancer (n=7). Patient advocates were patients with breast cancer too, but they were invited as representatives and stakeholders of BVN. These patient advocates were following an education program and had insight into the patients' needs and wishes from not only their own situations and experiences but also from many breast cancer types and treatment forms.

Recruitment was stopped when data saturation was achieved and no new information was obtained from the interviews. Patients with breast cancer from a large regional hospital with a breast cancer unit (Ziekenhuis Groep Twente [ZGT]) were included. In 2019, 306 patients from this breast cancer unit have had breast cancer surgery. The mean age of this patient population was 61.8 (SD 13.9) years, which is comparable with the Dutch benchmark from the Dutch Institute for Clinical Auditing with a mean age of 61.6 (SD 12.9) years. Inclusion criteria were that patients were diagnosed with breast cancer and had finished or almost finished receiving cancer treatment. The first patient had finished chemotherapy and was receiving radiotherapy. The second patient was in the last phase of chemotherapy. The third patient had finished treatment several years ago. Health care professionals who participated were a medical oncologist, oncology physiotherapist, nurse practitioner, surgical oncologist, general practitioner, and 2 oncology nurses. Participants were selected by the medical oncologist and approached via telephone by the researcher RW.

Data Collection

Semistructured face-to-face interviews were conducted by RW with patients, patient advocates, and health care professionals. The aim of the interviews was to gain more insight into the course of the fatigue, explore the attitude of the stakeholders toward CBM, collect the requirements for the training based on preferences, and lastly investigate how the training could be implemented into regular care. An overview of the interview topics and questions can be found in Table 1. Interview questions were generally the same for patients, patient advocates, and professionals and were focused on their expertise. In the interviews, the idea of CBM was introduced to the participants with an oral explanation. The literal translation of the original Dutch text is added in Multimedia Appendix 1. To help them understand the concept of CBM, the mobile approach avoidance CBM training "Breindebaas" for alcohol addiction was shown as an example [23]. Interviews lasted for 30-60 minutes and were recorded with a voice recorder. The interviews were conducted in Dutch and back-translated to English. After the interviews, the first version of the platform was developed by software engineers and shown to the same patients for feedback. Two out of 3 patients were willing to test the platform and provide feedback.

Table 1. Overview of the interview topics.

Interview topic	Topics of the questions[a]
Demographics and medical information	Age, occupation, diagnosis, and treatment
Fatigue	Impact and course of fatigue
Attitude	Attitude and experience supporting health by technology and opinion on cognitive bias modification training
Preferences	Preferred platform, content, layout, and explanation of the training
Implementation	Integration of training in regular care and the best moment for offering it to patients

[a]The literal translation of the main questions is given in Multimedia Appendix 1.

Data Analysis

The audio recordings were transcribed verbatim. After transcription, thematic analysis was conducted [24]. This process started with familiarization of the data and generation of the initial codes. Then, the codes were transformed into themes and subthemes. This was done until no new codes emerged, indicating that the saturation point was reached. Interrater reliability was ensured by means of consensus estimates with all authors [25]. Coding was an iterative process in which generating themes was both deductive and inductive [26]. Deductive coding was based on the Technology Acceptance Model [27], Persuasive System Design Model [28], and literature on cancer-related fatigue [29]. The following main themes were identified: (1) course of cancer-related fatigue, (2) attitude toward CBM, (3) preferences with regard to the training, and (4) implementation of the training in the regular treatment course. Qualitative data analysis and research software ATLAS.ti 8.4 (Scientific Software Development GmbH) was used.

Research Ethics

The Medical Ethical Review Committee of Medical Spectrum Twente judged that this study does not fall under the scope of the Medical Research Involving Human Subjects Act (WMO) (K18-47). The Advisory Committee on the Local Feasibility of Scientific Research ZGT agreed by providing a declaration of no objection (ZGT18-44). This study was also approved by the ethical committee of the Faculty of Behavioral, Management and Social Science, University of Twente (file number 18791). Written informed consent was obtained from all participants.

Results

Sample Population

This study comprised a purposive sample that consisted of 3 female patients with breast cancer with mean age of 51.33 (SD 4.71) years (a 48-year-old saleswomen from a fashion store, a 48-year-old secretary, and a 58-year-old youth doctor), 2 female patient advocates with mean age of 63 (SD 5) years, and 7 health care professionals with mean age of 50.43 (SD 7.5) years (females 6/7, 86% and males 1/7, 14%), that is, an oncologist, oncology physiotherapist, nurse practitioner, surgical oncologist, general practitioner, and 2 oncology nurses.

Course of Cancer-Related Fatigue

In the interviews, all patients reported to suffer or have suffered from fatigue, which they experienced as a large burden in daily, social, and occupational life.

> *…You lose your independence as a result. Because if you're so tired, then you won't even be able to go to the supermarket to get your groceries.* [Patient 1]

Professionals confirmed that fatigue is very common in patients with breast cancer.

> *…I think that fatigue is the only complaint that affects everyone.* [Nurse Practitioner]

Patients described fatigue in terms of mental aspects (2 of 3 patients, eg, not being able to concentrate), emotional aspects (2 of 3 patients, eg, frustration), and physical aspects (3 of 3 patients, eg, loss of endurance). They all reported cancer-related fatigue as different from regular fatigue.

> *…Normally you go to bed tired and wake up refreshed, but that is not the case. Everything takes a lot of effort and energy.* [Patient 1]

Furthermore, professionals mentioned that fatigue often happens suddenly.

> *…For example, they have to walk to the counter to get a knife and then they don't know how to get there.* [Surgical Oncologist]

Regarding the course of fatigue, general participants indicated that fatigue is prevalent before, during, and after chemotherapy. Professionals mentioned that fatigue sometimes starts before the chemotherapy because of the diagnosis, poor sleep, and many hospital appointments. Professionals reported that severe fatigue starts during chemotherapy.

> *…When we give chemotherapy, I think that fatigue occurs relatively quickly, after cycles one and two and maybe three.* [Nurse Practitioner]

Patients (Patient 1 and Patient 2) noticed patterns of fatigue within a treatment cycle.

> *…The first two days were pretty good, and the third day it started to collapse. On the fourth and fifth day I felt usually very bad and tired. After that I started to feel a bit better again.* [Patient 1]

However, professionals mentioned that the more treatment cycles patients have had, the longer it takes for patients to recover.

> *…As chemotherapy continues, they become more tired and it takes longer to recover.* [Oncology Nurse 2]

This fact was also acknowledged by all the patients.

> *…I was sick during the first course of chemotherapy, but I always recovered after a cycle. However, the second course was very tough. I did not recover anymore; the fatigue was progressive.* [Patient 3]

Professionals reported that fatigue often persists after treatment.

> *…It will stick around for a very long time after completing chemotherapy.* [Oncologist]

> *…A year after treatment, it is exceptional that fatigue is not presented as the main complaint.* [Surgical Oncologist]

This was confirmed by the patient who had completed treatment several years ago.

> *…I never became my old self again after chemotherapy. [...] I had hoped that the fatigue would disappear over time, but that did not happen. So I had to adjust my life to that.* [Patient 3]

Attitude Toward CBM

Health care professionals mainly reported positive opinions about the concept of CBM.

> *…That sounds very simple. [...] It is also a small time investment.* [General Practitioner]

Health care professionals indicated that they were curious about the results and that it could be fitting for patients with breast cancer.

> *…At a certain point, the fatigue also starts to belong to them. [...] Patients with breast cancer want to do everything they can, a good patient group in that regard.* [Nurse Practitioner]

Health care professionals also seemed hopeful.

> *…Even if it could reduce a little bit of the experienced fatigue, we have gained a lot, because it really is a big problem. Not only for the patient, but also for us as a society.* [Surgical Oncologist]

Furthermore, patient advocates reported that CBM could be interesting for patients who do not want to reach out for psychotherapeutic help.

> *…Some patients are reluctant to go to a psychologist. Maybe they can benefit in this way.* [Patient Advocate 2]

In general, all 3 patients reported to be open and, to some degree, positive toward the concept of CBM.

> *…There is no harm in trying.* [Patient 1]

> *…I think it will give me a boost.* [Patient 2]

They mentioned that they were curious and liked the simplicity of this concept. Two out of 3 patients reported that they would be open to receive CBM training.

> *…Because you want to do everything to endure the cancer treatment as good as possible.* [Patient 1]

One patient was very critical, reporting that she would not use the training without being informed about how it could work. She believed that fatigue was something that happened to her over which she had no control over herself, and therefore, this simple tool evoked resistance.

> *…As if fatigue is a choice.* [Patient 3]

Furthermore, she reported that it was shocking to hear that she could influence the fatigue symptoms herself, especially because she suffered from fatigue for many years after treatment. Nevertheless, she indicated that she was curious and might need some more substantiation before accepting the CBM training.

Preferences With Regard to the Training

In the interviews, a smartphone app was mentioned as the preferred platform for the CBM training by all the participants. The training is based on repeated pairing of vitality-related words (eg, energetic) with “self” and fatigue-related words (eg, exhausted) with “others” in order to strengthen a vitality rich self-concept. On a smartphone app, one can simply swipe the words toward the correct category on the screen instead of pressing a destined key on the keyboard. The interview outcomes revealed important requirements to be taken into account in the development of the app in cocreation. All user requirements are shown in Table 2.

Table 2. User requirements for the training based on the interview outcomes.

User requirements	Examples of interview quotes
The app contains an introduction video that explains the rationale behind the training and explains how to perform the training.	…*I think that a video and spoken explanation will stick better in mind.* [Oncology Nurse 1]
	…*It would be important to know the idea behind the swiping task, otherwise it doesn't make sense.* [Patient 1]
	…*Maybe I need a little more substantiation. The explanation why this would work.* [Patient 3]
One training session lasts no longer than 5 minutes (ie, no more than 100 words).	…*A training should not last longer than 5 minutes.* [Patient Advocate 2]
In the training, the categories are displayed on the top and bottom of the screen.	…*I can also imagine that you bring "energy" toward you and push "fatigue" away from you.* [Oncologist]
After completing a training session in the app, a positive message is displayed.	…*Simply a "thank you" or "well done" after you finished a session.* [Patient 2]
The app includes a progress bar that could be switched on/off.	…*That you can see how many words you have swiped.* [Patient 3]
The name and appearance of the app are positive, feminine, and simple.	….*It must be positive, but do justice to what people go through.* [Patient 3]
The app contains the option to set a daily reminder.	…*A reminder would be helpful. I notice that I quickly forget things.* [Patient 2]
The app is visually attractive and uses warm colors.	…*The more attractive it looks, the more fun it becomes to do the training.* [Patient Advocate 2]
	…*Use of warm colors is always good.* [Patient 2]
The app has a simple layout and is straightforward to use.	…*It should be mainly functional.* [Oncology Nurse 2]
	…*Keep it simple and clear.* [Oncology Physiotherapist]

Based on the requirements in Table 2, the development of the eHealth app "IVY" was started, which stands for Implicit VitalitY. The first version of the app was extensively tested and reviewed in a second iteration by 2 patients. This study continued with this version of the app, because no points of improvement were found.

Implementation of CBM

All health care professionals would recommend the training to their patients because it is a simple and practical tool. Participants were asked how the app can be integrated into regular care so that it becomes a part of daily practice. Multiple health care professionals mentioned that the app could be included in the information session that patients attend before the start of chemotherapy in the neoadjuvant setting.

> …*A nurse will give them information about chemotherapy on the basis of the patient information folder [...] and fatigue is also discussed there.* [Nurse Practitioner]

Professionals expected the app to be helpful.

> …*I think that it is pretty good to give nurses a practical tool to add to the information, that might help patients on their way.* [Nurse Practitioner]

Patient advocates mentioned that the introduction of the app is crucial.

> …*Everything depends on the explanation and introduction. I think it is very important to emphasize that the training can help patients to endure that phase of therapy better, and perhaps also has some positive consequences at a later stage.* [Patient Advocate 2]

Nurses indicated they would be open to introduce and explain the app to patients because it connects to the information that they already provide and requires only a small time investment. Although not necessary, they mentioned that they would like to keep track of the app use once in a while.

> …*We can give patients some support, by just asking about it or motivating them.* [Oncology Nurse 2]

With regard to actually start CBM training using the app, the best option was considered to be just before the second chemotherapy cycle.

> …*The first time they get chemotherapy they just do not know what to expect and are often quite scared. [...] At the start of the second cycle they are usually much calmer.* [Oncology Nurse 2]

Discussion

Summary

The aim of this study was to design a CBM training to prevent chronic fatigue and improve implicit vitality in patients with breast cancer. This was a cocreation process in which patients, patient advocates, and health care professionals were involved.

Time Point for Offering the CBM Training

The first aim was to discover the window of opportunity in the patient's journey to offer the CBM training pre-emptively. Interviews showed that fatigue seems to increase and become more persistent as chemotherapy continues, which confirms the findings of previous research studies [30,31]. In line with a

study by Spichiger et al [32], fatigue seems to be most prevalent from the third and fourth treatment cycle onwards. Our findings clearly show that the earliest suitable stage to introduce the CBM app is shortly after the start of chemotherapy but not before the first cycle. At this stage, patients have gone through the initial shock of receiving their diagnosis. They may be less apprehensive about chemotherapy after having experienced the first cycle. Additionally, for some patients, the first fatigue symptoms have already set in, motivating them to engage in a treatment. However, these findings about the preferred moment of introduction of IVY does not inform us about the point in the treatment for achieving the best effects regarding fatigue symptoms. In this regard, an effect study is needed.

Stakeholder Perspective on CBM Training for Vitality

The second aim was to examine the attitude of the stakeholders with regard to the concept of CBM in preventing fatigue, which has not been studied before. Overall, participants showed a positive, open, and curious attitude toward CBM, with remarkably little skepticism or resistance. The health care professionals perceived the training as a simple and useful tool for providing advice to their patients. The health care professionals found CBM training to be helpful in providing patients with something they can do by themselves in this stage. Patients often feel frustrated and helpless about undergoing treatment and feel unable to do something about reducing the side effects such as fatigue in this phase. Offering a simple tool to tackle fatigue themselves seems to enable patients regain a sense of self-control, regardless of the actual effects of the training. An interesting finding was that the CBM concept evoked resistance in 1 patient. This patient attributed the fatigue to external factors only and thus was hesitant about the idea that fatigue can be influenced by oneself. It was shocking for her that such a simple preventive task could have influenced fatigue—a symptom that she suffered from for several years that had forced her to adjust her life completely. This might be interpreted as the first indication for the beneficial effects of preventive use of CBM training for vitality in order to buffer the development of an unbeneficial cognitive bias for fatigue. However, it should be taken into account that the training can possibly lead to self-blaming in patients [19]. On the one hand, self-blaming might lead to negative affect. On the other hand, blaming oneself might enhance perceptions of control [33]. Nevertheless, it seems relevant to inform patients about the automaticity and lack of awareness that characterize cognitive biases to avoid harmful interpretations by patients. Such educational content might easily be included in the app.

The third aim of this study was to explore the preferences with regard to the CBM training. A smartphone app was the preferred platform for the training. Interestingly, participants spontaneously suggested a vertical positioning of the CBM categories, with the “Self-Vitality” combination at the bottom and the “Other-Fatigue” stimuli on top, thereby allowing a more intuitive swipe movement toward or away from their body accordingly. Although the self-as-fatigued identity CBM was originally based on the implicit association test paradigm, this adds an approach-avoidance mechanism that originates from a different CBM paradigm [34]. A classical implicit association test is considered to represent evaluative associations with the target—in this case, the concepts of fatigue and vitality, stored in memory after repeated previous experiences [35]—and is conceptually most closely related to an attitude as a cognitive concept [36]. The approach-avoidance task is thought to capture cognitive embodiment, which is conceptually closer to motoric impulses [37,38]. By combining both the evaluative conditioning from the implicit association test paradigm and the motoric tendencies from the approach-avoidance task paradigm, multiple mechanisms may be targeted, which could potentially enhance the effectiveness of the intervention.

Additional requirements that emerged from the interviews can be explained in the light of the persuasive system design model [28], which defines several persuasive design principles. From the category of primary task support, there seemed to be a need for self-monitoring (progress bar) and reduction (simple layout). From the category dialogue support, there seemed to be a need for liking (visually attractive), praise (positive messages), and a reminder. From the category credibility support, there seemed to be the need for expertise by offering a good explanation for the rationale behind the training. No findings related to the category of social support emerged in line with the individual nature of the training task. Overall, the simplicity of the training appeared to be the major strength. This implies that adding persuasive elements such as gaming features may not improve usability. In line with Hassenzahl’s hedonic/pragmatic model of user experience, this suggests that patients expect health technology to be pragmatic rather than hedonic in nature [39]. Based on all the requirements, the eHealth app IVY was created. The link to the app is shown in Multimedia Appendix 1. With regard to the implementation, it is recommended that IVY be aligned carefully with clinical treatment. Furthermore, it is recommended that health care professionals are involved to foster successful implementation. Nurses seemed willing to provide oral explanation about the app and mentioned that it could be included in patient education materials.

Strengths of This Study

A strength of this study is the user-centered design approach. The design of IVY was a cocreation process, involving a broad range of stakeholders, including patient advocates at an early stage in order to optimize successful implementation at a later stage. This study is innovative in multiple ways. First, this study is the first to develop a CBM intervention for fatigue for patients with cancer undergoing treatments. Second, this is the first study that focuses on developing a preventive CBM intervention instead of working on the existing reactive CBM interventions. Currently available CBM evidence is limited to attempts to correct already established maladaptive response tendencies. However, assuming that the concepts of fatigue and vitality are to some degree ends of a single dimension, it seems plausible that training patients toward the positive end would create a buffer against the development of adverse conditioning later on during the disease. Even when both concepts would be rather separate dimensions, strengthening vitality bias may still help attenuating adverse effects of fatigue bias, as has been shown convincingly by the two-continua approach in the positive psychology field: building positive mental health has shown to reduce the incidence of psychopathology or to protect against its impact on the quality of life [40,41]. Lastly, by creating a

mobile CBM intervention, this study is also innovative within the CBM research field.

Limitations of This Study

A limitation of this study is the small sample size, especially of the patient group (n=3). However, the patient advocates are patients with breast cancer too. They were invited as representatives and stakeholders of BVN. The mean age of the patients seemed representative, considering the regional hospital population and in relation to the Dutch Institute for Clinical Auditing registry. Therefore, our study includes a broader patient perspective than that based on the input of 3 patients participating on personal title. This sample is expected to be sufficient for the goal of this study, which is an intervention design. Furthermore, health care professionals were able to position themselves in the place of patients and as such, may also represent the patient perspective in certain way. Besides, the reach of saturation in the answers is a valid indicator of sufficient sample size [42]. Representativeness with regard to age is a limitation of this study since patients with breast cancer younger than 40 years or older than 70 years did not take part in the cocreation of IVY.

The second limitation of this study is that the interviews did not provide as much information about design and content as was expected. Participants were overall quickly satisfied and not very critical. A possible explanation for this is that participants had no prior knowledge about persuasive elements and simply had no idea about all the options. Future research studies should give more examples about persuasive elements to inspire patients to come up with their own ideas and preferences.

The last limitation of this study is that the verbal stimuli that are used in the app were not pretested among the participants of this study. However, the stimuli were collected, pretested, and validated in previous studies [43-45] in healthy individuals. The verbal stimuli were retrieved from both items in validated questionnaires for measuring the constructs of fatigue and vitality and from interviews. In this study, no negative responses to the stimuli were observed, confirming the face validity of these materials. However, in further cocreation research, it would be recommended to evaluate these stimuli among this patient population as well.

Clinical Implications

Patients suffering from fatigue are often reluctant to receive time-consuming and energy-consuming interventions such as cognitive behavioral therapy, especially during their treatment. Moreover, these therapies are not widely available owing to limited clinical capacities or insurance coverage. This simple CBM swipe training requires low levels of mental energy and literacy. Furthermore, patients can complete the training anywhere and anytime on their own smartphone, which makes it very time-saving and cost-effective. IVY is expected to be a low-threshold intervention that is suitable for many people. When the implicit vitality training shows to be effective, it could also be applied to other cancer types, populations, or chronic diseases accompanied with serious fatigue symptoms. The unique characteristic of IVY is that it can be used as a preventive intervention in the early stage of cancer treatment unlike most existing fatigue programs that target already manifested fatigue. The aim of this psychological conditioning task is that the self-concept becomes more focused on vitality, which aims to create a buffer to protect against the fatigue symptoms that will occur during treatment and recovery. This pro-vitality self-image will then contribute positively to recovery, promoting behavior, and to quality of life. Taking into account that fatigue is one of the most prevalent symptoms in patients with cancer, even small progress could be clinically relevant.

Acknowledgments

This work was supported by grants from ZGT Science voucher and resources of the University of Twente. The data that support the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available due to privacy or ethical restrictions.

Authors' Contributions

All authors contributed to the research idea and study design. Data acquisition was done by RW and ES. Data analysis and interpretation were performed by RW, CB, and MP. Each author contributed important intellectual content during manuscript drafting and revisions and accept accountability for the overall work by ensuring that questions on the accuracy or integrity of any portion of the work are appropriately investigated and resolved. All authors have read and approved the manuscript and agreed to act as guarantors of this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary material.
[DOCX File , 66 KB - formative_v5i3e18325_app1.docx]

References

1. Ryan JL, Carroll JK, Ryan EP, Mustian KM, Fiscella K, Morrow GR. Mechanisms of Cancer - Related Fatigue. The Oncologist 2007 May;12(S1):22-34. [doi: 10.1634/theoncologist.12-s1-22]
2. Fabi A, Falcicchio C, Giannarelli D, Maggi G, Cognetti F, Pugliese P. The course of cancer related fatigue up to ten years in early breast cancer patients: What impact in clinical practice? Breast 2017 Aug;34:44-52. [doi: 10.1016/j.breast.2017.04.012] [Medline: 28500901]
3. Morrow G. Cancer-related fatigue: causes, consequences, and management. Oncologist 2007;12 Suppl 1:1-3 [FREE Full text] [doi: 10.1634/theoncologist.12-S1-1] [Medline: 17573450]
4. Paalman C, van Leeuwen FE, Aaronson N, de Boer AGEM, van de Poll-Franse L, Oldenburg H, et al. Employment and social benefits up to 10 years after breast cancer diagnosis: a population-based study. Br J Cancer 2016 Jan 12;114(1):81-87 [FREE Full text] [doi: 10.1038/bjc.2015.431] [Medline: 26757424]
5. Curt G, Breitbart W, Cella D, Groopman J, Horning S, Itri L, et al. Impact of cancer-related fatigue on the lives of patients: new findings from the Fatigue Coalition. Oncologist 2000;5(5):353-360 [FREE Full text] [doi: 10.1634/theoncologist.5-5-353] [Medline: 11040270]
6. Custers JAE, Becker ES, Gielissen MFM, Van Laarhoven HWM, Rinck M, Prins JB. Selective attention and fear of cancer recurrence in breast cancer survivors. Ann Behav Med 2015 Feb;49(1):66-73. [doi: 10.1007/s12160-014-9632-9] [Medline: 24838871]
7. Crombez G, Van Ryckeghem DML, Eccleston C, Van Damme S. Attentional bias to pain-related information: a meta-analysis. Pain 2013 Apr;154(4):497-510. [doi: 10.1016/j.pain.2012.11.013] [Medline: 23333054]
8. Hughes A, Hirsch C, Chalder T, Moss-Morris R. Attentional and interpretive bias towards illness-related information in chronic fatigue syndrome: A systematic review. Br J Health Psychol 2016 Nov;21(4):741-763. [doi: 10.1111/bjhp.12207] [Medline: 27329758]
9. Strack F, Deutsch R. Reflective and impulsive determinants of social behavior. Pers Soc Psychol Rev 2004;8(3):220-247. [doi: 10.1207/s15327957pspr0803_1] [Medline: 15454347]
10. Mobini S, Reynolds S, Mackintosh B. Clinical Implications of Cognitive Bias Modification for Interpretative Biases in Social Anxiety: An Integrative Literature Review. Cogn Ther Res 2012 Mar 14;37(1):173-182. [doi: 10.1007/s10608-012-9445-8]
11. Grafton B, MacLeod C, Rudaizky D, Holmes EA, Salemink E, Fox E, et al. Confusing procedures with process when appraising the impact of cognitive bias modification on emotional vulnerability. Br J Psychiatry 2017 Nov;211(5):266-271. [doi: 10.1192/bjp.bp.115.176123] [Medline: 29092835]
12. Kakoschke N, Kemps E, Tiggemann M. Approach bias modification training and consumption: A review of the literature. Addict Behav 2017 Jan;64:21-28. [doi: 10.1016/j.addbeh.2016.08.007] [Medline: 27538198]
13. Harris K, Spiegelhalder K, Espie C, MacMahon K, Woods H, Kyle S. Sleep-related attentional bias in insomnia: A state-of-the-science review. Clin Psychol Rev 2015 Dec;42:16-27. [doi: 10.1016/j.cpr.2015.08.001] [Medline: 26284598]
14. Grumm M, Erbe K, von Collani G, Nestler S. Automatic processing of pain: the change of implicit pain associations after psychotherapy. Behav Res Ther 2008 Jun;46(6):701-714. [doi: 10.1016/j.brat.2008.02.009] [Medline: 18423571]
15. Markus H. Self-schemata and processing information about the self. Journal of Personality and Social Psychology 1977 Feb;35(2):63-78. [doi: 10.1037/0022-3514.35.2.63]
16. Ouimet A, Gawronski B, Dozois D. Cognitive vulnerability to anxiety: A review and an integrative model. Clin Psychol Rev 2009 Aug;29(6):459-470. [doi: 10.1016/j.cpr.2009.05.004] [Medline: 19552990]
17. Pincus T, Morley S. Cognitive-processing bias in chronic pain: a review and integration. Psychol Bull 2001 Sep;127(5):599-617. [doi: 10.1037/0033-2909.127.5.599] [Medline: 11548969]
18. Beard C, Weisberg R, Primack J. Socially Anxious Primary Care Patients' Attitudes Toward Cognitive Bias Modification (CBM): A Qualitative Study. Behav Cogn Psychother 2011 Nov 30;40(5):618-633. [doi: 10.1017/s1352465811000671]
19. Callebaut L, Molyneux P, Alexander T. The Relationship Between Self-Blame for the Onset of a Chronic Physical Health Condition and Emotional Distress: A Systematic Literature Review. Clin Psychol Psychother 2017 Jul;24(4):965-986. [doi: 10.1002/cpp.2061] [Medline: 27925335]
20. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008 Mar;4(1):5-18. [doi: 10.1080/15710880701875068]
21. van Gemert-Pijnen L, Kip H, Kelders SM, Sanderman R. Introducing eHealth. In: eHealth Research, Theory and Development. Milton Park: Routledge Taylor & Francis Group; 2018:23-46.
22. Tongco MDC. Purposive Sampling as a Tool for Informant Selection. Ethnobot. Res. App 2007 Dec 31;5:147-158. [doi: 10.17348/era.5.0.147-158]
23. Tactus H. Breindebaas (Version 1.1). URL: https://play.google.com/store/apps/details?id=com.tfe.tactus.cbm&hl=nl [accessed 2021-02-01]
24. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
25. Stemler S. A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. Practical Assessment, Research & Evaluation 2004;9(4):1-19.

26. Rivas C. Coding and analysing qualitative data. Researching society and culture, 3 2012:367-392. [doi: 10.1007/978-1-137-11022-0_12]
27. Venkatesh V, Davis F. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
28. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: Key issues, process model and system features. Communications of the Association for Information Systems. 2018. URL: https://www.researchgate.net/publication/257411281_Persuasive_Systems_Design_Key_Issues_Process_Model_and_System_Features [accessed 2020-01-02]
29. Berger AM, Mooney K, Alvarez-Perez A, Breitbart WS, Carpenter KM, Cella D, et al. Cancer-Related Fatigue, Version 2.2015. J Natl Compr Canc Netw 2015 Aug 13;13(8):1012-1039. [doi: 10.6004/jnccn.2015.0122]
30. Servaes P, Verhagen C, Bleijenberg G. Fatigue in cancer patients during and after treatment. European Journal of Cancer 2002 Jan;38(1):27-43. [doi: 10.1016/s0959-8049(01)00332-x]
31. Donovan KA, Jacobsen PB, Andrykowski MA, Winters EM, Balducci L, Malik U, et al. Course of fatigue in women receiving chemotherapy and/or radiotherapy for early stage breast cancer. Journal of Pain and Symptom Management 2004 Oct;28(4):373-380. [doi: 10.1016/j.jpainsymman.2004.01.012]
32. Spichiger E, Müller-Fröhlich C, Denhaerynck K, Stoll H, Hantikainen V, Dodd M. Prevalence of symptoms, with a focus on fatigue, and changes of symptoms over three months in outpatients receiving cancer chemotherapy. Swiss Med Wkly 2011;141:w13303. [doi: 10.4414/smw.2011.13303] [Medline: 22065282]
33. Bulman R, Wortman C. Attributions of blame and coping in the "real world": Severe accident victims react to their lot. Journal of Personality and Social Psychology 1977;35(5):351-363. [doi: 10.1037/0022-3514.35.5.351]
34. Wiers R, Rinck M, Kordts R, Houben K, Strack F. Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. Addiction 2010 Feb;105(2):279-287. [doi: 10.1111/j.1360-0443.2009.02775.x] [Medline: 20078486]
35. Wiers R, Gladwin T, Hofmann W, Salemink, Ridderinkhof K. Cognitive Bias Modification and Cognitive Control Training in Addiction and Related Psychopathology. Clinical Psychological Science 2013 Jan 09;1(2):192-212. [doi: 10.1177/2167702612466547]
36. Greenwald A, McGhee, Schwartz JLK. Measuring individual differences in implicit cognition: The implicit association test. Journal of Personality and Social Psychology 1998;74(6):1464-1480. [doi: 10.1037/0022-3514.74.6.1464]
37. Neumann R, Förster, Strack F. Motor compatibility: the bidirectional link between behavior and evaluation. In: The psychology of evaluation: Affective processes in cognition and emotion. Mahwah, NJ: Lawrence Erlbaum; 2003:371-391.
38. Macy JT, Chassin L, Presson CC, Sherman JW. Changing implicit attitudes toward smoking: results from a web-based approach-avoidance practice intervention. J Behav Med 2015 Feb;38(1):143-152 [FREE Full text] [doi: 10.1007/s10865-014-9585-2] [Medline: 25059750]
39. Hassenzahl M. The Thing and I: Understanding the Relationship Between User and Product. In: Blythe MA, Overbeeke K, Monk AF, Wright PC, editors. Funology. Human-Computer Interaction Series, vol 3. Dordrecht: Springer; 2003.
40. Fredrickson B. The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. American Psychological Association 2001;56(3):218-226. [doi: 10.1037/0003-066x.56.3.218]
41. Westerhof GJ, Keyes CLM. Mental Illness and Mental Health: The Two Continua Model Across the Lifespan. J Adult Dev 2010 Jun;17(2):110-119 [FREE Full text] [doi: 10.1007/s10804-009-9082-y] [Medline: 20502508]
42. Brod M, Tesler LE, Christensen TL. Qualitative research and content validity: developing best practices based on science and experience. Qual Life Res 2009 Nov;18(9):1263. [doi: 10.1007/s11136-009-9540-9] [Medline: 19784865]
43. Klaus O. First indications of the effectiveness of an IAT in modifying attentional bias regarding tiredness or vitality. Master's thesis, University of Twente. 2016. URL: https://essay.utwente.nl/71296/ [accessed 2020-01-02]
44. Pieterse ME, Bode C. Cognitive biases involved in fatigue: measuring and modifying a self-identity fatigue bias to improve self-reported symptoms. 2018 Presented at: 7th ARPH Conference; 24-25 Jan 2018; Tilburg.
45. Pieterse ME, Bode C. Cognitive bias modification to relieve fatigue symptoms: preliminary results from a self-identity fatigue IAT. In: Psychology, Health & Technology. 2018 Presented at: the 32nd EHPS Conference; August 21-25, 2018; Galway, Ireland URL: https://research.utwente.nl/en/activities/cognitive-bias-modification-to-relieve-fatigue-symptoms-prelimina

Abbreviations

BVN: Dutch breast cancer patient association
CBM: cognitive bias modification
IVY: Implicit VitalitY
ZGT: Ziekenhuis Groep Twente

Edited by G Eysenbach; submitted 19.02.20; peer-reviewed by D Ryckeghem, N Kakoschke; comments to author 29.03.20; revised version received 30.06.20; accepted 11.11.20; published 10.03.21.

Please cite as:
Wolbers R, Bode C, Siemerink E, Siesling S, Pieterse M
Cognitive Bias Modification Training to Improve Implicit Vitality in Patients With Breast Cancer: App Design Using a Cocreation Approach
JMIR Form Res 2021;5(3):e18325
URL: https://formative.jmir.org/2021/3/e18325
doi:10.2196/18325
PMID:33688833

Original Paper

Attitudes of Australian Patients Undergoing Treatment for Upper Gastrointestinal Cancers to Different Models of Nutrition Care Delivery: Qualitative Investigation

Kate Furness[1,2,3], BSc, BHSc; Catherine Elizabeth Huggins[4], BSc, PhD; Helen Truby[5], BSc (Hons), M, PhD; Daniel Croagh[6,7], PhD, MBBS; Terry Peter Haines[2,3], BPhys, PhD

[1]Monash Health, Nutrition and Dietetics, Monash Medical Centre, Melbourne, Australia

[2]School of Primary and Allied Health Care, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia

[3]Department of Physiotherapy, School of Primary and Allied Health Care, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia

[4]Department of Nutrition, Dietetics and Food, School of Clinical Sciences, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia

[5]School of Human Movement and Nutrition Sciences, University of Queensland, Brisbane, Australia

[6]Upper Gastrointestinal and Hepatobiliary Surgery, Monash Health, Monash Medical Centre, Melbourne, Australia

[7]Department of Surgery, School of Clinical Sciences, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia

Corresponding Author:
Kate Furness, BSc, BHSc
Monash Health
Nutrition and Dietetics, Monash Medical Centre
246 Clayton Road, Clayton
Melbourne, 3168
Australia
Phone: 61 95944180
Email: kate.furness@monashhealth.org

Abstract

Background: Adults diagnosed with cancers of the stomach, esophagus, and pancreas are at high risk of malnutrition. In many hospital-based health care settings, there is a lack of systems in place to provide the early and intensive nutritional support that is required by these high-risk cancer patients. Our research team conducted a 3-arm parallel randomized controlled trial to test the provision of an early and intensive nutrition intervention to patients with upper gastrointestinal cancers using a synchronous telephone-based delivery approach versus an asynchronous mobile app–based approach delivered using an iPad compared with a control group to address this issue.

Objective: This study aims to explore the overall acceptability of an early and intensive eHealth nutrition intervention delivered either via a synchronous telephone-based approach or an asynchronous mobile app–based approach.

Methods: Patients who were newly diagnosed with upper gastrointestinal cancer and who consented to participate in a nutrition intervention were recruited. In-depth, semistructured qualitative interviews were conducted by telephone and transcribed verbatim. Data were analyzed using deductive thematic analysis using the Theoretical Framework of Acceptability in NVivo Pro 12 Plus.

Results: A total of 20 participants were interviewed, 10 from each intervention group (synchronous or asynchronous delivery). Four major themes emerged from the qualitative synthesis: participants' self-efficacy, low levels of burden, and intervention comprehension were required for intervention effectiveness and positive affect; participants sought a sense of support and security through relationship building and rapport with their dietitian; knowledge acquisition and learning-enabled empowerment through self-management; and convenience, flexibility, and bridging the gap to hard-to-reach individuals.

Conclusions: Features of eHealth models of nutrition care delivered via telephone and mobile app can be acceptable to those undergoing treatment for upper gastrointestinal cancer. Convenience, knowledge acquisition, improved self-management, and support were key benefits for the participants. Future interventions should focus on home-based interventions delivered with simple, easy-to-use technology. Providing participants with a choice of intervention delivery mode (synchronous or asynchronous) and allowing them to make individual choices that align to their individual values and capabilities may support improved outcomes.

Trial Registration: Australian and New Zealand Clinical Trial Registry (ACTRN) 12617000152325; https://tinyurl.com/p3kxd37b.

(JMIR Form Res 2021;5(3):e23979) doi:10.2196/23979

KEYWORDS

qualitative; upper gastrointestinal; cancer; nutrition; mobile phone

Introduction

Background

Adults diagnosed with cancers of the stomach, esophagus, and pancreas are at a high risk of malnutrition [1-9]. The nature of upper gastrointestinal (UGI) cancers directly affects digestive capacity and function due to tumor location and obstruction and inability to tolerate adequate volumes of oral intake confer risks of malnutrition, which are exacerbated by oncological, radiological, and surgical treatments [8,10]. Malnutrition increases the risk of complications. These include immune impairment leading to increased risk of infection, treatment toxicity resulting in dose reductions or cessation of chemotherapy, loss of muscle mass and decreased strength, increased complications associated with surgery, more hospital admissions, longer length of stay with decreased quality of life, and increased mortality [6,9,11,12].

In many hospital-based health care settings, there is a lack of systems in place to provide the early and intensive nutritional support that is required by these high-risk cancer patients. Our research team conducted a 3-arm parallel randomized controlled trial (RCT) to test the provision of an early and intensive nutrition intervention to patients with UGI cancers using a synchronous telephone-based delivery approach versus an asynchronous mobile app–based approach delivered using an iPad compared with a control group to address this issue [13].

A proposed definition of acceptability has been theorized by Sekhon [14] describing it as "a multi-faceted construct that reflects the extent to which people delivering or receiving a health care intervention consider it to be appropriate, based on anticipated or experienced cognitive and emotional responses to the intervention." When we achieve acceptability with trials, we are more likely to see intervention adherence and engagement, and the proposed health outcomes of the intervention are more likely to be realized [15,16].

The majority of patients diagnosed with UGI cancers are aged over 60 years [17,18]. Given the increasing use of information and communication technologies (ICTs) such as mobile telephones, home computers, and the internet to provide health care, we need to be sure that older adults are not only willing but also able to use these technologies to receive their health care. Although these technologies are viewed as a bridge to the cost inefficiencies of in-person health care, allowing patients to access health care in their own homes, reducing inequities, and meeting their unmet health needs, they are potentially very costly interventions if patients do not accept and engage with the intervention being delivered [19]. A qualitative study delivered face-to face interviews with 123 people aged over 60 years examining the use of technology to evaluate their readiness to adopt health-related ICT found that increasing age negatively impacted ICT use, they needed ready access to support such as a spouse or family member, they were reluctant to adopt new technologies unless they were convinced of significant benefits, and their use of ICT was very basic [20]. Overall health was a moderating factor on ICT use, so the people who are most likely to benefit were the people least likely to use it [20].

Telehealth feasibility and acceptance have been examined by many reviews of telephone-delivered health-related interventions. Telehealth can be delivered via home phone or mobile telephone, so it has the potential to capture a larger audience of the population. Cancer-specific telehealth, looking at symptom management during head and neck cancer treatment, improves patients' ability to self-manage their disease, including side effects from aggressive treatment [21]. It also provides patients with support and security [21]. A systematic review of 22 studies examining cancer survivors' experience with telehealth health interventions found that it improved accessibility, that patients could raise concerns otherwise difficult to raise, and that it provided a safety net to receive support where required [22]. Other patients felt that it was time consuming, an additional burden, impersonal, and not sufficiently tailored to their individual needs and that lower engagement was found where the patient and health professional had not met in person before commencement of telehealth [22]. Another systematic review of telehealth interventions for cancer patients' quality of life revealed a statistically significant improvement in quality of life and improved access to care, and telephone-based interventions were found to be superior to internet-based interventions [23]. COVID-19 has resulted in the rapid adoption of telehealth services to enable remote delivery of health care to comply with physical distancing laws and to keep vulnerable patients safe [24,25]. Thus, the impetus to understand the best way to deliver these services to enhance acceptability, enhance engagement, and improve patients' health care outcomes has become even more urgent.

Objectives

This study aims to explore the overall acceptability of an early and intensive nutrition intervention delivered via a synchronous telephone-based approach or an asynchronous mobile app–based approach.

Methods

This qualitative study was set within a 3-arm RCT that examined two nutrition service delivery models, using a synchronous telephone-based approach or an asynchronous mobile app–based approach, to deliver a standardized early and intensive nutrition intervention to patients with UGI cancers (esophageal, gastric,

and pancreatic) close to diagnosis, to determine improvements in quality of life, the protocol is available [13].

RCT Interventions

The research dietitian contacted participants with their randomization information and set a date for the completion of their initial nutrition assessment. For some participants, this was carried out on the same day as the randomization call. The 18-week intervention commenced as soon as it was practicable after the participants' diagnosis, consent, recruitment, and baseline data collection.

Participants received either intensive weekly or fortnightly collaborative and individually tailored nutrition interventions during their interaction with the dietitian. The synchronous telephone intervention was delivered using the participants' home or mobile telephone. The mobile app intervention was delivered using a preexisting mobile app, MyPace. All participants were offered an iPad (Apple Air 2) device with internet connectivity prepaid from the study team that could be used to run the app; however, participants were also permitted to install the app on their own device, if they preferred to do so. MyPace allowed a messaging function for the participant and dietitian to communicate asynchronously and daily reminders to assist with self-monitoring of weight and completion of scheduled small steps (goals). Behavior change techniques used throughout the delivery of interventions in both groups were taken from the Behavior Change Technique Taxonomy v1 [26]. At the end of the 18-week intervention period, patients were able to access usual care.

Design

Semistructured individual interviews were used to gather data from the perspective of participants in each of the intervention groups of the overarching RCT. The Standards for Reporting Qualitative Research were used in this study [27].

Eligibility

Participants were eligible to participate if they were randomized into either of the intervention groups of the overarching RCT.

Setting

The participants were drawn from 5 health services, including 3 tertiary public hospitals and 2 private hospitals in Victoria, Australia.

Participants

Convenience sampling was used to select a group of participants with similar numbers involved in both the intervention arms of the overarching research trial.

Procedures

One-on-one recorded semistructured interviews were conducted by a researcher (CH) via telephone. Interviews were conducted between September 2017 and June 2019. The interviews were designed to be completed within 30 minutes to ensure that they were not overly taxing to the participants.

Method of Approach

At the conclusion of active interventions, a convenience sample of participants was asked if they would consent to participate in a postintervention interview with another researcher (CH). Of those who accepted, their data was entered into an intervention completion data file with their name, contact details, and intervention group details. The researcher (CH) contacted the participants by phone. Participants who did not respond to the first phone call were then contacted with another follow-up call, email, or text message. Participants who did not respond to the follow-up were not contacted again (17/37, 46%).

Measurements

An interview guide (Table 1) was developed by the researcher (CH) after feedback from the overarching research trial investigator team. Interviews were conducted between September 2017 and June 2019. Immediately at the conclusion of each interview, the interviewer (CH) made reflective field notes.

Table 1. Semistructured postintervention interview questions.

Questions	Logic
As someone who has cancer, what is it like for you managing your nutrition?	Living with cancer
Tell me about the experience you had as a participant in this study. • Did it meet your nutritional needs?	Relevance to the patient
What was it like for you being contacted by the dietitian frequently? • Tell me what was challenging. • Would you change anything (throw something out, add something in?) • What did you like? • (iPad group) What was it like learning a new App?	Self-management practice
Tell me what it was like communicating with a health profession using the phone (or iPad). • What helped or hindered communication between you and the dietitian? • What would have made this experience better for you? • Describe any challenges you had communicating with the dietitian. • What do you need to facilitate communication? • What could we have improved the way we delivered the nutrition to you?	Communication
If you could design this service, what would be the key features of the service? • Tell me about the scheduling of the consultations. • How important is flexibility? • What could we have done to support you better?	Unmet care needs
What motivated you to take part in this study?	Motivation
Is there anything else you'd like to tell me about that relates to your experience throughout the intervention?	Overall experience
What role did your family play in your nutrition care during the study period?	Social influences
Did you contact the dietitian as often as you wanted to? • What motivated or stopped you from using the app?	Contact
Did you have any problems using the app or contacting the dietitian? • If you had problems, what were they? • How did you solve the problem?	Technical problems
Did any of your family members help you with the app or dietetic consultations? • How important was it to you to be able to involve someone else in this service? • Would they like to share with me their experience of the intervention?	Family/career engagement

Data Preparation

The recordings of the interviews were transcribed verbatim and deidentified for analysis. Interview recordings and transcriptions were stored in a secure cloud-based repository.

Researcher Positioning

The researcher KF is a senior clinical dietitian in general, UGI, and hepatobiliary surgery at one of the tertiary hospitals in this study. The delivery of nutrition intervention in the concurrent intervention study was also conducted by this researcher (KF).

Reflexivity

As an insider in this research study, KF was deeply embedded with emotional investment in the provision of health care to the intervention population and a member of the health professional team [28]. KF continually evaluated her subjective and automatic responses and how they were intertwined with how she ultimately interpreted and actively constructed knowledge throughout the research process [29].

Rigor

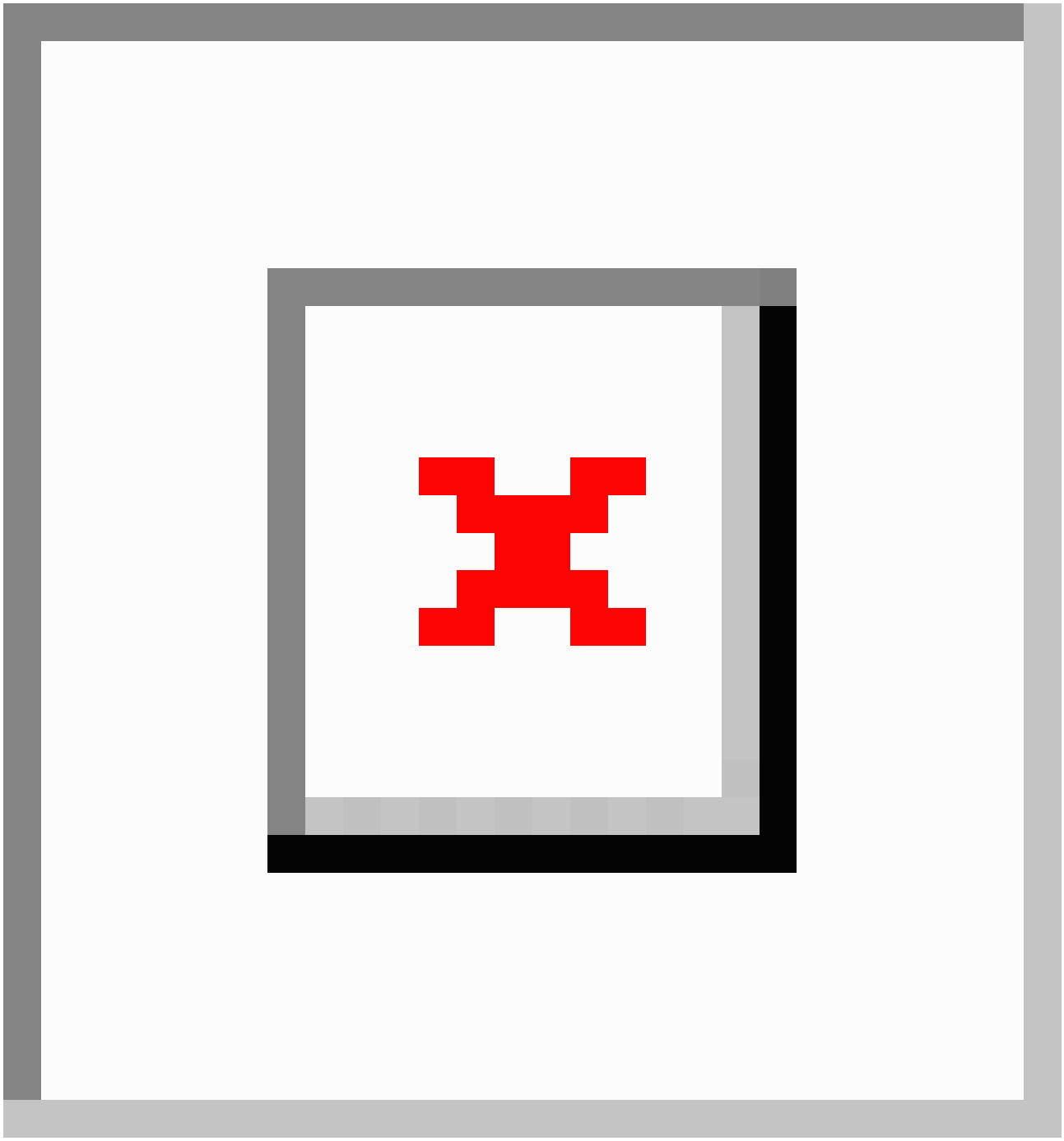
Standardized interview questions were developed before the commencement of interviews and refined throughout to promote rigor in data collection. A wide variety of participants across

both intervention groups with different cancer types, ages, and genders increased the dependability of the data collection. Verbatim transcription of interview audio recordings and data analysis verification by a second author (CH) supported accuracy and reliability.

Analysis

The data analysis framework was based on the 7 constructs of the Theoretical Framework of Acceptability: affective attitude, burden, intervention coherence, ethicality, opportunity costs, perceived effectiveness, and self-efficacy (Figure 1) [15]. Hand coding of all interviews was used as the first pass of the deductive coding of framework constructs. Researchers (KF) and (CH) met to discuss the coding accuracy of the two interviews until a consensus was reached. NVivo Pro 12 Plus was used to sort and analyze the data. This process allowed coding to be reexamined carefully.

Figure 1. The Theoretical Framework of Acceptability. Data were collected retrospectively and explored using each of the 7 acceptability constructs.



Ethics Approval

Conduct of the study at all sites was approved and reviewed by the Monash Health Human Research Ethics Committee (HREC/16/MonH/290). Site-specific authorization has been granted for all sites (Monash Health, Jessie McPherson Private Hospital, Cabrini Health, Eastern Health, and Peninsula Health). Participants provided written informed consent to participate in the study.

Consent to Publish

Participants provided written informed consent to participate in the study, which is the approved consent process by the main ethics committee (Monash Health Human Research Ethics) and the other sites (Cabrini Health, Eastern Health, and Peninsula Health).

Data Availability

Where requested by publishers, deidentified participant data may be stored in a public repository. These data will be restricted and deidentified so that combinations of data cannot be used to potentially identify participants.

Results

Participant Sample

A total of 37 participants consented to be interviewed, and a total of 20 interviews were completed, providing a total response rate of 54%. There were 15 participants who were unable to be contacted via telephone despite consenting to be contacted for postintervention interviews; these participants were nonresponders. Two participants subsequently withdrew their initial consent to participate. All interviews were conducted via telephone.

Information power was used as a tool to assist with sample size determination, as described by Malterud [30], which indicates that lower numbers of participants are required when more relevant information is held by the sample.

Demographics of the interview sample are shown in Table 2.

Table 2. Characteristics of the interview sample.

Participant characteristics	Total participants (N=20), n (%)	Telephone (n=10)	Mobile app (n=10)
Gender			
Female	7 (35)	3	3
Male	12 (60)	7	6
Other	1 (5)	0	1
Age (years)			
50-59	5 (25)	3	2
60-69	8 (40)	3	5
70-79	7 (35)	4	3
Cancer type			
Gastric	6 (30)	3	3
Esophageal	5 (25)	3	2
Pancreatic	9 (45)	4	5
Education			
Year 1-12	8 (40)	4	4
>Year 12	12 (60)	6	6
Baseline Patient-Generated Subjective Global Assessment[a] score			
<9	9 (45)	5	4
≥9	11(55)	5	6
Mortality	9 (45)	5	4

[a]Patient-Generated Subjective Global Assessment: scores ≥9 indicate a critical need for nutrition intervention and symptom management [31].

There were a total of 20 participants, 10 participants from each group (7 female, 12 male, and 1 other). Five participants were aged between 50 and 59 years, 8 were aged between 60 and 69 years, and 7 were aged between 70 and 79 years.

Self-Efficacy, Low Levels of Burden, and Intervention Comprehension Required for Intervention Effectiveness and Positive Affect

Participants perceived the intervention as effective when particular conditions related to the construct of acceptability were met. They positively viewed the intervention when there were low levels of participant burden, when they understood how the intervention worked, when they felt that it aligned with their value system, and when they were confidently able to perform the tasks and behaviors required of them. If these conditions were not met, the intervention was perceived as having little benefit. This was described more commonly by the participants in the mobile app group. People felt that their age, skill level, and familiarity with technology affected their participation in the intervention. The introduction of new

technology often did not correspond to their daily habits and normal behaviors. Therefore, it required them to learn a new skill and incorporate this into their lives, at the same time as receiving treatment for cancer. This was often too challenging for participants and was a significant barrier to engaging in eHealth interventions. Even when participants had high levels of technology skill, if the intervention delivered via the mobile app did not align with their value system of communication, they were less inclined to use it:

> *I've learnt so much from interacting with [Dietitian], and her suggestions to how I can improve my diet and what have you. I have learnt a lot from her, and I appreciate that, and I think that's the best part of it, that a patient has a chance to interact with a professional like that, and be able to help themselves.* [Participant 10, 61-year-old female]

> *I'm at that age where although I understand computers and I can do just about anything with them, to me there's nothing better than sitting down and having a face-to-face or a voice-to-voice conversation. So maybe it's a bit of an age thing with me.* [Participant 1, 70-year-old male]

> *–Na, not bothering with an app, because I don't even have anything on my phone. I've got app things there and stuff I've never even looked at them. All these things pop up and I'm like, "You know what? I don't want to be on me phone."* [Participant 4, 71-year-old female]

Many participants used a family member or spouse or partner to increase their engagement with the intervention on their behalf. This occurred in both the synchronous and asynchronous groups. The burden of the intervention was perceived to be higher when English was not their first language or when symptoms from the disease and/or treatment were high. This was often coupled with lack of self-efficacy and skills to use technology:

> *He didn't do any of it. He is good on the computer, but he's limited, he doesn't sort of muck around as much as I do.* [Participant 8, 73-year-old male]

For the mobile app group, it was found that ease of use of hardware and operating systems was important for the acceptance of the mode of delivery. Participants who chose to use their own devices instead of the provided iPad had issues with service delivery, as the app was not supported by Android devices. Participants who experienced service disruption due to technical issues with the technology itself, including issues with loading the app on their home computers, tablets, and smartphones; logging in; and internet connectivity, became disengaged and frustrated.

There was a small subgroup of participants who did not realize that they were not participating in the intervention as intended, reflecting that they poorly understood the intervention. They perceived the intervention as low burden and believed they had high self-efficacy to complete what was required of them, resulting in affective attitude and perceived effectiveness. The intervention was not actually delivered to them as was prescribed because they did not fully understand the requirements for complete participation:

> *I think I was weighing myself twice a week.* [Participant 9, 61-year-old female]

The analysis using the Theoretical Framework of Acceptability constructs has been illustrated in Table 3.

Table 3. Analysis using the Theoretical Framework of Acceptability constructs.

Theoretical framework of acceptability: construct	Definition	Participants, n (%)	Code frequency: for	Defining quote	Code frequency: against	Defining quote
Affective attitude	How an individual feels about the intervention	18 (90)	122	"The dietitian was quite competent in answering everything overall over that 18 weeks and I think she just kind of set the groundwork for me and now I kind of know what I need to do to best look after myself." (Participant 14, 78-year-old male)	12	"It was not important to have any external information about my food and my diet because I knew it myself." (Participant 13, 71-year-old female)
Burden	The perceived amount of effort required to participate in the intervention	19 (95)	32	"I think really important because you don't have to go anywhere. You don't have to look respectable, when it's going to an appointment for the chemo or for the doctors or oncologists et cetera. Just to have that informal phone call, is great." (Participant 1, 70-year-old male)	31	"Well, I didn't really go on the app at all. I'm not really a technical person." (Participant 18, 72-year-old male)
Ethicality	The extent to which the intervention has a good fit with the individual's value system	6 (30)	9	"Well, all I can say is that that was very positive. You know, I was more than happy with speaking to [Kate] via the telephone. I didn't feel as though I needed to sit across the desk from anybody and speak face to face. No, I'm very happy with it, you know, that mode of communication." (Participant 1, 70-year-old male)	2	"I'm at that age where although I understand computers and I can do just about anything with them, to me there's nothing better than sitting down and having a face-to-face or a voice-to-voice conversation. So maybe it's a bit of an age thing with me." (Participant 7, 68-year-old female)
Intervention coherence	The extent to which the participant understands the intervention and how it works	13 (65)	17	"Primarily it's support to patients with proper diet and maintaining weight and all that sort of thing." (Participant 3, 75-year-old female)	6	"I didn't access anything, I had you or somebody ring me up and we went through the questions verbally on the phone." (Participant 10, 61-year-old female)
Opportunity costs	The extent to which benefits, profits, or costs must be given up to engage in the intervention	6 (30)	3	"And you know, I made myself available. Otherwise, I was always able to leave her a message. I had her number and was able to leave her a message and you know, we could reconvene at another time." (Participant 2, 63-year-old male)	3	"I prefer the phone – because I don't have the time." (Participant 4, 71-year-old female)
Perceived effectiveness	The extent to which the intervention is perceived as likely to achieve its purpose	17 (85)	120	"Well, like I said, (the dietitian) would send me an email and suggest various things that I can do. And I emailed her and told her what I was doing, and she encouraged me to do those things, because they were important. So, it was a case of between the two of us, we bounced back on each other. And I got what I needed out of it." (Participant 12, 70-year-old female)	19	"The small steps thing that I was supposed to be filling out, I couldn't record it every day." (Participant 16, 61-year-old male)
Self-efficacy	The participant's confidence that they can perform the behavior(s) required to participate in the intervention	11 (55)	14	"No, it wasn't a problem, because I'm very tech-savvy, I don't have a problem with any computer or iPad." (Participant 8, 73-year-old male)	13	"I'm 80 darling and I would stuff it up." (Participant 5, 79-year-old male)

A Sense of Support and Security Through Relationship Building and Rapport

Participants who engaged well with the intervention frequently reported feeling a sense of support and security and felt that someone was there if they needed them. This was built through the frequent nature of interactions with participants, which facilitated the formation of a trusting relationship. Participants eagerly awaited their appointments with the dietitian, and the cessation of the intervention at 18 weeks shocked and saddened them, as they felt a keen sense of loss. These psychological and emotional effects on participants at the conclusion of trials have been reported in a number of other qualitative articles examining postintervention transition to usual care, with participants reporting a sense of loss, disappointment, anxiety, and isolation [32,33]. This highlights the need to incorporate how to manage the psychological impacts of abrupt cessation of intensive interventions on trial participants in the future [33]. Despite this sense of loss, they also felt well prepared to manage their nutrition needs moving forward even if more treatment was planned (eg, additional rounds of chemotherapy). Participants often shared intimacies of their symptoms that they may have found embarrassing to discuss without the close rapport and nature of non-face-to-face interventions where some level of anonymity was maintained. The conduct of the dietitian delivering the intervention in both groups was remarked as important for building rapport. Participants valued what they referred to as high levels of empathy through nonjudgmental, kind, and encouraging communication. Furthermore, participants reported that the strong rapport enabled them to offload their concerns and reduce their loneliness without burdening those close to them:

> *When the dietitian used to ring, she would listen, she was never judgmental. If I was feeling rotten and I swore on the phone, she didn't say anything. It's a very lonely journey.* [Participant 1, 70-year-old male]

> *I think it's benefited us, and him in particular, because he felt like he had someone watching over him, in respect of that, and it didn't cost us anything, it didn't take up too much time, and it was a great help. I felt like it was a guardian angel over him. So, to have that support, it was wonderful.* [Participant 8, 73-year-old male]

> *I really missed it. I really missed that conversation via the iPad because I could express myself quite well, I thought.* [Participant 2, 63-year-old male]

Knowledge Acquisition and Learning Enabled Empowerment Through Self-Management

Participants reported developing knowledge on how to make appropriate decisions to self-manage their nutrition impact symptoms and often complex dietary modifications alongside their treatment. They felt this was through the provision of supportive, targeted educational contact that was easy for them to understand with guidance and repeated messaging. Many participants had a high symptom burden, including fatigue, anorexia, and diarrhea, many of which were considered very distressing, reducing their quality of life. Some participants reported that they initially thought that these were inevitable, that they needed to cope with them, and that they could not be remedied. Participants described that the detailed explanations of disease process and treatment effects on the body allayed their anxieties, whereas timely identification and intervention as a team allowed them to take more control of their symptom management:

> *And the body not doing what I was telling it. It was doing what it wanted to. And tough luck with what I was thinking, virtually. So that required understanding, information, and support.* [Participant 4, 71-year-old female]

> *The dietitian was quite competent in answering everything overall over that 18 weeks and I think she just kind of set the groundwork for me and now I kind of know what I need to do to best look after myself.* [Participant 14, 78-year-old male]

Convenience, Flexibility, and Bridging the Gap of Hard-to-Reach Individuals

Given the physical and psychological burden of cancer treatment, participants overwhelmingly preferred the convenience, flexibility, and accessibility of a home-based intervention. Many people lived far away from tertiary hospitals where many of their face-to-face health provider interactions took place, which was a significant burden for these participants. They expressed the favorable conditions of not having to get dressed or have a formal appointment, and these synchronous or asynchronous interactions likely increased their level of comfort and relaxation with the intervention, particularly related to information sharing. Most participants did not allude to the financial ramifications of seeing health care providers, but one participant saw the benefit of being offered a free service. Participants in both groups enjoyed the flexibility of being able to schedule appointments that fit with their lives or to send messages via the app late at night when they were experiencing concerns or anxieties, knowing that it would be attended too quickly. Some participants thought that the 18-week intervention should have been delivered as 18 individual sessions when they required it, rather than over consecutive weeks. As they were in and out of hospital for treatment, they may have missed weeks, and they felt that 18 weeks did not often cover their treatment duration:

> *If I'm worried about something at 10 o'clock at night I can send an email off and the dietitian will get back to me the next day. That takes the anxiety out of that situation I think as much as anything, so it was good.* [Participant 16, 61-year-old male]

> *You know, some days it's a struggle to go to the garage and jump in the car let alone, you know, go to – go sit in the car for half an hour or 45 min to get somewhere, it's just hard to do, so to have that flexibility of being able to say, you know, well can we do it on such-and-such a date, yeah, makes a hell of a difference.* [Participant 15, 52-year-old male]

One participant described the requirement to escalate concerns as they arose via the telephone despite being randomized into the mobile app group. The store-and-forward nature of the

mobile app group meant that concerns may not have been actioned in a period that was deemed fast enough to respond to urgent issues:

> *Yes. I suppose I felt like my message was being heard more urgently if I spoke to someone, rather than just did it via the app. I don't know that I had that confidence that it was going to be picked up straight away. I suppose that's just me, I shouldn't have done that, but anyway.* [Participant 1, 70-year-old male]

Discussion

Principal Findings

Delivery of an early and intensive nutrition intervention via telephone or mobile app were both largely acceptable delivery modes for people with UGI cancers. However, it was apparent that specific requirements need to be met for this to be the case. Participants required a perception that their self-efficacy was high, that the intervention was of low burden, and that they understood what was required of them to feel that the intervention was acceptable. This led to participants having positive affect related to the study, and they viewed the intervention as effective in meeting their needs. Some participants did not find the mobile app delivery mode acceptable. They cited their age, perceived and actual skill level, technological savviness, and delivery mode requiring a significant shift from normal daily activities as major barriers to engagement. Adapting the intervention to improve perceived fit with the individual may include modification of technology and mobile app with extensive consumer engagement, allowing people to choose their own delivery mode to suit their communication preference and allowing participants to choose the timing and quantity of dietitian contact. Ascertaining whether participant autonomy in choosing the timing and delivery of the contact for both delivery approaches would impact the effectiveness of the nutrition intervention requires further examination.

To date, few studies have investigated the application of eHealth during the active treatment phase of cancer, with even fewer studies using it in the UGI cancer cohort [34]. The penetration of smartphones in Australia is at an all-time high, with 91% of the population having a smartphone, including 77% of people aged over 55 years [35]. This aligns well with the UGI cancer patient population, where most diagnoses are made in those aged over 60 years [17,36,37]. Having a smartphone does not necessarily correlate with the high-level use of the technology embedded within them. Older adults are the least likely age group to engage with technology, as they are often late adopters of new innovations [38,39]. This may be a transient concern as technology becomes increasingly embedded in our everyday lives.

Low burden through ease of use was highlighted as an important consideration for our participants, supported by evidence that individuals may cease use of technology if the benefits are not perceived early on in its use [38]. Participants reported issues with loading the app, logging on, and internet connectivity, all of which would have increased the burden of use for many participants unfamiliar with the use of technology. This, in turn, may have increased motivational barriers and decreased self-efficacy. Older people are less likely to ask for help, as they do not want to burden others despite technical assistance being offered by the research team [19]. Many of our study participants required the assistance of family members, such as their spouses, to undertake the communication element of the intervention on their behalf when they deemed engaging with the eHealth component too challenging for them individually; interestingly, this was predominantly female partners/carers. A study of older Dutch adults exploring teleconferencing intention and capability to use found that family support positively impacted frequency of use and self-efficacy through mastery [40]. Engaging family members more actively may be an important mechanism to enhance participants' use of technology, and it needs to be considered in intervention delivery in the future.

Participants' nutrition, symptom, and pharmacological management information needs were high, and they requested easy-to-understand, tailored information provided by a confident and supportive dietitian to enhance their self-management. This aligns with patients with cancer generally wanting to receive information that helps them [41]. These aspects of nutrition intervention delivery are essential to meet the acceptability constructs of affective attitude and perceived effectiveness. A survey of 185 patients with cancer exploring satisfaction with the information provided found that only 50% were happy with the level and amount of information they were given throughout their treatment journey [42]. Many of our study participants required ongoing, repeated messaging throughout the course of their 18-week treatment journey to actualize self-management. They found that both methods of delivery allowed them to communicate better with their dietitian, which enabled shared decision making.

Limitations of this study include the possibility of selection bias due to participant mortality before the opportunity to participate in this evaluative aspect of the overall study. Similarly, the absence of data from those who were uncontactable and those who withdrew their consent may also have introduced bias. Participants who were amenable to participating in the interviews may have been more likely to view the intervention positively. There was a risk that people provided obsequious responses in this study, as the same investigative team that developed and delivered the intervention conducted this evaluation. To minimize this potential risk, an investigator who was not involved in the delivery of the intervention with participants undertook the data collection interviews. Similarly, there was a risk that investigators involved in the analysis of our data may have sought to overplay positive feedback and underplay negative feedback, as they were involved in the delivery of the intervention. We verified the data analysis and coding by a second author who was not involved in the delivery of the intervention to mitigate this risk.

Future research needs to focus on evaluating the relationships between acceptability, engagement, and use to ensure that eHealth nutrition interventions are effective in their intended health outcomes, both long and short term. Similarly, evaluating the cost-effectiveness of these interventions will be critical to inform future uptake of novel eHealth service delivery models.

Central service delivery models versus those delivered through individual health services need to be explored in detail.

Conclusions

This study has shown that early and intensive eHealth nutrition models delivered via telephone and mobile app are acceptable to patients undergoing treatment for UGI cancers when the core requirements for individual fit are realized, including self-efficacy, low levels of burden, and comprehension of the intervention to bring about effectiveness and positive affect. Convenience, knowledge acquisition, improved self-management, and support were key benefits for participants to engage with eHealth. Future interventions of this nature should focus on home-based interventions as an adjunct to usual health care. Simple, easy-to-use technology with technical support, which allows individual choice with respect to the mode of delivery of the intervention so that it aligns with the individual's intrinsic value system, will enhance acceptability.

Acknowledgments

The authors wish to thank the study participants who generously shared their experiences. The concurrent RCT study is funded by the Victorian Cancer Agency (project identification number HSR15007). The Victorian Cancer Agency is based at 50 Lonsdale St, Melbourne, CH is supported by a National Health and Medical Research Fellowship (Translating Research into Practice). There was no role of the funding agency in study design. Involvement in collection, management, analysis, or interpretation of data will be completed independent of the funding body. Submission of reports for publication are independent of the funding organization.

Authors' Contributions

CH conceived the interview guide and interviewed all participants. KF completed all thematic analyses of the interview transcripts. CH completed the theme analysis of two manuscripts to initiate theme development. KF drafted the manuscript. All authors provided critical review of the manuscript. All authors read and approved the final manuscript. All authors conducted the RCT on which this RCT is based.

Conflicts of Interest

None declared.

References

1. Andreyev H, Norman A, Oates J, Cunningham D. Why do patients with weight loss have a worse outcome when undergoing chemotherapy for gastrointestinal malignancies? European Journal of Cancer 1998 Mar;34(4):503-509. [doi: 10.1016/s0959-8049(97)10090-9]
2. Bauer JD, Ash S, Davidson WL, Hill JM, Brown T, Isenring EA, et al. Evidence based practice guidelines for the nutritional management of cancer cachexia. Nutrition & Dietetics 2006 Sep;63(s2):S3-S32. [doi: 10.1111/j.1747-0080.2006.00099.x]
3. Bower MR, Martin RC. Nutritional management during neoadjuvant therapy for esophageal cancer. J. Surg. Oncol 2009 Apr 16;100(1):82-87. [doi: 10.1002/jso.21289]
4. Gavazzi C, Colatruglio S, Sironi A, Mazzaferro V, Miceli R. Importance of early nutritional screening in patients with gastric cancer. Br J Nutr 2011 Jun 17;106(12):1773-1778. [doi: 10.1017/s0007114511002509]
5. Murry D, Riva L, Poplack D. Impact of nutrition on pharmacokinetics of anti-neoplastic agents. Int J Cancer Suppl. PMID 1998;11:48-51. [doi: 10.1002/(sici)1097-0215(1998)78:11+<48::aid-ijc14>3.0.co;2-3]
6. Reim D, Friess H. Feeding Challenges in Patients with Esophageal and Gastroesophageal Cancers. Gastrointest Tumors 2016 Jan 8;2(4):166-177. [doi: 10.1159/000442907]
7. Farreras N, Artigas V, Cardona D, Rius X, Trias M, González JA. Effect of early postoperative enteral immunonutrition on wound healing in patients undergoing surgery for gastric cancer. Clinical Nutrition 2005 Feb;24(1):55-65. [doi: 10.1016/j.clnu.2004.07.002]
8. Laviano A, Meguid MM, Rossi-Fanelli F. Cancer anorexia: clinical implications, pathogenesis, and therapeutic strategies. The Lancet Oncology 2003 Nov;4(11):686-694. [doi: 10.1016/s1470-2045(03)01247-6]
9. Marshall KM, Loeliger J, Nolte L, Kelaart A, Kiss NK. Prevalence of malnutrition and impact on clinical outcomes in cancer services: A comparison of two time points. Clinical Nutrition 2019 Apr;38(2):644-651. [doi: 10.1016/j.clnu.2018.04.007]
10. Nitenberg G, Raynard B. Nutritional support of the cancer patient: issues and dilemmas. Critical Reviews in Oncology/Hematology 2000 Jun;34(3):137-168. [doi: 10.1016/s1040-8428(00)00048-2]
11. Skeie E, Tangvik RJ, Nymo LS, Harthug S, Lassen K, Viste A. Weight loss and BMI criteria in GLIM's definition of malnutrition is associated with postoperative complications following abdominal resections – Results from a National Quality Registry. Clinical Nutrition 2020 May;39(5):1593-1599. [doi: 10.1016/j.clnu.2019.07.003]
12. Van Cutsem E, Arends J. The causes and consequences of cancer-associated malnutrition. European Journal of Oncology Nursing 2005 Jan;9:S51-S63. [doi: 10.1016/j.ejon.2005.09.007]

13. Hanna L, Huggins CE, Furness K, Silvers MA, Savva J, Frawley H, et al. Effect of early and intensive nutrition care, delivered via telephone or mobile application, on quality of life in people with upper gastrointestinal cancer: study protocol of a randomised controlled trial. BMC Cancer 2018 Jul 3;18(1). [doi: 10.1186/s12885-018-4595-z]
14. Sekhon M, Cartwright M, Francis JJ. Acceptability of health care interventions: A theoretical framework and proposed research agenda. Br J Health Psychol 2018 Feb 16;23(3):519-531. [doi: 10.1111/bjhp.12295]
15. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1). [doi: 10.1186/s12913-017-2031-8]
16. Hommel KA, Hente E, Herzer M, Ingerski LM, Denson LA. Telehealth behavioral treatment for medication nonadherence. European Journal of Gastroenterology & Hepatology 2013;25(4):469-473. [doi: 10.1097/meg.0b013e32835c2a1b]
17. Understanding Stomach and Oesophageal Cancers 2019. Cancer Council. URL: https://www.cancervic.org.au/cancer-information/types-of-cancer/pancreatic_cancer/pancreatic-cancer-overview.html [accessed 2020-07-08]
18. Understanding Pancreatic Cancer 2020. Cancer Council. URL: https://www.cancervic.org.au/cancer-information/types-of-cancer/pancreatic_cancer/pancreatic-cancer-overview.html [accessed 2020-07-08]
19. Joe J, Demiris G. Older adults and mobile phones for health: A review. Journal of Biomedical Informatics 2013 Oct;46(5):947-954. [doi: 10.1016/j.jbi.2013.06.008]
20. Heart T, Kalderon E. Older adults: Are they ready to adopt health-related ICT? International Journal of Medical Informatics 2013 Nov;82(11):e209-e231. [doi: 10.1016/j.ijmedinf.2011.03.002]
21. Head BA, Keeney C, Studts JL, Khayat M, Bumpous J, Pfeifer M. Feasibility and Acceptance of a Telehealth Intervention to Promote Symptom Management during Treatment for Head and Neck Cancer. The Journal of Supportive Oncology 2011 Jan;9(1):e1-e11. [doi: 10.1016/j.suponc.2010.12.006]
22. Cox A, Lucas G, Marcu A, Piano M, Grosvenor W, Mold F, et al. Cancer Survivors? Experience With Telehealth: A Systematic Review and Thematic Synthesis. J Med Internet Res. PMID 2017;19(1):A. [doi: 10.1136/bmjopen-2019-qhrn.51]
23. Larson JL, Rosen AB, Wilson FA. The Effect of Telehealth Interventions on Quality of Life of Cancer Patients: A Systematic Review and Meta-Analysis. Telemedicine and e-Health 2018 Jun;24(6):397-405. [doi: 10.1089/tmj.2017.0112]
24. Wosik J, Fudim M, Cameron B, Gellad Z, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. Journal of the American Medical Informatics Association. 2020;PMID 2020 Jun 01;27(6):957-962. [doi: 10.1093/jamia/ocaa067]
25. Mann D, Chen J, Chunara R, Testa P, Nov O. COVID-19 transforms health care through telemedicinevidence from the field. Journal of the American Medical Informatics Association. 2020PMID 2020 Jul 01:E. [doi: 10.1093/jamia/ocaa072]
26. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions. ann. behav. med 2013 Mar 20;46(1):81-95. [doi: 10.1007/s12160-013-9486-6]
27. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for Reporting Qualitative Research. Academic Medicine 2014;89(9):1245-1251. [doi: 10.1097/acm.0000000000000388]
28. Kanuha VK. 'Being' native versus 'going native': conducting social work research as an insider. Soc Work 2000 Oct 1;45(5):439-447. [doi: 10.1093/sw/45.5.439] [Medline: 11029899]
29. Finlay L. 'Outing' the Researcher: The Provenance, Process, and Practice of Reflexivity. Qual Health Res 2016 Jul;12(4):531-545. [doi: 10.1177/104973202129120052]
30. Malterud K, Siersma VD, Guassora AD. Sample Size in Qualitative Interview Studies. Qual Health Res 2016 Jul 10;26(13):1753-1760. [doi: 10.1177/1049732315617444]
31. Isenring E, Bauer J, Capra S. The scored Patient-generated Subjective Global Assessment (PG-SGA) and its association with quality of life in ambulatory patients receiving radiotherapy. Eur J Clin Nutr 2003 Feb 6;57(2):305-309. [doi: 10.1038/sj.ejcn.1601552]
32. Naidoo N, Nguyen VT, Ravaud P, Young B, Amiel P, Schanté D, et al. The research burden of randomized controlled trial participation: a systematic thematic synthesis of qualitative evidence. BMC Med 2020 Jan 20;18(1). [doi: 10.1186/s12916-019-1476-5]
33. Lawton J, Blackburn M, Rankin D, Werner C, Farrington C, Hovorka R, et al. Broadening the Debate About Post-trial Access to Medical Interventions: A Qualitative Study of Participant Experiences at the End of a Trial Investigating a Medical Device to Support Type 1 Diabetes Self-Management. AJOB Empirical Bioethics 2019 Apr 15;10(2):100-112. [doi: 10.1080/23294515.2019.1592264]
34. Jongerius C, Russo S, Mazzocco K, Pravettoni G. Research-Tested Mobile Apps for Breast Cancer Careystematic Review. JMIR Mhealth Uhealth. PMID 2019;7(2):e10930. [doi: 10.2196/preprints.10930]
35. Deloitte. Mobile Consumer Survey 2019 ? The Australian Cut. Deloitte 2020:2019.
36. About Cancer, Types of Cancer: Pancreatic Cancer 2019. Cancer Council Australia. URL: https://www.cancer.org.au/about-cancer/types-of-cancer/pancreatic-cancer.html [accessed 2020-08-10]
37. About Cancer, Types of Cancer: Oesophageal Cancer 2019. Cancer Council Australia. URL: https://www.cancer.org.au/about-cancer/types-of-cancer/oesophageal-cancer.html [accessed 2020-08-10]
38. Cosco TD, Firth J, Vahia I, Sixsmith A, Torous J. Mobilizing mHealth Data Collection in Older Adults: Challenges and Opportunities. JMIR Aging 2019 Mar 19;2(1):e10019. [doi: 10.2196/10019]

39. Gücin N, Berk . Technology Acceptance in Health Care: An Integrative Review of Predictive Factors and Intervention Programs. Procedia - Social and Behavioral Sciences 2015 Jul;195:1698-1704. [doi: 10.1016/j.sbspro.2015.06.263]
40. 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. [doi: 10.2196/jmir.8407]
41. Odeh B, Kayyali R, Nabhani-Gebara S, Philip N. Optimizing cancer care through mobile health. Support Care Cancer 2015 Feb 4;23(7):2183-2188. [doi: 10.1007/s00520-015-2627-7]
42. Keinki C, Seilacher E, Ebel M, Ruetters D, Kessler I, Stellamanns J, et al. Information Needs of Cancer Patients and Perception of Impact of the Disease, of Self-Efficacy, and Locus of Control. J Canc Educ 2015 May 22;31(3):610-616. [doi: 10.1007/s13187-015-0860-x]

Abbreviations

ICT: information and communication technology
RCT: randomized controlled trial
UGI: upper gastrointestinal

Edited by G Eysenbach; submitted 30.08.20; peer-reviewed by E Laing, R Poluru; comments to author 31.10.20; revised version received 29.11.20; accepted 19.12.20; published 12.03.21.

Please cite as:
Furness K, Huggins CE, Truby H, Croagh D, Haines TP
Attitudes of Australian Patients Undergoing Treatment for Upper Gastrointestinal Cancers to Different Models of Nutrition Care Delivery: Qualitative Investigation
JMIR Form Res 2021;5(3):e23979
URL: https://formative.jmir.org/2021/3/e23979
doi:10.2196/23979
PMID:33709939

Original Paper

Exploration of Gender-Sensitive Care in Vocational Rehabilitation Providers Working With Youth With Disabilities: Codevelopment of an Educational Simulation

Sally Lindsay[1,2], BA, MA, PhD; Kendall Kolne[1], BSc, MSc, PhD; Donna J Barker[2], BSc, MSc; Angela Colantonio[2], BSc, MSc, PhD; Jennifer Stinson[3,4], BSc, MSc, PhD; Sandra Moll[5], BSc, MSc, PhD; Nicole Thomson[6], BHSc, MHSc, PhD

[1]Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, ON, Canada

[2]Department of Occupational Science & Occupational Therapy, University of Toronto, Toronto, ON, Canada

[3]Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada

[4]Hospital for Sick Children, Toronto, ON, Canada

[5]School of Rehabilitation Sciences, McMaster University, Hamilton, ON, Canada

[6]Centre for Addiction and Mental Health, Toronto, ON, Canada

Corresponding Author:
Sally Lindsay, BA, MA, PhD
Bloorview Research Institute
Holland Bloorview Kids Rehabilitation Hospital
150 Kilgour Road
Toronto, ON, M4G 1R8
Canada
Phone: 1 416 425 6220 ext 3654
Fax: 1 416 425 1634
Email: slindsay@hollandbloorview.ca

Abstract

Background: Although research shows that there is a need for gender-specific vocational support to help youth with disabilities find employment, health care providers often report needing more training in this area. Currently, there are no existing educational simulations of gender-sensitive care within vocational rehabilitation for clinicians who provide care to youth with disabilities. Therefore, developing further educational tools that address gender-sensitive care could help them enhance the care they provide while optimizing patient outcomes.

Objective: This study aims to codevelop an educational simulation and identify issues relevant to providing gender-sensitive care within the context of vocational rehabilitation for youth with disabilities.

Methods: We used a qualitative co-design approach with a purposive sampling strategy that involved focus group discussions and journal reflections to understand and address issues relevant to gender-sensitive care within vocational rehabilitation for those working with youth with disabilities. A total of 10 rehabilitation providers participated in two sessions (5 participants per session) to design the web-based simulation tool. The sessions (2.5 hours each) were audio recorded, transcribed, and analyzed thematically.

Results: Two main themes arose from our analysis of codeveloping a simulation focusing on gender-sensitive care. The first theme involved the relevance of gender within clinical practice; responses varied from hesitance to acknowledging but not talking about it to those who incorporated gender into their practice. The second theme focused on creating a comfortable and safe space to enable gender-sensitive care (ie, included patient-centered care, effective communication and rapport building, appropriate language and pronoun use, respecting gender identity, awareness of stereotypes, and responding to therapeutic ruptures).

Conclusions: Our web-based gender-sensitive care simulation that addressed vocational rehabilitation among youth with disabilities was cocreated with clinicians. The simulation highlights many issues relevant to clinical practice and has potential as an educational tool for those working with young people with disabilities.

(JMIR Form Res 2021;5(3):e23568) doi:10.2196/23568

KEYWORDS

continuing education; gender-identity; gender-sensitive care; rehabilitation

Introduction

Background

Within Canada, there are more than 200,000 youth aged between 15 and 25 years who have a disability [1]. Such youth often encounter significant challenges in finding employment, including discriminatory attitudes, inadequate transportation, and inaccessible workspaces [2,3]. Consequently, youth with disabilities often have higher unemployment rates than those without disabilities [4]. Focusing on the youth is salient because many disadvantages are often compounded for those who start life with a disability or acquire it at an early age [5,6]. For example, youth with disabilities often encounter challenges with developmental tasks, social development, and role functioning [5,7]. In addition, this developmental period is characterized by identity exploration, instability, and continued development of executive functioning, which is important for enhancing job skills and independence [5]. Furthermore, young adulthood is a critical time for optimizing work-based identities and positive behaviors [2,7]. Unfortunately, youth with disabilities often do not receive the appropriate support they need to find meaningful employment. For example, even though clinicians who work with such youth acknowledge gender differences within their practice, many admit to not providing appropriate or gender-specific support to help them with their vocational goals [8]. Furthermore, recent research shows that sex-specific vocational support is needed to help youth as they transition into employment roles [8,9]. Therefore, further training tools are needed to help clinicians provide gender-sensitive care in vocational rehabilitation for youth with disabilities.

Gender and Vocational Rehabilitation for Youth With Disabilities

Gender plays an important role within vocational rehabilitation for youth with disabilities. Research shows that gender often shapes how people with disabilities cope with engaging in vocational training and employment [8-12]. For example, a recent systematic review on the role of gender in employment among youth with disabilities found that young women with disabilities continue to lag behind their male peers on several health and social outcomes, including lower employment rates and multiple forms of discrimination [8,9,13,14]. Young women with disabilities often encounter particular barriers in career development, including limited vocational training, lower family expectations, disability stereotypes, and decreased self-confidence [15,16]. Although gender is becoming increasingly important, the intersection of disability and youth employment has received little attention [17]. Understanding this intersection is critical because inequalities in employment are significant for individuals with disabilities who identify as men, women, gender fluid, or nonbinary [18].

Of the limited research exploring gender and employment among youth with disabilities, the focus is often on employment outcomes (eg, working or unemployed) or pay differences [8,9]. Such studies also tend to only look at gender from a binary (ie, man or woman) perspective, which is problematic because it does not capture the diversity of gender identity experiences. Thus, there is a need for further research to unpack the complex relationship between gender, vocational rehabilitation, and employment.

Although there are several systematic reviews focusing on youth with disabilities and employment, there is very little, if any, mention of gender [13,17,19-21]. Meanwhile, several studies highlight the important need for gender-specific vocational support for youth with disabilities [13,22-24]. Exploring gender within the context of vocational rehabilitation for youth with disabilities is important for decision making, enhanced communication, and engagement in programs and interventions [25].

Gender-Sensitive Care

Despite the importance of gender (ie, socially constructed roles, behaviors, and identities) within health care, many health care providers report that they lack knowledge on how to provide gender-sensitive care (ie, knowledge and competence about the role of gender within clinical practice) [26], including within rehabilitation settings [27]. Such a shortage of knowledge in this area is concerning because it can influence health outcomes, health inequalities, incorrect or delayed diagnoses, and suboptimal therapies [28-30].

Gender-sensitive approaches aim to incorporate specific health care needs of men and women while aiming to address gender-based health inequities and methods of transforming harmful gender norms, roles, and relations while also centering on promoting gender equality [31]. Gender sensitivity involves understanding gendered patterns of individual experiences of boys and men as well as girls and women and the unique experiences of those who have identities that are nonbinary or gender fluid; it also considers their morbidity and mortality while reflecting on the broader sociopolitical and cultural context of where health care takes place [26,32]. Although research on gender-sensitive care is increasing, it mainly focuses on adult and acute care populations, with little mention of pediatrics or rehabilitation, especially within the context of vocations [33].

Research consistently highlights that gender plays a critical role in the incidence, clinical presentation, manifestation, and health outcomes; yet, youth with disabilities are often viewed and treated as without gender and asexual [34-38]. For example, youth who identify as lesbian, gay, bisexual, transgender, or queer (LGBTQ+) have been largely ignored in vocational rehabilitation research [8,26,28]. This trend is concerning because sexual and gender minority groups often encounter substantial barriers within the health care system including discrimination, which could affect their health and vocational outcomes [39-44].

A recent review on gender-sensitive educational interventions for health care providers found that few studies focused on rehabilitation care providers and none focused on vocational

rehabilitation for youth with disabilities. Currently, there are no existing gender-sensitive care educational simulations within vocational rehabilitation for clinicians who provide care to youth with disabilities [28]. Therefore, developing further educational tools in gender-sensitive care could help them enhance the care they provide while optimizing patient outcomes.

Simulations as a Professional Development Tool

There is currently a lack of consensus regarding the best educational approaches to teach gender-sensitive care [45,46]. One potential way to address the need for further education to health care providers is through simulations, which refer to life-like environments and contrived social situations mimicking problems or conditions arising in professional encounters [47]. Simulations are a critical tool for interprofessional training that can help health care providers enhance their clinical skills, teamwork, and communication [48,49]. Research highlights that simulation-based learning can be a positive educational experience that enhances patient outcomes [50-52]. Although health care providers frequently participate in simulations, they are rarely engaged in the design and development process [47,53,54]. Including health care providers in building a simulation, as opposed to just watching or participating in one, could also help to create relevant scenarios [53,55] while reinforcing clinical skills [47]. This study aims to involve health care providers in cocreating simulations.

Our simulation is novel because it was informed by several needs assessments from both health care providers and youth perspectives and is one of the first to explore gender-sensitive care among vocational rehabilitation care providers who work with youth with disabilities [56]. Developing further training for health care providers could help address gender-based health inequalities [46].

Objective

The objective of our study was to codevelop an educational simulation and identify issues relevant to providing gender-sensitive care within the context of vocational rehabilitation for youth with disabilities.

Methods

Design

We used an interpretive descriptive qualitative design, which is particularly valuable in applied clinical health studies [57]. We conducted focus group discussions and participant journal reflections to understand participants' experiences and perspectives of building the simulation. This design was appropriate because it is consistent with the development of professional simulation tools [49,53].

Sample and Recruitment

This study was conducted at an academic pediatric rehabilitation hospital located in a large urban center where we received institutional ethics approval. All participants provided written consent before taking part following approval by the local research ethics board. Using a purposive sampling strategy, participants were recruited through invitation letters, referrals, or advertisements at a pediatric rehabilitation hospital. Participants who were interested were screened to meet the following eligibility criteria: a health care provider, practitioner, or trainee who had relevant experience in helping young people with disabilities to find employment. A total of 10 participants (9 women and 1 man) participated in the simulation build exercises (5 in each build session; Table 1). Note that not all participants took part in both sessions. We recognize that the gender composition of most of the participants in this sample is female; however, this is consistent with the gender composition of pediatric rehabilitation care providers [58]. Our sample size aligns with recommendations for the optimal size for conducting simulations and focus group methodology [59,60].

Table 1. Overview of participants.

Participant	Profession
1	Occupational therapist
2	Occupational therapist
3	Social worker
4	Social worker
5	Social worker
6	Vocational rehabilitation counselor
7	Vocational rehabilitation counselor
8	Occupational therapist
9	Vocational rehabilitation counselor
10	Occupational therapist

Procedure for Developing the Web-Based Simulation

The simulation codevelopment sessions occurred over two 2.5-hour sessions, held in March and November 2019, and were facilitated by researchers (ie, the first two authors, SL and KK) who were certified in *SIM-one* simulations (ie, briefing, debriefing, and facilitating). The first session (n=5) focused on building the simulation scenario content, whereas the second

session (n=5) piloted the scenario with live actors (ie, simulated participants) who were trained for their character roles before participating in the session. The authors had no previous relationship with the study participants.

First Build Session

The initial discussion with participants was informed by needs assessments and systematic reviews conducted by our team, focusing on the role of gender among youth with disabilities [8,9,13,17,61-64] and health care providers [13,64-66]. The rationale for this was to share evidence-informed, first-hand experiences. In particular, our discussion addressed research evidence about gender-sensitive approaches to engaging youth in rehabilitation (eg, rapport development, gender role expectations, self-advocacy, disclosure decisions) and self-reflection on gender in clinical practice, in addition to the lived experience perspectives of young people with disabilities. After giving participants an overview of the topic, we led a focus group discussion that allowed participants to reflect on their own clinical experiences regarding gender and share relevant issues.

Next, we gave participants a brief orientation on the process of building a simulation, which was led by a professionally trained simulation educator. We then followed a simulation scenario template (Multimedia Appendix 1), which guided participants through the learning objectives for building a simulation. Then participants were asked to help develop the content for the simulation by using a template, while a simulation educator and researchers guided a discussion about what the content would entail. Participants shared their experiences of how they addressed their gender in their practice. After the completion of each simulation, we had a focus group discussion (ie, simulation debrief) [59] (Multimedia Appendix 2 [53]). Participants then shared their reactions to the simulation content, its relevance to health care providers, and any recommendations they had for further development. After the first simulation session, the researchers then worked with a simulation educator and standardized participants to finalize a scenario script. We also met with a youth who has a disability and who also identified as transgender, and they provided feedback on our scenario, which we incorporated into the second build session.

Second Build Session

In the second simulation build session, we described the scenario template that the participants built in the first session. This session involved simulated participants who were trained for their character roles and then piloted the simulation with feedback from the health care provider participants. We then piloted the disability disclosure scenario, with feedback from participants through a guided focus group discussion (Multimedia Appendix 2). Next, we incorporated all participant feedback into the final version of the simulation, which will be filmed for further professional development purposes.

Data Analysis

Both simulation codevelopment sessions were audio recorded, transcribed verbatim, and verified by the researchers who attended the simulation sessions. Two researchers independently reviewed the transcripts and developed the codes using an open-coding thematic approach [67]. The research question guided the analysis, where we looked for codes regarding gender-sensitive care. This approach involved the first two authors (with backgrounds in pediatric rehabilitation and gender-sensitive care) independently reading the transcripts and familiarizing themselves with the data, generating initial codes, and revising and defining the themes. We developed a list of preliminary codes while noting the patterns between them and then met to compare the codes. During this discussion, we split, merged, and relabeled codes until we reached an agreement in the coding tree. Team discussions also helped resolve any discrepancies in the organization of the themes. We checked for thematic saturation and felt that this was achieved in our codes and themes [67,68]. The first author, with experience in qualitative research, applied the final coding scheme to the transcripts. Relevant quotes that reflected each theme and subtheme were extracted [67]. Strategies to enhance rigor and trustworthiness of the findings included prolonged engagement, descriptive participant accounts, and peer debriefing discussions among the research team who have expertise in rehabilitation, pediatrics, and gender [68]. We also kept notes on key decisions made throughout the data analysis. The research team also reflected on their own biases and assumptions and interests in this topic and how this may have influenced their selection of the themes, which we noted in our audit trail [67,68].

Results

Overview

Two main themes emerged from our analysis of codeveloping a simulation on gender-sensitive care within vocational rehabilitation for youth with disabilities including (1) the relevance of gender within clinical practice and (2) creating a comfortable and safe space to enable gender-sensitive care (an overview of themes and subthemes is given in Multimedia Appendix 3).

The Relevance of Gender Within Clinical Practice

During the first build session, most participants reflected on the extent to which gender was addressed within their clinical practice, ranging from hesitance to acknowledging but not talking about it to acknowledging and incorporating gender into practice.

Hesitant and Resistant

In the first build session, many participants resisted the idea that gender even mattered to their work. For example, one participant said:

> *I was sort of initially struggling with reflecting on this...I don't think it is ever the first thing in my mind...I don't think I necessarily have come across a blatant difference.* [#2, simulation 1]

Some participants hesitated at first about the possibility of gender differences and the role of gender within health care, despite learning about evidence of the importance of such differences. To illustrate, one health care provider reported:

> *I don't see there's a gender difference...I do not make any distinctions between genders.* [#3, simulation 1]

Others mentioned that they had never really considered the role of gender in their practice until participating in this study. Meanwhile, some participants stated that they considered gender within their practice but did not notice any gender-related patterns. For example, one said:

> *I don't think about it enough systematically... I'm not sure I've ever noticed anything at the systems level surrounding gender; Or, that it's ever been front of mind in any discussion with any families and youth.* [#1, simulation 1]

Gender Is Acknowledged But Not Openly Discussed

Some participants noted that gender was sometimes acknowledged within their practice but rarely explicitly discussed. For instance, one health care provider described, "at some level it may connect to gender. I've never brought that out into the open and I never had anyone else bring that into the open" (#2, simulation 1). Another health care provider concurred: "It's interesting in the sense it's being talked about, but not directly at the same time" (#5, simulation 1).

Gender in Clinical Practice

Some participants reported noticing gender differences (mostly binary) in their practice. For instance, a provider described, "over time we have possibly seen a slight increase in young men we're serving in our programs" (#1, simulation 1). Another participant shared:

> *Autism, we do find in terms of diagnoses that females are diagnosed much older and later in life. Given a lot of the symptoms are masked...We're seeing some differences there. With boys it seems a bit easier...the age of diagnosis is a bit younger...I'm just realizing that I am working mostly with males than females.* [#3, simulation 1]

Meanwhile, some participants mentioned gender differences around safety concerns and parental overprotection, particularly for female patients. To illustrate, one participant said:

> *I have families who will say, their experience outside of home and school the first things on their mind are around safety and vulnerability in the community...I can perhaps see a connection to gender because sometimes individuals have that concern more so for females.* [#1, simulation 1]

Furthermore, some participants noticed patterns in the career pathways of youth with disabilities, with females sometimes going into social sciences and males into computers. For example, a health care provider remarked: "I'm more often working with young men who are interested in computers and I don't think I've had a young lady in my career who has said that's what she's dreamt of doing" (#4, simulation 1). The journal reflections indicated that building a simulation on gender-sensitive care helped encourage participants to share their reflections on this topic from an interprofessional perspective while also learning from each other.

Creating a Comfortable and Safe Space to Enable Gender-Sensitive Care

The second main theme involved creating a comfortable and safe space to enable gender-sensitive care, which included patient-centered care, effective communication and rapport building, appropriate language and pronouns, respecting gender identity, and responding to therapeutic ruptures.

Patient-Centered Care

Participants highlighted the importance of applying patient-centered principles in providing gender-sensitive care. For example, one participant explained:

> *How we could enable people to hear and value and know the person first...sort of a person-centered start to a therapeutic relationship...We want to maximize their engagement.* [#1, simulation 1]

Another health care provider, who had some informal training in gender, explained how she also used a patient-centered approach to create a gender-sensitive environment:

> *People can easily sense when you're uncomfortable; Or, you're not sure how to approach the situation...It helps to open up the conversation as clinicians become more comfortable having a discussion and often times people don't open up to others...Maybe comments and phrases people use that they don't feel it's a safe space...It comes from your authenticity and how you communicate that.* [#5, simulation 1]

Meanwhile, one participant shared how it might be helpful to use a patient-centered or solution-focused approach:

> *If they're a little harder to read it shows how we have to do some exploring...I would make that more open-ended, like in solution-focused coaching we ask what are your best hopes? This is what I'm here for...What are you hoping of getting out of meeting today? and giving more open-ended questions so it's giving space for that to come forward.* [#2, simulation 2]

Effective Communication and Building Rapport

Health care providers highlighted the importance of providing a comfortable and safe space that involved effective communication, developing rapport, listening, and understanding. Most participants agreed that developing rapport was the building block for effective communication. For instance, one participant said:

> *You have to establish rapport. So, the best, easiest way to do that from this perspective is to get people talking about themselves.* [#6, simulation 2]

Others agreed:

> *I don't think you're going to get anywhere without rapport.* [#8, simulation 2]

Another health care provider mentioned how to start the conversation:

> *Creating an open space where somebody can tell us their wishes and how to best support them...That lends*

> *itself to the conversation about gender...If we foster those conversations that were based on interests regardless of whether it is your son or your daughter...putting it in a neutral way...will give them the opportunity to explore this for themselves...having an atmosphere that doesn't limit by gender.* [#4, simulation 1]

Participants described how their role involves getting patients to open up and they can do this by communicating effectively and building rapport. Others concurred that communication and rapport were important in providing gender-sensitive care. To illustrate, one health care provider shared:

> *In this whole process, we are probing for information. So, we are trying to poke holes at where can I connect with this person...trying to help them with their goal...and to get them to open up.* [#10, simulation 2]

Appropriate Language and Gender-Related Pronouns

Another part of creating a comfortable and safe space to enable gender-sensitive care involves using appropriate gender-related pronouns with patients. For example, one participant shared:

> *A lot of emphasis is being placed on gender and specifically on pronoun use and transgender. People in the program have shared lived experiences, and being misgendered...adds a lot of expectations that society might place on somebody thinking you're female, but they might not identify as female.* [#5, simulation 1]

Participants described how exploring gender identity might play out while considering when they might ask patients about pronouns. For example:

> *It depends who you're working with, like this client also seems really shy and quiet so you wouldn't want to be like, outright; but sometimes at the beginning, especially now, it's happening a lot more; You can ask people their pronouns, on their forms, but it would obviously be super awkward. I would not recommend saying "oh never mind what are your pronouns?"; because that is kind of weird, but people are asking now what are your pronouns.* [#7, simulation 2]

One participant explained how they would respond if they had to deal with the issue of completing paper work and identifying a patient's gender identity. They said:

> *I have to submit your information to the government in order for you to participate in our program. They have to check male or female or other. So, I see you checked male. When we interact together do you have a preferred name that is maybe different than your legal name? and a preferred pronoun, that may be different than what you wrote on this form? I have found that the individuals who are highly politicized we use language like "this is my identity; this is my pronoun. I don't identify as...my gender x." They will use language like that.* [#6, simulation 2]

Respecting Gender Identity

Creating a comfortable and safe space involved participants respecting a patient's gender identity including their values and preferences. Some participants described their experiences of working with gender-diverse patients and the potential challenges of misgendering a patient. To illustrate, one health care provider explained, "It's trying really hard to be sensitive around identifying people with the way they see themselves and being respectful" (#5, simulation 1).

In the second simulation build session, several participants debated whether the gender identity of the patient was relevant. For instance, one participant said, "I wasn't sure when the youth (i.e., simulated participant) walked in whether or not I could put them into a category of male or female" (#9, simulation 2). Some participants commented that patients they know who are nonbinary are "used to people making assumptions, misgendering and so on. So, the best thing to do is catch it, as soon as you can" (#6, simulation 2).

Meanwhile, others described how it can be difficult to know the gender identity of a patient when there are often only binary choices (ie, male or female) on their chart, which may have been checked off by their parents when the child was originally registered at the hospital at a much younger age.

Awareness of Gender Stereotypes and Gender-Diverse Clients

Codeveloping a simulation on gender-sensitive care encouraged clinicians to share their reflections on this topic from an interprofessional perspective while learning from each other. For example, clinicians have demonstrated an awareness of gender stereotypes and gender-diverse clients. One clinician stated:

> *I'm certainly sensitive to it if I hear a parent try to compartmentalize their child...Oh, you're a girl; you can't possibly be a NASA space astronaut...hoping to not limit people based on gender.* [#4]

Others described the experience of working with gender-diverse patients and the potential problems of misgendering a patient. To illustrate, one said, "It's trying really hard to be sensitive around identifying people with the way they see themselves and being respectful of that" (#5). Others have explained the potential problems of misgendering a client. For instance, a clinician explained:

> *Given that gender can be a big part of someone's identity when someone has been misgendered based on how they look, you've kind of reduced their identity just to how they look and that makes conversations difficult.* [#3]

Responding to Therapeutic Ruptures

Participants told us how mistakes or "therapeutic ruptures" could happen as a result of a lack of knowledge about gender-sensitive care, particularly for patients who identify as gender diverse. A health care provider explained:

> *If you have made a mistake and maybe used the wrong pronoun, I think it's how you create this space if*

> *you're genuinely coming from the place of curiosity and wanting to learn. Then, often times in my experience people have been very receptive.* [#5, simulation 1]

Another shared:

> *Unexpectedness may be something that happens, if we have two major genders: male and female and someone brings us a new way of describing themselves and identifying themselves and having the language to engage and handle our own response in the professional and sensitive way...How to fix it, and how to do it right?* [#3, simulation 1]

Participants emphasized the importance of acknowledging any communication mistakes related to gender and working to rebuild a rapport with patients. Furthermore, providing gender-sensitive care involved recognizing and dealing with a potential unconscious bias. A health care provider shared:

> *Before we see a client, we look in the file. The first things we see are their name, age, gender, and diagnosis...Even noticing in my own practice we have lots of unconscious biases. I think, I'm going to see a sixteen-year-old boy with this diagnosis then I have a sense of this is what they're going to be like. I do find myself not really exploring their gender identity if it's already stated. However, if it said "other" I don't have a lot of experience with, transgender female or male. I'd be more keen to explore, what is that like?* [#5, simulation 1]

In summary, participants described how important it was for them to acknowledge any mistakes or therapeutic ruptures related to gender identity.

Discussion

Principal Findings

Our study addressed an important gap in the literature by exploring how issues relevant to gender-sensitive care among vocational rehabilitation care providers within the context of building a simulation for people working with youth with disabilities. Focusing on gender-sensitive care is important because it is a key social determinant of health and a component of patient-centered care. Offering training in this area for health care providers could help to address health inequalities and optimize patient outcomes [46,69]. Our findings showed that the relevance of gender within clinical practice ranged considerably from hesitance to acknowledging it but not talking about it to incorporating gender into their practice. Research indicates that health care providers need an ongoing awareness of their values, assumptions, and beliefs and how these relate to their patients, which may be developed through reflective practice [53,70]. Furthermore, studies also show that having knowledge about gender is often associated with more positive attitudes and enhanced patient care [26,71,72].

Our findings highlight that creating a comfortable and safe space is a salient aspect of gender-sensitive care. In particular, we found the following elements to be important for creating a safe space: patient-centered care, effective communication and rapport building, appropriate language and pronoun use, respecting gender identity, and responding to therapeutic ruptures. These findings are consistent with culturally sensitive care models that emphasize the importance of self-awareness and sensitivity to one's own values, biases, and power differences with patients while also understanding their values and priorities [73,74]. Other research demonstrates the importance of the awareness of health care providers of any biases they have because it could influence health care decisions [75].

Our findings also revealed that addressing therapeutic ruptures regarding gender, and particularly misgendering someone, was an important aspect of gender-sensitive care. Our results are consistent with other research highlighting that failing to identify a patient by affirming their name and pronoun within a medical setting can impact patient satisfaction and quality of care, particularly for LGBTQ+ patients [76]. Gender-related therapeutic ruptures could be a result of binary documentation (ie, male or female-only options), which could potentially misgender a patient, causing stress or stigmatization [28]. It is important that health care providers address gender diversity in their clinical practice. The World Health Organization advocates that sex and gender are addressed both within and outside of health care [31].

Our results highlight that effective communication is an important component of gender-sensitive care. This finding aligns with the principles of patient-centered care and culturally sensitive care, which require health care providers to be open to information from patients about their needs, expectations, and preferences while listening and respecting the patients [77]. Although many clinicians are trained in patient-centered care, it may be worthwhile to have more specific training in gender-sensitive care. Research indicates that communication plays a critical role in therapeutic sessions, ongoing patient-clinician relationships, and patient satisfaction [77,78]. Effective communication could help health care providers to better understand patients' needs and priorities, which could help them to tailor their recommendations [77].

Most of our findings focused on issues related to gender identity and, to some extent, gender roles; however, there was little discussion of gender relations or institutionalized gender (eg, power relations between men and women) [25]. It is important to recognize that gender is part of a larger sociopolitical and cultural context and that health care organizations themselves are gendered [26]. Surprisingly, few participants reflected on how their own gender could potentially impact the development of rapport with patients. Further unpacking this trend is important because most Canadian health care providers in pediatric rehabilitation are female, whereas most pediatric rehabilitation patients are male [79]. In addition, recognizing gender differences within the health care context along with understanding the role of one's own gender within professional practice are critical for avoiding and reinforcing gender stereotypes [29,80,81].

Participants in our study had the opportunity to reflect on their own skills in gender-sensitive care. Building a simulation (Multimedia Appendix 1) allowed participants to reflect on this

topic within an interprofessional setting while listening to, and learning from, other participants and researchers. Participants identified gender-sensitive approaches by sharing experiences and perspectives regarding gender within their practice within the format of building a professional development simulation, which created an opportunity for participants to work directly with live simulated participants. Our study is novel in that few studies have involved simulation facilitators in the process of building a simulation but instead rely on predeveloped character roles [49]. Other research shows that using patient narratives is a useful tool for reflection among health care providers [82]. Our findings were consistent with other studies indicating that interprofessional discussions of case scenarios helped to broaden individual perspectives from different professions within a safe environment [77].

Limitations

Our study has some limitations that are important to acknowledge. First, the majority of the participants in our study were women; however, this gender-unbalanced sample reflects the gender composition of health care providers within pediatric rehabilitation (ie, majority are women) [58,83]. Second, we used a small, qualitative sample drawn from one city. Thus, further testing with a larger sample and long-term outcomes (ie, changes in knowledge and practice) are needed. Future studies should consider evaluating the impact of the simulation on practices of health care providers. Third, the participants who chose to participate in our study may have had more of an interest in or experience with disability and were not representative of all pediatric rehabilitation health care providers. In addition, it is also important to recognize that gender norms and values often vary by culture, which is an area of further exploration.

Conclusions

This study explored perspectives of health care providers on relevant issues to provide gender-sensitive care within pediatric rehabilitation within the context of building a professional development simulation tool. Our findings showed two key themes, including the relevance of gender within clinical practice and creating a comfortable and safe space (ie, practicing patient-centered care, effective communication and rapport building, respecting gender identity and being aware of stereotypes, and responding to therapeutic ruptures) to enable gender-sensitive care. Future research should consider exploring the impact that simulations may have on behaviors in clinical practice.

Acknowledgments

The authors would like to thank the project partners for their contributions to this toolkit and the staff and trainees in the Transitions and Inclusive Environments laboratory for their support in this project.

This study was supported, in part, by the Canadian Institutes of Health Research-Social Sciences and Humanities Research Council (CIHR-SSHRC) Partnership Grant (01561-000 and 895-2018-4002) awarded to SL and the Kimel Family Fund through the Holland Bloorview Kids Rehabilitation Hospital.

Authors' Contributions

SL, DB, AC, JS, SM, and NT conceived the study and developed the initial study protocol. KK assisted with data collection and analysis. SL wrote the protocol for publication and provided an overview of data collection during the study. All authors read and approved the final protocol for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Simulation scenario template.
[DOCX File , 22 KB - formative_v5i3e23568_app1.docx]

Multimedia Appendix 2
Focus group discussion guide.
[DOCX File , 20 KB - formative_v5i3e23568_app2.docx]

Multimedia Appendix 3
Table overview of themes.
[DOCX File , 15 KB - formative_v5i3e23568_app3.docx]

References

1. Disability in Canada: A 2006 Profile. Statistics Canada. URL: https://www.canada.ca/en/employment-social-development/programs/disability/arc/disability-2006.html [accessed 2021-03-04]

2. Lindsay S, McDougall C, Sanford R, Menna-Dack D, Kingsnorth S, Adams T. Exploring employment readiness through mock job interview and workplace role-play exercises: comparing youth with physical disabilities to their typically developing peers. Disabil Rehabil 2015;37(18):1651-1663. [doi: 10.3109/09638288.2014.973968] [Medline: 25323394]
3. Lindsay S. Discrimination and other barriers to employment for teens and young adults with disabilities. Disabil Rehabil 2011;33(15-16):1340-1350. [doi: 10.3109/09638288.2010.531372] [Medline: 21067349]
4. Breaking barriers to employment for people with disabilities. National Disability Council. URL: https://dds.dc.gov/page/breaking-barriers-employment-people-disabilities [accessed 2019-02-20]
5. Arnett JJ. Emerging adulthood: what is it, and what is it good for? Child development perspectives. Child Dev Perspect 2007;1(2):68-73. [doi: 10.1111/j.1750-8606.2007.00016.x]
6. Poppen M, Lindstrom L, Unruh D, Khurana A, Bullis M. Preparing youth with disabilities for employment: an analysis of vocational rehabilitation case services data. J Vocat Rehabil 2017 Feb 16;46(2):209-224. [doi: 10.3233/jvr-160857]
7. Lindsay S, McPherson AC, Maxwell J. Perspectives of school-work transitions among youth with spina bifida, their parents and health care providers. Disabil Rehabil 2017 Apr;39(7):641-652. [doi: 10.3109/09638288.2016.1153161] [Medline: 26987863]
8. Lindsay S, Cagliostro E, Albarico M, Mortaji N, Srikanthan D. Gender matters in the transition to employment for young adults with physical disabilities. Disabil Rehabil 2019 Feb;41(3):319-332. [doi: 10.1080/09638288.2017.1390613] [Medline: 29041821]
9. Lindsay S, Cagliostro E, Albarico M, Srikanthan D, Mortaji N. A systematic review of the role of gender in securing and maintaining employment among youth and young adults with disabilities. J Occup Rehabil 2018 Jun;28(2):232-251. [doi: 10.1007/s10926-017-9726-x] [Medline: 28836051]
10. Lindsay S, Proulx M, Maxwell J, Hamdani Y, Bayley M, Macarthur C, et al. Gender and transition from pediatric to adult health care among youth with acquired brain injury: experiences in a transition model. Arch Phys Med Rehabil 2016 Feb;97(2 Suppl):S33-S39. [doi: 10.1016/j.apmr.2014.04.032] [Medline: 25660004]
11. Lindstrom L, Doren B, Post C, Lombardi A. Building Career PATHS (Postschool Achievement Through Higher Skills) for young women with disabilities. Career Dev Q 2013 Dec 09;61(4):330-338 [FREE Full text] [doi: 10.1002/j.2161-0045.2013.00059.x]
12. Gibson BE, Mistry B, Smith B, Yoshida KK, Abbott D, Lindsay S, et al. Becoming men: gender, disability, and transitioning to adulthood. Health (London) 2014 Jan;18(1):95-114. [doi: 10.1177/1363459313476967] [Medline: 23456143]
13. Lindsay S, Cagliostro E, Carafa G. A systematic review of workplace disclosure and accommodation requests among youth and young adults with disabilities. Disabil Rehabil 2018 Dec;40(25):2971-2986. [doi: 10.1080/09638288.2017.1363824] [Medline: 28797182]
14. Powers K, Hogansen J, Geenen S, Powers L, Gil-Kashiwabara E. Gender matters in transition to adulthood: a survey study of adolescents with disabilities and their families. Psychol Schs 2008 Apr;45(4):349-364 [FREE Full text] [doi: 10.1002/pits.20297]
15. Doren B, Gau J, Lindstrom L. The role of gender in the long-term employment outcomes of young adults with disabilities. J Vocat Rehabil 2011;34(1):35-42. [doi: 10.3233/jvr-2010-0532]
16. Lindstrom L, Harwick RM, Poppen M, Doren B. Gender gaps. Career development and transition for exceptional individuals 2012 Jan;35(2):108-117. [doi: 10.1177/2165143412437737]
17. Hanif S, Peters H, McDougall C, Lindsay S. A systematic review of vocational interventions for youth with disabilities. Disab Soc Sci 2017;10. [doi: 10.1108/s1479-354720170000010008]
18. Roebroeck ME, Jahnsen R, Carona C, Kent RM, Chamberlain MA. Adult outcomes and lifespan issues for people with childhood-onset physical disability. Dev Med Child Neurol 2009 Aug;51(8):670-678 [FREE Full text] [doi: 10.1111/j.1469-8749.2009.03322.x] [Medline: 19627341]
19. Brohan E, Henderson C, Wheat K, Malcolm E, Clement S, Barley EA, et al. Systematic review of beliefs, behaviours and influencing factors associated with disclosure of a mental health problem in the workplace. BMC Psychiatry 2012 Feb 16;12:11 [FREE Full text] [doi: 10.1186/1471-244X-12-11] [Medline: 22339944]
20. McDowell C, Fossey E. Workplace accommodations for people with mental illness: a scoping review. J Occup Rehabil 2015 Mar;25(1):197-206. [doi: 10.1007/s10926-014-9512-y] [Medline: 24841728]
21. Rusch FR, Dattilo J. Employment and self-management: a meta-evaluation of seven literature reviews. Intellect Dev Disabil 2012 Feb;50(1):69-75. [doi: 10.1352/1934-9556-50.1.69] [Medline: 22316227]
22. Lindsay S, McDougall C, Sanford R. Disclosure, accommodations and self-care at work among adolescents with disabilities. Disabil Rehabil 2013;35(26):2227-2236. [doi: 10.3109/09638288.2013.775356] [Medline: 23594050]
23. Sung C, Sánchez J, Kuo H, Wang C, Leahy MJ. Gender differences in vocational rehabilitation service predictors of successful competitive employment for transition-aged individuals with autism. J Autism Dev Disord 2015 Oct;45(10):3204-3218. [doi: 10.1007/s10803-015-2480-z] [Medline: 26060047]
24. Schmidt MA, Smith DL. Individuals with disabilities perceptions on preparedness for the workforce and factors that limit employment. Work 2007;28(1):13-21. [Medline: 17264416]
25. Tannenbaum C, Greaves L, Graham ID. Why sex and gender matter in implementation research. BMC Med Res Methodol 2016 Oct 27;16(1):145 [FREE Full text] [doi: 10.1186/s12874-016-0247-7] [Medline: 27788671]

26. Celik H, Lagro-Janssen TALM, Widdershoven GGAM, Abma TA. Bringing gender sensitivity into healthcare practice: a systematic review. Patient Educ Couns 2011 Aug;84(2):143-149. [doi: 10.1016/j.pec.2010.07.016] [Medline: 20719461]
27. Hanafy S, Amodio V, Haag HL, Colquhoun H, Lewko J, Quilico E, et al. Is it prime time for sex and gender considerations in traumatic brain injury? Perspectives of rehabilitation care professionals. Disabil Rehabil 2020 Jun 23:1-9. [doi: 10.1080/09638288.2020.1774670] [Medline: 32574090]
28. Lindsay S, Rezai M, Kolne K, Osten V. Outcomes of gender-sensitivity educational interventions for healthcare providers: a systematic review. Health Educ J 2019 Jun 27;78(8):958-976 [FREE Full text] [doi: 10.1177/0017896919859908]
29. Hamberg K. Gender bias in medicine. Womens Health (Lond) 2008 May;4(3):237-243 [FREE Full text] [doi: 10.2217/17455057.4.3.237] [Medline: 19072473]
30. Böckers A, Grab C, Waller C, Schulze U, Gerhardt-Szep S, Mayer B, et al. Implementation and sex-specific analysis of students' attitudes toward a longitudinal, gender-specific medical curriculum - a pilot study. Educ Health (Abingdon) 2017;30(1):50-59 [FREE Full text] [doi: 10.4103/efh.EfH_338_15] [Medline: 28707637]
31. World Health Organization. World Health Organization gender responsive assessment scale: criteria for assessing programs and policies. WHO Gender mainstreaming manual for health managers: A practical approach. URL: http://www.who.int/gender/mainstreaming/GMH_Participant_GenderAssessmentScale.pdf [accessed 2021-03-04]
32. Miers M. Developing an understanding of gender sensitive care: exploring concepts and knowledge. J Adv Nurs 2002 Oct;40(1):69-77. [doi: 10.1046/j.1365-2648.2002.02341.x] [Medline: 12230531]
33. Dworkin S, Fleming P, Colvin C. The promises and limitations of gender-transformative health programming with men: critical reflections from the field. Cult Health Sex 2015;17 Suppl 2:S128-S143 [FREE Full text] [doi: 10.1080/13691058.2015.1035751] [Medline: 25953008]
34. Abbott D, Carpenter J, Gibson B, Hastie J, Jepson M, Smith B. Disabled men with muscular dystrophy negotiate gender. Disabil Soc 2019 Mar 18;34(5):683-703 [FREE Full text] [doi: 10.1080/09687599.2019.1584093]
35. McMillan I. Gender-sensitive services vital for women with mental health problems. Learn Disabi Pract 2008 Jul 23;11(6):7. [doi: 10.7748/ldp.11.6.7.s13]
36. Milligan M, Neufeldt A. The myth of asexuality: a survey of social and empirical evidence. Sexual Disabi 2001;19(2) [FREE Full text] [doi: 10.1023/A:1010621705591]
37. Shah S. “Disabled People Are Sexual Citizens Too”: supporting sexual identity, well-being, and safety for disabled young people. Front Educ 2017 Sep 04;2:1-5 [FREE Full text] [doi: 10.3389/feduc.2017.00046]
38. Toft A, Franklin A, Langley E. Young disabled and LGBT+: negotiating identity. J LGBT Youth 2019 Jan 08;16(2):157-172 [FREE Full text] [doi: 10.1080/19361653.2018.1544532]
39. Parkhill AL, Mathews JL, Fearing S, Gainsburg J. A transgender health care panel discussion in a required diversity course. Am J Pharm Educ 2014 May 15;78(4):81 [FREE Full text] [doi: 10.5688/ajpe78481] [Medline: 24850943]
40. Sequeira GM, Chakraborti C, Panunti BA. Integrating Lesbian, Gay, Bisexual, and Transgender (LGBT) content into undergraduate medical school curricula: a qualitative study. Ochsner J 2012;12(4):379-382 [FREE Full text] [Medline: 23267268]
41. Porter KE, Krinsky L. Do LGBT aging trainings effectuate positive change in mainstream elder service providers? J Homosex 2014;61(1):197-216. [doi: 10.1080/00918369.2013.835618] [Medline: 24313259]
42. Dispenza F, Harper LS, Harrigan MA. Subjective health among LGBT persons living with disabilities: a qualitative content analysis. Rehabil Psychol 2016 Aug;61(3):251-259. [doi: 10.1037/rep0000086] [Medline: 26950832]
43. Eriksson SES, Safer JD. Evidence-based curricular content improves student knowledge and changes attitudes towards transgender medicine. Endocr Pract 2016 Jul;22(7):837-841. [doi: 10.4158/EP151141.OR] [Medline: 27042742]
44. Safer JD, Pearce EN. A simple curriculum content change increased medical student comfort with transgender medicine. Endocr Pract 2013;19(4):633-637. [doi: 10.4158/EP13014.OR] [Medline: 23425656]
45. Dielissen P, Verdonk P, Waard MW, Bottema B, Lagro-Janssen T. The effect of gender medicine education in GP training: a prospective cohort study. Perspect Med Educ 2014 Nov;3(5):343-356 [FREE Full text] [doi: 10.1007/s40037-014-0122-3] [Medline: 24895104]
46. Lindsay S, Rezai M, Kolne K, Osten V. Outcomes of gender-sensitivity educational interventions for healthcare providers: a systematic review. Health Educ J 2019 Jun 27;78(8):958-976 [FREE Full text] [doi: 10.1177/0017896919859908]
47. Issenberg SB, McGaghie WC, Petrusa ER, Gordon LD, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach 2005 Jan;27(1):10-28. [doi: 10.1080/01421590500046924] [Medline: 16147767]
48. McGaghie WC, Issenberg SB, Cohen ER, Barsuk JH, Wayne DB. Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. Acad Med 2011 Jun;86(6):706-711 [FREE Full text] [doi: 10.1097/ACM.0b013e318217e119] [Medline: 21512370]
49. Stanley C, Lindsay S, Parker K, Kawamura A, Samad Zubairi M. Value of collaboration with standardized patients and patient facilitators in enhancing reflection during the process of building a simulation. J Contin Educ Health Prof 2018;38(3):184-189. [doi: 10.1097/CEH.0000000000000198] [Medline: 29746322]

50. Khanduja PK, Bould MD, Naik VN, Hladkowicz E, Boet S. The role of simulation in continuing medical education for acute care physicians: a systematic review. Crit Care Med 2015 Jan;43(1):186-193. [doi: 10.1097/CCM.0000000000000672] [Medline: 25343571]
51. Orledge J, Phillips WJ, Murray WB, Lerant A. The use of simulation in healthcare: from systems issues, to team building, to task training, to education and high stakes examinations. Curr Opin Crit Care 2012 Aug;18(4):326-332. [doi: 10.1097/MCC.0b013e328353fb49] [Medline: 22614323]
52. Hales BM, Hawryluck L. An interactive educational workshop to improve end of life communication skills. J Contin Educ Health Prof 2008;28(4):241-249. [doi: 10.1002/chp.191] [Medline: 19058258]
53. Zubairi MS, Lindsay S, Parker K, Kawamura A. Building and participating in a simulation: exploring a continuing education intervention designed to foster reflective practice among experienced clinicians. J Contin Educ Health Prof 2016;36(2):127-132. [doi: 10.1097/CEH.0000000000000065] [Medline: 27262157]
54. Eppich W, Howard V, Vozenilek J, Curran I. Simulation-based team training in healthcare. Simul Healthc 2011 Aug;6 Suppl:S14-S19. [doi: 10.1097/SIH.0b013e318229f550] [Medline: 21817858]
55. Ryan DP, Marlow B. Build-a-case: a brand new continuing medical education technique that is peculiarly familiar. J Contin Educ Health Prof 2004;24(2):112-118. [doi: 10.1002/chp.1340240208] [Medline: 15279136]
56. Arnold JL, McKenzie FRD, Miller JL, Mancini ME. The many faces of patient-centered simulation: implications for researchers. Simul Healthc 2018 Jun;13(3S Suppl 1):51-55. [doi: 10.1097/SIH.0000000000000312] [Medline: 29771815]
57. Colorafi KJ, Evans B. Qualitative Descriptive Methods in Health Science Research. HERD 2016 Jul;9(4):16-25 [FREE Full text] [doi: 10.1177/1937586715614171] [Medline: 26791375]
58. Canada's health care providers: Provincial profiles, 2008-2017. Canadian Institutes of Health Information. Ottawa: Canadian Institute for Health Information; 2008. URL: www.cihi.ca › document › hcp-2017-data-tables-en-web [accessed 2021-03-11]
59. Nestel D, Bearman M. Simulated patient methodology: theory, evidence and practice. Hoboken, New Jersey, United States: Wiley-Blackwell; 2014:1-168.
60. Gill P, Stewart K, Treasure E, Chadwick B. Methods of data collection in qualitative research: interviews and focus groups. Br Dent J 2008 Mar 22;204(6):291-295. [doi: 10.1038/bdj.2008.192] [Medline: 18356873]
61. Lindsay S, Cagliostro E, Albarico M, Mortaji N, Karon L. A systematic review of the benefits of hiring people with disabilities. J Occup Rehabil 2018 Dec;28(4):634-655. [doi: 10.1007/s10926-018-9756-z] [Medline: 29392591]
62. Lindsay S, Cagliostro E, Leck J, Shen W, Stinson J. Disability disclosure and workplace accommodations among youth with disabilities. Disabil Rehabil 2019 Aug;41(16):1914-1924. [doi: 10.1080/09638288.2018.1451926] [Medline: 29558221]
63. Lindsay S, R Hartman L, Fellin M. A systematic review of mentorship programs to facilitate transition to post-secondary education and employment for youth and young adults with disabilities. Disabil Rehabil 2016 Jul;38(14):1329-1349. [doi: 10.3109/09638288.2015.1092174] [Medline: 26497325]
64. Lindsay S, Duncanson M, Niles-Campbell N, McDougall C, Diederichs S, Menna-Dack D. Applying an ecological framework to understand transition pathways to post-secondary education for youth with physical disabilities. Disabil Rehabil 2018 Feb;40(3):277-286. [doi: 10.1080/09638288.2016.1250171] [Medline: 27868448]
65. Lindsay S, Cancelliere S. A model for developing disability confidence. Disabil Rehabil 2018 Sep;40(18):2122-2130. [doi: 10.1080/09638288.2017.1326533] [Medline: 28503957]
66. Lindsay S, Cagliostro E, Leck J, Shen W, Stinson J. Employers' perspectives of including young people with disabilities in the workforce, disability disclosure and providing accommodations. J Vocat Rehabil 2019 Mar 21;50(2):141-156. [doi: 10.3233/jvr-180996]
67. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
68. Rolfe G. Validity, trustworthiness and rigour: quality and the idea of qualitative research. J Adv Nurs 2006 Feb;53(3):304-310. [doi: 10.1111/j.1365-2648.2006.03727.x] [Medline: 16441535]
69. Weyers S, Vervoorts A, Dragano N, Engels M. The Gender Lens: development of a learning aid to introduce gender medicine. GMS J Med Educ 2017;34(2):Doc17 [FREE Full text] [doi: 10.3205/zma001094] [Medline: 28584865]
70. Kumagai AK, Lypson ML. Beyond cultural competence: critical consciousness, social justice, and multicultural education. Acad Med 2009 Jun;84(6):782-787. [doi: 10.1097/ACM.0b013e3181a42398] [Medline: 19474560]
71. Bristol S, Kostelec T, MacDonald R. Improving emergency health care workers' knowledge, competency, and attitudes toward lesbian, gay, bisexual, and transgender patients through interdisciplinary cultural competency training. J Emerg Nurs 2018 Nov;44(6):632-639. [doi: 10.1016/j.jen.2018.03.013] [Medline: 29704979]
72. Dielissen PW, Bottema BJAM, Verdonk P, Lagro-Janssen TLM. Incorporating and evaluating an integrated gender-specific medicine curriculum: a survey study in Dutch GP training. BMC Med Educ 2009 Sep 08;9:58 [FREE Full text] [doi: 10.1186/1472-6920-9-58] [Medline: 19737396]
73. Lindsay S, Tétrault S, Desmaris C, King GA, Piérart G. The cultural brokerage work of occupational therapists in providing culturally sensitive care. Can J Occup Ther 2014 Apr;81(2):114-123. [doi: 10.1177/0008417413520441] [Medline: 25004587]
74. Yan M, Wong Y. Rethinking self-awareness in cultural competence: toward a dialogic self in cross-cultural social work. Families in Society 2018 May 03;86(2):181-188 [FREE Full text] [doi: 10.1606/1044-3894.2453]

75. Ahlgren C, Fjellman-Wiklund A, Hamberg K, Johansson EE, Stålnacke BM. The meanings given to gender in studies on multimodal rehabilitation for patients with chronic musculoskeletal pain - a literature review. Disabil Rehabil 2016 Nov;38(23):2255-2270. [doi: 10.3109/09638288.2015.1127435] [Medline: 26730507]
76. Mizock L, Lewis T. Trauma in transgender populations: risk, resilience, and clinical care. J Emotional Abuse 2008 Aug 26;8(3):335-354. [doi: 10.1080/10926790802262523]
77. King G, Desmarais C, Lindsay S, Piérart G, Tétreault S. The roles of effective communication and client engagement in delivering culturally sensitive care to immigrant parents of children with disabilities. Disabil Rehabil 2015;37(15):1372-1381. [doi: 10.3109/09638288.2014.972580] [Medline: 25323397]
78. Pelzang R. Time to learn: understanding patient-centred care. Br J Nurs 2010;19(14):912-917. [doi: 10.12968/bjon.2010.19.14.49050] [Medline: 20647984]
79. Dielissen P, Bottema B, Verdonk P, Lagro-Janssen T. Attention to gender in communication skills assessment instruments in medical education: a review. Med Educ 2011 Mar;45(3):239-248. [doi: 10.1111/j.1365-2923.2010.03876.x] [Medline: 21299599]
80. van der Meulen F, Fluit C, Albers M, Laan R, Lagro-Janssen A. Successfully sustaining sex and gender issues in undergraduate medical education: a case study. Adv Health Sci Educ Theory Pract 2017 Dec;22(5):1057-1070 [FREE Full text] [doi: 10.1007/s10459-016-9742-1] [Medline: 28050653]
81. Tannenbaum C, Clow B, Haworth-Brockman M, Voss P. Sex and gender considerations in Canadian clinical practice guidelines: a systematic review. CMAJ Open 2017;5(1):E66-E73 [FREE Full text] [doi: 10.9778/cmajo.20160051] [Medline: 28401121]
82. Kilgour J. Power of the patient voice in medical education. Clin Teach 2016 Dec;13(6):451-453. [doi: 10.1111/tct.12487] [Medline: 26749474]
83. Dear everybody: your impact matters- impact report 2017-2018. Toronto: Holland Bloorview Kids Rehabilitation Hospital; 2018. URL: http://yourimpactmatters.hollandbloorview.ca [accessed 2021-03-04]

Abbreviations

LGBTQ+: lesbian, gay, bisexual, transgender, queer, and other sexual identities

Edited by G Eysenbach; submitted 16.08.20; peer-reviewed by V Manchaiah, M Saud; comments to author 15.10.20; revised version received 15.10.20; accepted 17.01.21; published 15.03.21.

Please cite as:
Lindsay S, Kolne K, Barker DJ, Colantonio A, Stinson J, Moll S, Thomson N
Exploration of Gender-Sensitive Care in Vocational Rehabilitation Providers Working With Youth With Disabilities: Codevelopment of an Educational Simulation
JMIR Form Res 2021;5(3):e23568
URL: https://formative.jmir.org/2021/3/e23568
doi:10.2196/23568
PMID:33720023

Original Paper

Using a Personal Health Library–Enabled mHealth Recommender System for Self-Management of Diabetes Among Underserved Populations: Use Case for Knowledge Graphs and Linked Data

Nariman Ammar[1], MSc, PhD; James E Bailey[2], MD, MPH; Robert L Davis[1], MD, MPH; Arash Shaban-Nejad[1], MSc, MPH, PhD

[1]Oak Ridge National Laboratory Center for Biomedical Informatics, Department of Pediatrics, College of Medicine, The University of Tennessee Health Science Center, Memphis, TN, United States

[2]Center for Health System Improvement, College of Medicine, The University of Tennessee Health Science Center, Memphis, TN, United States

Corresponding Author:
Arash Shaban-Nejad, MSc, MPH, PhD
Oak Ridge National Laboratory Center for Biomedical Informatics, Department of Pediatrics
College of Medicine
The University of Tennessee Health Science Center
50 N Dunlap Street
Memphis, TN, 38103
United States
Phone: 1 901 287 5863
Email: ashabann@uthsc.edu

Abstract

Background: Traditionally, digital health data management has been based on electronic health record (EHR) systems and has been handled primarily by centralized health providers. New mechanisms are needed to give patients more control over their digital health data. Personal health libraries (PHLs) provide a single point of secure access to patients' digital health data and enable the integration of knowledge stored in their digital health profiles with other sources of global knowledge. PHLs can help empower caregivers and health care providers to make informed decisions about patients' health by understanding medical events in the context of their lives.

Objective: This paper reports the implementation of a mobile health digital intervention that incorporates both digital health data stored in patients' PHLs and other sources of contextual knowledge to deliver tailored recommendations for improving self-care behaviors in diabetic adults.

Methods: We conducted a thematic assessment of patient functional and nonfunctional requirements that are missing from current EHRs based on evidence from the literature. We used the results to identify the technologies needed to address those requirements. We describe the technological infrastructures used to construct, manage, and integrate the types of knowledge stored in the PHL. We leverage the Social Linked Data (Solid) platform to design a fully decentralized and privacy-aware platform that supports interoperability and care integration. We provided an initial prototype design of a PHL and drafted a use case scenario that involves four actors to demonstrate how the proposed prototype can be used to address user requirements, including the construction and management of the PHL and its utilization for developing a mobile app that queries the knowledge stored and integrated into the PHL in a private and fully decentralized manner to provide better recommendations.

Results: To showcase the main features of the mobile health app and the PHL, we mapped those features onto a framework comprising the user requirements identified in a use case scenario that features a preventive intervention from the diabetes self-management domain. Ongoing development of the app requires a formative evaluation study and a clinical trial to assess the impact of the digital intervention on patient-users. We provide synopses of both study protocols.

Conclusions: The proposed PHL helps patients and their caregivers take a central role in making decisions regarding their health and equips their health care providers with informatics tools that support the collection and interpretation of the collected knowledge. By exposing the PHL functionality as an open service, we foster the development of third-party applications or services and provide motivational technological support in several projects crossing different domains of interest.

(JMIR Form Res 2021;5(3):e24738) doi:10.2196/24738

KEYWORDS

personal health library; mobile health; personal health knowledge graph; patient-centered design; personalized health; recommender system; observations of daily living; Semantic Web; privacy

Introduction

Overview

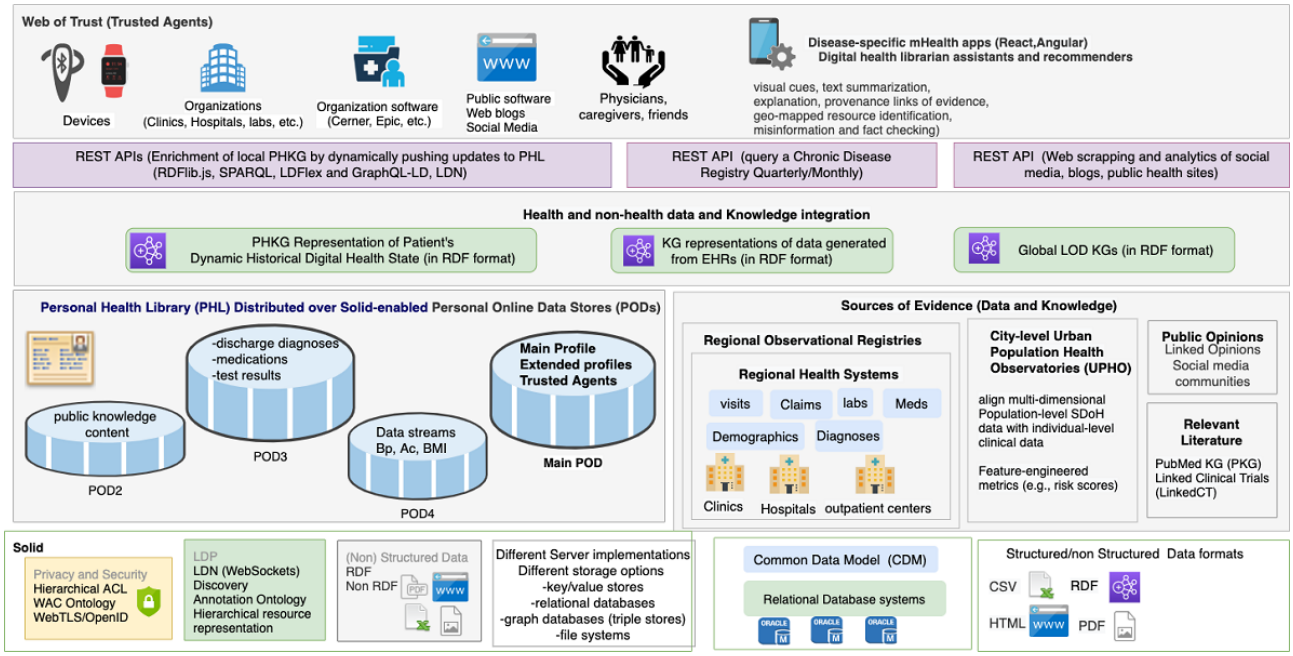
Historically, medicine has been largely health care provider–centered rather than patient-centered [1-3]. However, the new trend is moving toward incorporating patients' social and behavioral characteristics into electronic health records (EHRs) [4]. This combination of medical, social, behavioral, and lifestyle information about the patient is essential to facilitate understanding of medical events in the context of one's life and, conversely, to allow lifestyle choices to be considered jointly with that patient's medical context [5]. This data is generated over time by patients, their caregivers, and their providers and is potentially useful to all parties for decision making [6]. Patients are increasingly frustrated by the lack of EHR interoperability among fragmented systems and platforms dictated by providers or insurers, and they have expressed their needs to have an active role in managing their own health care data [7-12]. Improved interoperability and support for patient-provider communication have the potential to improve patient satisfaction and, evidence suggests, could even help detect and prevent medical errors [12].

In the field of personal digital health management, we often distinguish between a health *state* and a health *process* [12]. Health *state* is a digital representation of the patients' health at a given point in time, including their prescribed and over-the-counter medications, test results, exercise regimens, diets, appointments with providers, and clinical outcomes. Patients also receive other digital health information through other communication modalities and from diverse sources. These data can often come in different formats depending on their source, including EHRs, family histories, data streams from activity trackers, published research documents and data sets, websites, social media platforms, and videos. A health state changes over time as data is acquired through ongoing processes and events embedded within those processes. A health-related intervention is a *process*, which could either be therapeutic or preventive. For instance, an intervention for self-management of diabetes mellitus (commonly referred to simply as diabetes) focused on promoting change or changes in health behavior to improve clinical outcomes is an example of a preventive intervention. Early interventions are the best way to prevent the progression of a negative health outcome to its end stage. Digital interventions through mobile Health (mHealth) apps can serve as an effective tool in chronic disease self-management [13]. An integrated *personal health library* (PHL) can facilitate the building of mHealth apps by maintaining a historical digital representation of a patient's health state from diagnosis to monitoring and integrating local knowledge about patients with global web-based knowledge while providing patients with full ownership of their digital health states. Applications can process the knowledge stored in the library to generate intelligent personalized interventions that can help improve patients' health state.

Objective

We recently proposed both the conceptualization and initial implementation plan of a PHL [14,15]. The proposed PHL architecture (Figure 1) is the first of its kind to incorporate privacy, data ownership, integration, interoperability, portability, dynamic knowledge discovery, social determinants of health (SDoH) [16], and observations of daily living (ODLs) [5] into an end-to-end framework. In this paper, we provide a thematic assessment of patient requirements in a PHL, demonstrate how our proposed PHL meets these requirements, and describe an mHealth app that queries the PHL to deliver intelligent recommendations for improving self-care behaviors in diabetic adults.

Figure 1. A PHL that leverages the semantic technologies and decentralized privacy and security mechanisms of Social Linked Data (Solid) to enable true ownership, data integration, interoperability, portability, and dynamic knowledge discovery. The PHL enables building Hybrid mHealth Recommenders and Digital Librarians (HRDLs). ACL: access control list; API: application programming interface; Bp: blood pressure; DWPC: Diabetes Wellness and Prevention Coalition; ED: emergency department; LDN: Linked Data Notifications; LDP: Linked Data Platform; LOD: Linked Open Data; mHealth: mobile health; OWL: Web Ontology Language; PKG: personal knowledge graph; RDF: Resource Description Framework; REST: representational state transfer; SPARQL: SPARQL Protocol and RDF Query Language; WAC: Web Access Control.



Review of Relevant Literature

To the best of our knowledge, little data exist on the implementation of PHLs. Barr et al [17,18] proposed the Audio-PaHL project that utilizes text, audio, and image mining, natural language processing (NLP), and social network analysis to integrate audio-recordings of clinic visits into a library. They link medical terms that appear in the recordings to trustworthy patient resources, which can be retrieved, organized, edited, and shared by patients. Their project enables self-management in caregivers and older adults with multimorbidity. Several other researchers have utilized Semantic Web technologies and techniques to enhance EHRs or build research prototypes that can be integrated within clinical workflows. Table 1 describes some of the efforts in this area, including the rationale, the methods used, the health outcomes on which they focused, and whether they incorporate PHLs, SDoH, ODLs, and privacy. These works utilized a mix of Semantic Web and Machine Learning techniques. Most of these prior studies have predominantly focused on diabetes as the health outcome of interest. However, none of them incorporated privacy, SDoH, and ODLs in an integrated end-to-end framework.

Table 1. Comparison with existing methods.

Reference	Rationale	Method	Health outcome	Utilizes a PHL[a]	Privacy-aware	Incorporates ODLs[b]	Incorporates SDoH[c]
Audio-PaHL Barr et al [17,18]	Self-management	Audio text retrieval Natural Language Processing (NLP) Social Network Analysis	Multiple	Yes	No	No	No
PHD[d] Backonja et al [19,20]	EHR[e] Enhancements	Collection of structured data using standards and protocols	Multiple	No	Yes	Yes	No
Ralston et al [21]	Self-management	Web-based Interactive EHR	Diabetes	No	No	Yes	No
PerKApp [22]	Health promotion	Semantic inference and knowledge representation	Diabetes	No	No	No	No
PHKG[f] [23,24]	Health promotion	Semantic inference and knowledge representation	Diabetes	No	No	No	No
kHealth Sheth et al [25]	Early warning decision support system	Declarative knowledge-based reasoning and machine learning	Asthma	No	No	No	No
Seneviratne et al [26]	Disease characterization	Semantic inference and knowledge representation	Breast cancer	No	No	No	No
Chari et al [27]	Treatment recommendations	Knowledge integration	Diabetes	No	No	No	No

[a]PHL: personal health library.

[b]ODLs: observations of daily living.

[c]SDoH: social determinants of health.

[d]PHD: Project HealthDesign.

[e]EHR: electronic health record.

[f]PHKG: personal health knowledge graph.

Efforts to Enhance EHRs

There have been efforts to move EHRs from being mere data stores to being platforms that provide actionable information to patients, their caregivers, and health care providers. The Project HealthDesign program introduced ODLs as a major component of such enhanced EHR platforms [5,19,20]. Ralston et al describe a web-based disease self-management platform based on an interactive EHR that incorporates a diabetes module [21].

Applications for Promoting Healthy Behavior

Some standalone applications utilize semantic inference and knowledge representation capabilities for promoting healthy behavior. For example, the PerKApp [22] leverages augmented domain knowledge in ontologies as well as reasoning rules to implement a persuasive platform targeted toward health promotion in the workplace that monitors employees' dietary and physical activity habits and sends interactive messages to persuade them to change their behaviors. A few studies leverage the notion of a personal health knowledge graph (PHKG) for patients that enables them to monitor and self-manage their chronic diseases while incorporating their ODLs [23,24]. They utilize knowledge representation to define ontologies and perform intelligence tasks on top of Resource Description Framework (RDF) graphs.

Knowledge Integration for Disease Characterization

Some researchers have combined statistical and machine learning approaches with knowledge representation approaches. For example, Sheth et al proposed a knowledge-enabled approach to health data analytics that combines declarative knowledge-based models with probabilistic machine learning models [25]. They developed the kHealth project and deployed it as an mHealth app for decision support in patients with asthma. Seneviratne et al developed a semantic end-to-end prototype for cancer characterization [26]. Their tool utilizes a cancer staging ontology to aid physicians to quickly stage a new patient and identify risks, treatment options, and monitoring plans. Physicians can also restage existing patients or patient populations, allowing them to find patients whose stage has changed within a given patient cohort. They applied knowledge integration by converting a patient's EHR to an RDF knowledge graph and perform deductive reasoning to infer the stage of a tumor.

Applications for Treatment Recommendations Based on Cohort Characteristics

Chari et al utilized semantic technologies and knowledge graphs to implement a tool that allows users to quickly derive clinically relevant inferences about study populations [27]. They developed a prototype workflow that utilizes an ontology to

expose population descriptions in research studies through visual aids. Their goal is to enable physicians to better understand the applicability and generalizability of treatment recommendations within clinical practice guidelines.

Innovative Technologies Used in The PHL Implementation

The PHL leverages several innovative technologies that were inspired by the requirements that we identified in the literature. We highlight some of the relevant efforts undertaken over time, including protocol specifications, vocabularies, standards, and technologies to support those requirements. We also introduce the terminology used throughout the paper, and in the results section we include several code snippets to illustrate these technologies.

Linked Open Data

An abundance of scientific evidence and open data sets are available on the Web in different formats. To enhance *discoverability* and *linkability* by enabling both humans and machines to access such data, Berners-Lee proposed the *Linked Open Data* (LOD) project, which aimed to make open data available on the Web as linked data (eg, Bio2RDF) [28]. LOD is a way of connecting resources located throughout the Web by establishing a URI for each piece of data and explicitly stating how they are related to one another. LOD has led the research community to transform life sciences data sets into semantic format and make them available on the Web. LinkedCT, for example, is a ClinicalTrials.gov LOD data set that defines concepts related to diseases and interventions. The *Linked Open Research* (LOR) project leverages the LOD principles by providing an infrastructure to semantically represent research artifacts and to connect resources and the activity around them using notifications and visualizations to facilitate scientometric studies and decision making. By leveraging LOD, the proposed PHL simplifies the process of systematically and dynamically adding typed data relating to unique health and nonhealth concepts (eg, ODLs and SDoH) to the patients' digital health profiles, thereby *reducing the effort* needed for both patients and health care professionals to collect data and understand it, respectively. In addition, by leveraging the LOR initiative, the PHL enables physicians and patients to conduct scientific activities by combining global scientific knowledge discovered on the Web with local knowledge stored in their own library.

Web Annotation Specification

Annotations are typically used to convey information about a resource or associations between resources. The *Web Annotation* specification [29] describes a structured data model to enable annotations to be shared and reused across different platforms. It provides a specific format for creating annotations and consuming them based on a conceptual model and a set vocabulary of terms that accommodate a certain use case. One research challenge is to explore the potentially large number of annotations to discover patterns that capture semantic knowledge not only about individual nodes and their connections but also about groups of related nodes. For example, annotating clinical trials to look for patterns is an active research area [30], and there are open data sets that can be used for that purpose. Therefore, there is a need for more automatic tools to support scientists in pattern discovery, including link predictions or discovering complex patterns of annotation (eg across multiple disease conditions and drug interventions). By leveraging LOD, the PHL enables the mining of data sets that are semantically annotated with controlled vocabulary terms and concepts (eg, risk factors) and properties (eg, risk factors associated with a disease) encoded in ontologies. Through annotations, the PHL enables a user (eg, a physician) to convey information about a resource or associations between resources (eg, a tag on a lab test or image or a comment on a blog post about a research article). Annotations also help them capture scientific knowledge and use it as a basis for conducting focused literature reviews or planning new clinical trials.

Representational State Transfer

Representational state transfer (REST) is an architectural design pattern [31] for client-server communication that is centered around the following principles. First, each piece of content on the Web (both data and functionality) is considered a resource with a unique URI that provides a global addressing space for resource and service discovery. Second, resources are considered documents acted upon by Web application programming interface (API) operations (GET, POST, UPDATE, and DELETE) to manipulate those resources using HTTP as a communication protocol. Third, resources are decoupled from their representations, so their content can be accessed in a variety of formats (HTML, XML, JSON, plaintext, JPEG, PDF, etc). Finally, Web content should be designed as a network of resources that link to each other following the Hypermedia as the Engine of Application State principle. This principle enables discoverability using hypermedia controls that indicate to the resource requester a set of actions that are available to them on that resource as well as the URLs on which those actions can be performed. APIs using RESTful architecture have been widely adopted for software-to-software communication across heterogeneous distributed environments. RDF, the model driving the Linked Data Platform (LDP), follows the REST principles of identifying resources by URIs, which facilitates managing resources via HTTP operations on their URIs. It also enables hypermedia-based discovery [32]. We follow this approach for adding, deleting, and updating resources in the PHL. However, our approach solves the lack of substitutability with non-native RESTful APIs in current EHR implementations, which often hinders systems programmed for a specific API task (eg, adding a resource to Cerner at Hospital 1) from performing that same task with another incompatible API (eg, adding that same resource to Cerner at Hospital 2). Solid adopts a pattern-based approach to API design that enables applications to be compatible with APIs beyond those for which they were explicitly programmed [33].

Web-Scale Semantic Querying

The LOD stack (RDF, URIs, and SPARQL [SPARQL Protocol and RDF Query Language]) makes any piece of data accessible and queryable on the Web. SPARQL can be used to execute federated queries across many endpoints on the Web. The most straightforward technique for accessing LOD data is following

a URL of an RDF document through a process called *dereferencing*, which involves using the HTTP protocol to retrieve a representation of a resource identified by a URL. SPARQL endpoints offer interfaces that permit selection of data in a granular way with the ability to perform complex data retrieval from multiple *personal online data stores* (PODs) via link-following SPARQL. Solid exposes data in a document-oriented way (RDF) and provides a uniform interface to query this data.

Federated Linked Data Querying

Federated queries are used to achieve web-scale integration and interoperability. Executing federated Linked Data queries on the Web requires accessing multiple data sources, which involves the discovery of data sources and determining relevant ones. Researchers have proposed discovery approaches for Linked Data interfaces based on hypermedia links and controls and applied them to federated query execution with Triple Pattern Fragments [34]. SPARQL endpoints are expensive for the server and not always available for all data sets. Downloadable dumps are expensive for clients and do not allow live querying on the web. The Linked Data Fragments framework enables client-side SPARQL querying of live Linked Data on the Web and federated querying through a triple-pattern interface, providing a much faster, less expensive solution [34].

Methods

We conducted a thematic assessment of patients' requirements and used the identified requirements to develop a use case scenario that motivates the need for a PHL. We explain the technological infrastructures used to build the PHL platform, including the Solid platform and knowledge graphs.

Thematic Assessment of Patient Requirements

The patient requirements for a PHL that we identified in the literature [7-12] fall into three broad themes: (1) construction and management of the library, (2) dynamic discovery and integration of new knowledge related to data and types stored in the library, and (3) the ability to leverage the knowledge stored in the library through digital interventions (Table 2). We assessed the technological innovations required to meet those requirements. To address R1.2 and R3.2, we need an infrastructure that supports privacy by design. To enable patients to effectively share data and knowledge, we need a platform that supports sharing not only with individuals (R5, R11) but also with organizations (R6) and devices (R10, R13.2). To enable patients to selectively define and store *types* of knowledge or data (R1.3, R3.1), we need to leverage Semantic Web technologies. Finally, to incorporate dynamic discovery and knowledge enrichment, we need to store patient's data using a KG (R4.1, R4.2) in a machine-readable format. Besides these functional requirements, the platform should support nonfunctional requirements, including security, integration, and interoperability. Our innovative technologies meet all of these requirements.

Table 2. Some requirements for a PHL (from a patient perspective, per the literature).

Requirement	Description
I. Construction and management of digital health state	
R1.1 Integration R1.2 Security and privacy R1.3 Semantic technologies: ontologies	*Construct* a PHL[a] by *bringing* a patient's *data together* in a *trustworthy*, *usable*, and *useful* library by gathering different *types* of *knowledge* into a *single resource*
R2 Management: a. RESTful[b] resources b. CRUD[c] operations	*Manage* the PHL by *creating*, *reading*, *updating*, or *deleting* resources
R3.1 Semantic technologies: ontologies **R3.2 Privacy:** a. What: resource b. Who: agent	*Decide* what *types* of data should be kept and who *has access* to that data
II. Dynamic knowledge discovery and integration	
R4.1 Dynamic knowledge discovery R4.2 Knowledge enrichment	*Seek* health data from *constantly changing public sources*, enriched with *new streams* and *types* of *data*
R5 Knowledge sharing with individuals	Decide what *types* of data are important to *collect*, *manage*, and *share*
R6 Knowledge sharing with organizations	*Share* data with *citizen science* and *research initiatives*
III. Processing digital health state (digital interventions/mHealth[d] apps)	
Interaction-based usage	
R7.1 Searching **R7.2 Semantic technologies:** a. Unique resource representation b. Vocabulary mapping	*Search* through the PHL using *intelligent mapping* for *vocabulary* used to describe *resources* in the patients' *profiles*
R8.1 Semantic technologies: annotations	*Annotate* results from patient's participation in clinical trials to *look for patterns*
R8.2 AI[e]: pattern detection	
Notification-based usage	
R9 Dynamic knowledge discovery	Receive *alerts* about *new* data *related* to *topics* covered in their PHL
R10 Wearable device agents (ODL[f] data)	Play an active role in staying healthy by *monitoring* their *progress*
R11 Knowledge sharing with individuals	*Stay current* with *treatment options* and *clinical trials* for a *family member* with a debilitating condition
R12 Semantic technologies: a. Text summarization b. Knowledge mapping	*Find* and *use* information including *text summarization*, *knowledge mapping*, etc
R13.1 Intelligent mHealth apps: a. Digital assistants b. Recommender systems	Access *digital assistance via personalized alerts* and suggestions, text summarization, *literacy* aids, *translations*, etc
R13.2 Software agents	

[a]PHL: personal health library.

[b]REST: representational state transfer.

[c]CRUD: creating, reading, updating, or deleting.

[d]mHealth: mobile health.

[e]AI: artificial intelligence.

[f]ODL: observations of daily living.

Scenario

We present a scenario that involves Bob as the main actor and Alice and Mary as supporting actors to demonstrate how a diabetic patient can benefit from the PHL and use the mHealth app for diabetes self-management (Textbox 1).

Textbox 1. Use case scenario of self-management for a diabetic patient.

Bob is an African American adult with diabetes equipped with a smartwatch and smartphone that collect physiological data (eg, step counts) in real time. The social app on his smartphone queries his decentralized personal health library (PHL) to deliver tailored push notifications to support behavior change related to chronic disease self-care. Depending on sensor readings and other information in his PHL, the app provides personalized and tailored recommendations for healthy eating, physical activity, medication taking, and/or visiting health care providers. The health recommendations also take into account different characteristics of the nearby points-of-interest. For example, as Bob has another comorbidity (asthma), running activities must be avoided.

Alice is a patient who is under treatment for cancer and is also part of Bob's social network. She would like to report acute health conditions and side effects of using medications by sharing a notepad with her physician.

Mary is a physician who follows up with both Bob and Alice. Through the PHL, she would like to interact with her patients and to be able to access their test results easily. She would also like to conduct research about their health conditions using the content generated through their PHLs and her own by tracking publications and scientific observations from trusted public knowledge sources.

A clinic or lab would like to access the PHLs of Alice and Bob to share the results of their lab tests and to follow up on the test and visit history. It also wants to share those tests with their physician, Mary.

In the following, we describe two of the main technological infrastructures used to construct, manage, and integrate the types of knowledge stored in the PHL. Namely, we describe the Solid platform and the role of personal knowledge graphs in designing the PHL and the mHealth app.

Social Linked Data (Solid)

To extend the current functionality of the Web (World Wide Web Consortium [W3C] standards and protocols) by applying LOD principles, Berners-Lee proposed the Solid project [35,36]. Solid is shaping the future vision of the decentralized Web (Web 3.0) by enhancing the technologies used to build Web 2.0 while bringing back the privacy and freedom-oriented values of Web 1.0. Solid uses the *WebID specification* [37] to implement a global identity management architecture based on the notion of decentralized identity providers. Coupled with the WebID-TLS decentralized authentication protocol, a WebID enables a global web-scale single sign-on. Moreover, Solid follows a unique architecture for building applications by separating users' data from the applications that use that data, which guarantees not only privacy and true ownership but also flexibility. Users can store their data among several Solid-compatible PODs hosted on Solid-enabled Web servers (Figure 1) that users can either install on their own machines or obtain from a listed provider and selectively authenticate applications to access and process specific resources within those servers. When users register for an identity, their WebID profile document and associated cryptographic key is stored on their main POD server (Figure 1). By leveraging Solid, the underlying storage for patients' digital data in the proposed PHL can be implemented in several ways, for example, file systems, key-value stores, and relational or graph database systems (Figure 1). In addition, accessing physical data and the metadata will be performed through an allocated semantic layer so that changing or reorganizing the data sources does not cause interruption in the application. Solid is a stack of protocols and standards, so any mHealth app that utilizes the PHL will enable users to maintain control of their data so long as the app is built in a Solid-compatible way.

Personal Knowledge Graphs

A personal knowledge graph (PKG) [38] provides a new research frontier toward building intelligent applications. Assembling data from distributed sources is often challenging, but by leveraging the LOD stack (RDF, URIs, and SPARQL), our PHL platform can achieve such assembly by building an RDF representation of a PKG for each patient that maintains a historical representation of that patient's digital health state. Applications can query those graphs to render different aspects through visual aids and aggregate data by accessing the PODs of the target patient, PODs belonging to other users, and external Web resources (Figure 1).

Results

Summary

We present the results of assessing the requirements in the previous scenario, the prototype of the PHL and mHealth app, and our evaluation plan. We follow the three main themes identified in Table 2 to demonstrate how the different innovative technologies incorporated into the PHL meet these requirements through actions performed by the actors in our scenario (summarized in Textbox 1). In Textbox 2, we distill the most important requirements that the three actors and the involved organization or organizations will need in the proposed PHL and show how we leverage Solid and PKGs, among other technologies, to achieve those requirements.

Textbox 2. Requirements identified in the use case scenario from Textbox 1.

Functional requirements:

Agents

- Device agents (smartwatch, smartphone)
- Person agents (Mary/physician, Alice/another patient)
- Organization agents (clinic)
- Software agents (blogging app, calendar app)

Integrate different types of knowledge from different sources

- Registries: comorbidity (asthma)
- Global knowledge: blogs (HTML), articles (PDF)
- SDoH: neighborhood with low walkability score
- ODL: healthy eating, physical activity, medication taking, and visiting health care providers

Usage patterns

- Annotation of research articles and clinical trials
- Sharing knowledge and resources

Nonfunctional requirements:

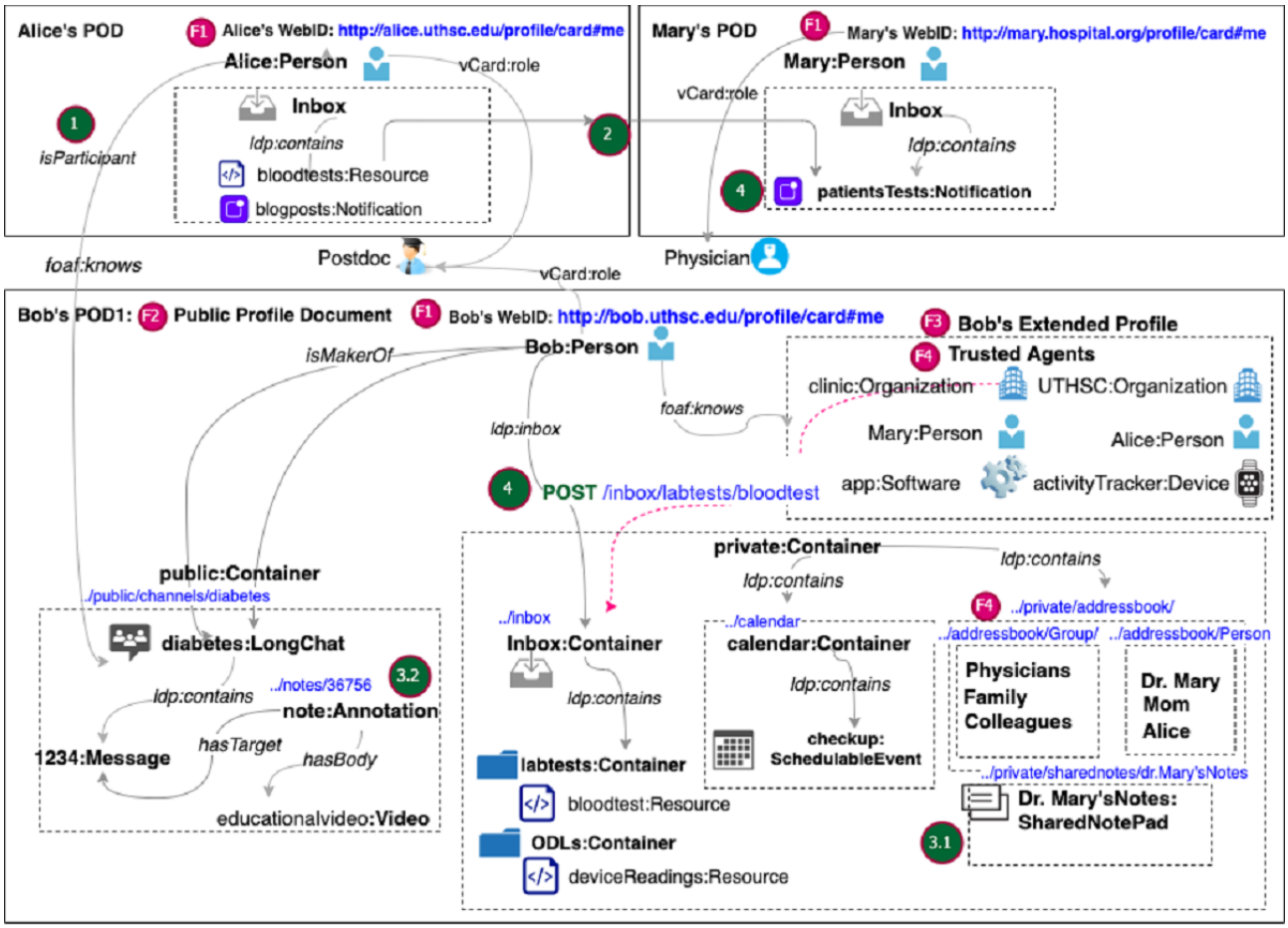
Security, privacy, interoperability

- Clinic and PHL software integration (posting test results)
- PHL and EHR software integration (patient's self-reported outcome)

Construction and Management of the PHL

We describe how the three actors in our scenario (Textboxes 1 and 2) can use our platform to construct their PHLs (Figure 2). This includes decentralized identity through WebIDs, main and extended profile documents (Figure 2, F1-F3), trusted agents (Figure 2, F4), and resource management (Figure 2, F5).

Figure 2. Main PHL features that meet some of the patients' requirements (Table 2) demonstrated through the PODs of Alice, Bob, and Mary. Bob's POD contains his main profile document in RDF-based KG representation. Social interactions within the PHL ecosystem include: (1) Alice and Mary can subscribe to Bob's channel using their WebIDs. (2) Alice can share her lab tests by pushing them to Mary's inbox. (3.1) Alice can share a notepad with Mary to discuss her lab results. (3.2) Alice can add annotations or comments to message content in Bob's diabetes channel. (4) Software from a clinic or other provider can share test results with Alice by performing a POST Web API operation on the unique URI of her inbox. POD: personal online data store.



Decentralized Identity Through WebIDs

First, the three actors generate unique WebIDs (Figure 2, F1) to securely log in to their main RDF-based PHL *profile document* (Figure 2, F2). Once they set up their main PHL profiles, they can build a *Web of Trust* using the FOAF (Friend of a Friend) vocabulary (eg, foaf:knows) by linking their main profiles to one or more extended profiles (Figure 2, F3). In their extended profiles, our actors can keep lists of vCard URIs of trusted agents and selectively grant them access to content in their PHL. For example, Bob can create an extended document in which he stores his friends list (Figures 3 and 4). Once they set up their profiles, users can manage content within those documents. In the following section, we explain how Bob manages content under his PHL and how he adds trusted agents.

Figure 3. Bob's main PHL profile document with a reference to his friends' extended profiles.

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
<> PersonalProfileDocument ;
<#me> a: Person ;
  :name "Bob";
  rdfs:seeAlso https://bob.uthsc.edu/friends
```

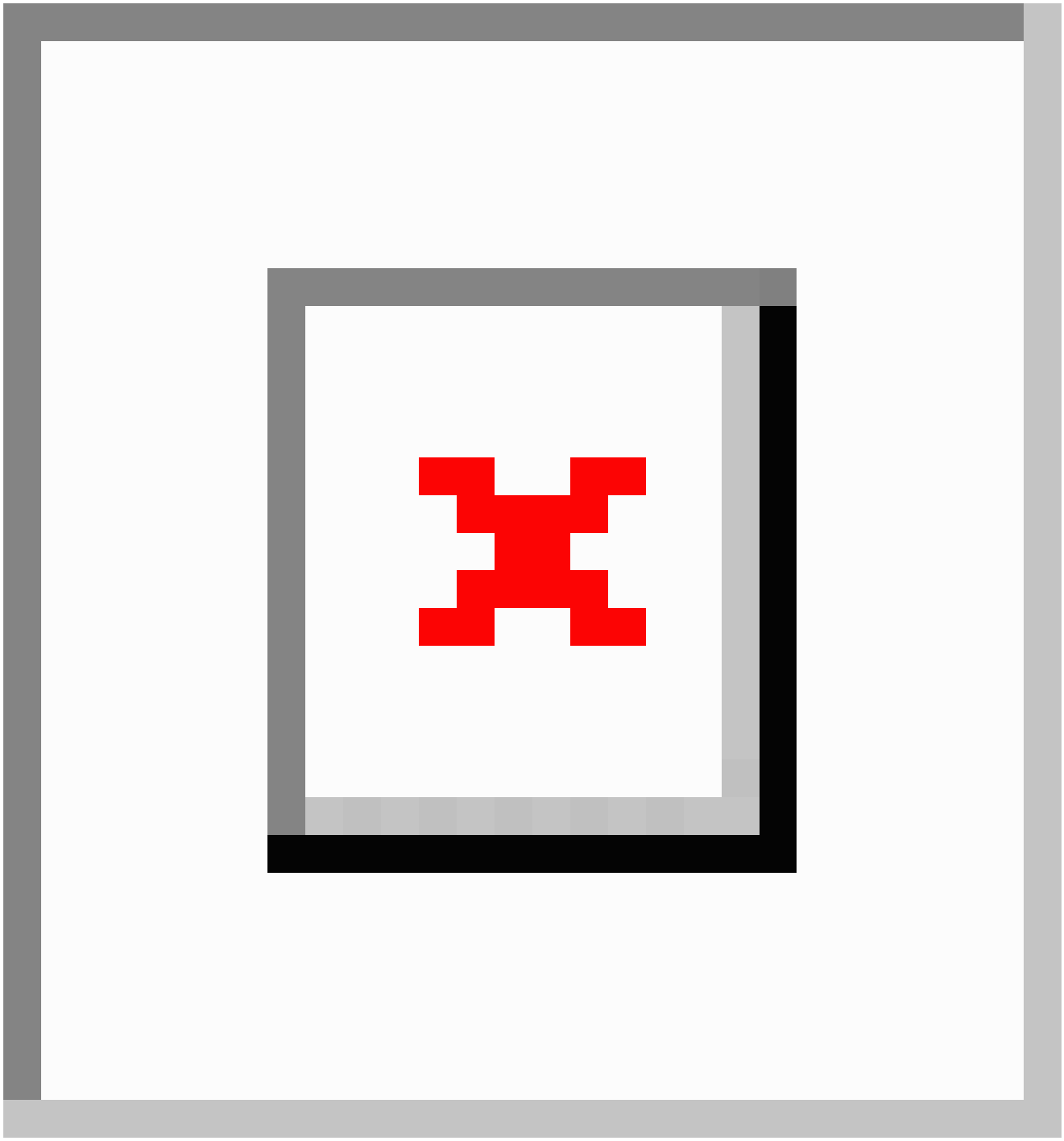
Figure 4. Bob's extended profile document (friends) that identifies Alice and Mary as trusted agents by establishing foaf:knows relations with their WebIDs.

```
@prefix foaf: <http://xmlns.com/foaf/0.1/>.
<> a foaf:PersonalProfileDocument ;
  foaf:maker <https://bob.uthsc.edu/profile#me> ;
..
  foaf:knows <https://uthsc.edu/p/Alice#MSc> ;
  foaf:knows <https://hospital.org/people/Mary/card#me> .
```

Hierarchical Resource Representation

Whether it is a person, an inbox, a file, an image, a notification, or a relationship, content within the PHL is represented as a collection of Web resources. Patients can organize resources in their PHL as a hierarchy of nested *containers*. For example, an event is nested inside a calendar and a message is nested within a chat channel. Both containers and resources conform to the LDP BasicContainer specification. The different actors in our scenario can start with default containers provided by the PHL. For example, the inbox container gets created in the PHL as a default container preconfigured with live notifications. The patient's inbox is also discoverable through the ldp:inbox property specified in the Linked Data Notification (LDN) specification [39]. Beyond default containers, patients can define their own resources or containers. For example, Bob can add the calendar and diabetes folders under his public or private folders (Figure 5). He can also set up a chat channel about diabetes as a resource of type LongChat and nest that within the Diabetes folder (Figure 2, F5).

Figure 5. Hierarchy of containers under Bob's PHL.



Flexible Data Representations

The PHL supports reading and writing resources in different formats: (1) structured Linked Data resources (eg, RDF, HTML+RDFa (RDF in Attributes), etc), (2) binary data (eg, images, videos, webpages), and (3) non–linked data structured text. While the PHL enables building applications with nonlinked resources, using RDF-based linked data provides extra benefits in terms of interoperability with the rest of the ecosystem.

Trusted Agents

Agents can be persons, organizations, devices, or software. For example, Bob can add the clinic as an organization agent, Alice and Mary as person agents, his mobile phone as a device agent, and the diabetes self-management mobile app as a software agent (Figure 2, F4). Access can be granted either to individual agents or agent groups. For example, Bob can define two work-groups (Physicians and Caregivers) in an extended document (work-groups) (Figure 6). In that document, he can list Mary and Alice as members using their WebIDs. He can then grant each of these agent groups fine-grained access permissions to his shared-notepad resource. Bob can reference a group as a resource under the work-groups document (Figure 7).

In addition, since he is interested in tracking and gathering data about diabetes from a blogging app, he can add the app under the trusted apps section of his extended profile document (Figure 8).

Figure 6. Bob's work-groups document that defines Physicians and Caregivers as groups.

```
@prefix  vcard:  <http://www.w3.org/2006/vcard/ns#>.
<#Physicians>
    a                 vcard:Group;
    vcard:hasUID      <urn:uuid:1234:ABGroup>;
    # Physicians group members:
    vcard:hasMember   <https://Hospital.org/profile/card#me>;
<#Colleagues>
    a                 vcard:Group;
    vcard:hasUID      <urn:uuid:5678:ABGroup>;
    # Caregivers group members:
    vcard:hasMember   <https://alice.uthsc.edu/profile/card#me>.
```

Figure 7. Individual and Group authorizations to Bob's notepad.

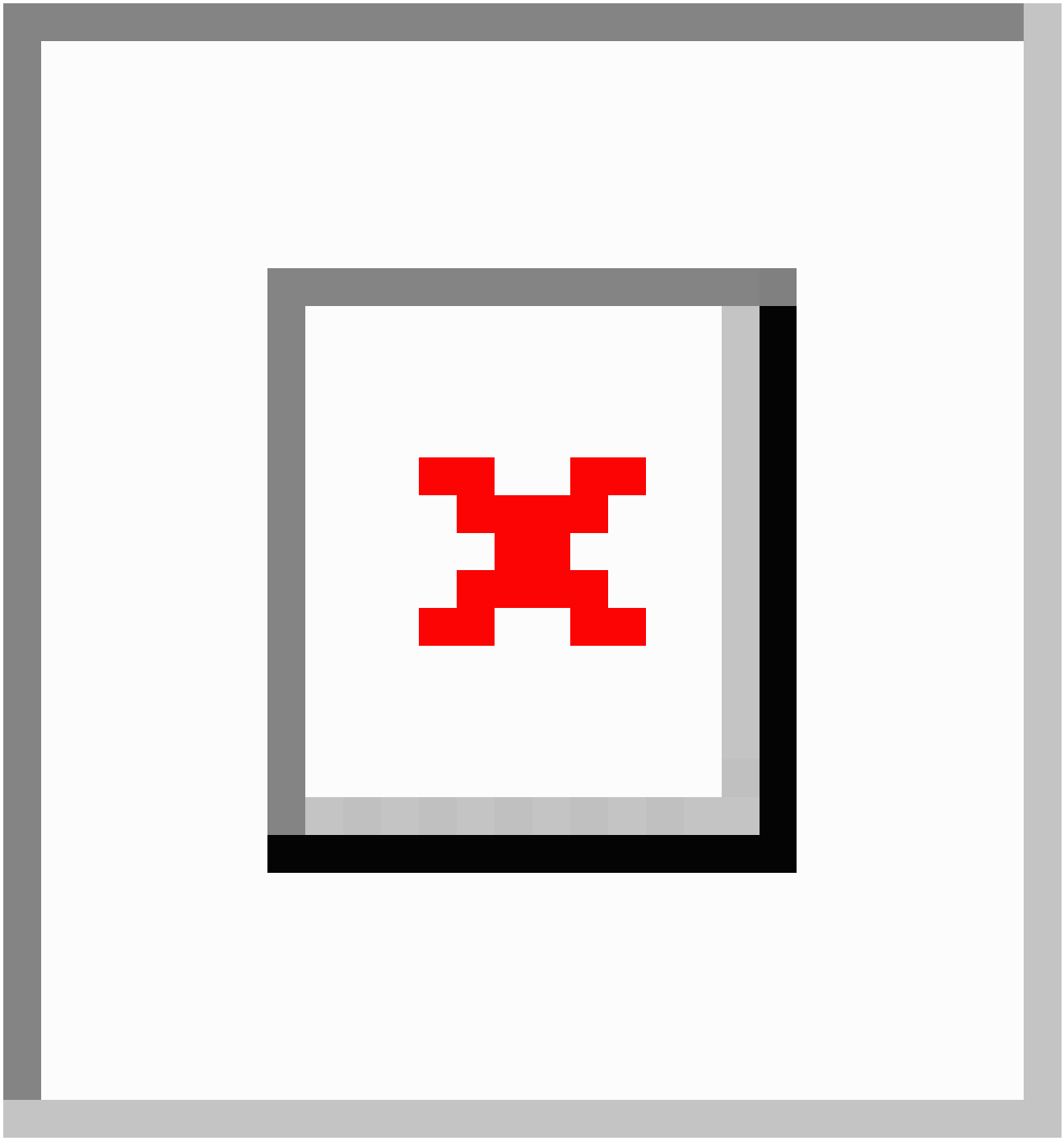


Figure 8. Bob's trusted apps.

```
<#me> acl:trustedApp [ acl:origin  <https://calendar.app.com>;
                       acl:mode    acl:Read,
                                   acl:Append].
 <#me> acl:trustedApp [ acl:origin  <https://blogging.app.com>;
                       acl:mode    acl:Read,
                                   acl:Write].
<#me> acl:trustedApp [ acl:origin  <https://diabetesSelfManagement.app.com>;
                       acl:mode    acl:Read,
                                   acl:Write].
```

Access Control Lists

The PHL uses the W3C Web Access Control (WAC) ontology to describe Read, Write, Control, and Append access control modes (eg, acl:mode acl:Read, Figure 9) at the level of a container or resource. Each resource can have an associated access control list (ACL) resource. If a container or resource does not have an ACL, it inherits the authorization of its parent container. For example, the default ACL on the inbox container is append-only by the public. Bob can associate the lab_test resource within his inbox with a corresponding ACL resource (lab_test.acl). Then, within that ACL resource, he can specify trusted agents and their corresponding access modes. For example, he can limit access to his friends' list extended profile document (Figure 4) by defining an ACL rule using the WAC ontology (Figure 9).

Figure 9. An ACL rule granting Read permission to Alice and Mary on Bob's "Friends" document.

```
@prefix acl: <http://www.w3.org/ns/auth/acl#> .
<#FriendsOnly>;
   acl:accessTo <https://bob.uthsc.edu/friends> ;
   acl:agent <http://uthsc.edu/people/Alice#Msc>,
             <http://hospital.org/people/Mary/card#me> ;
   acl:mode acl:Read .
```

Dynamic Knowledge Discovery and Integration

In this section, we describe how dynamic discovery and integration can be achieved through linkability, the ability to interact with the PHL content through annotations, rich embedding, and social interactions between the different actors in our scenario.

Linkability

Resources generated by each of the three actors get stored in their PHLs, with the possibility of linking resources in their PHLs to resources in other users' PHLs (Figure 2, F7). For example, if Alice comments on a message stored under the diabetes channel resource in Bob's PHL, her message will be stored under her PHL but links to Bob's message using the hasTarget Link type defined in the Web Annotation Ontology (WAO) (Figure 10).

Figure 10. Alice's comment is linked to Bob's message using the hasTarget link type of the Web Annotation Ontology.

```
@prefix wao: <http://www.w3.org/ns/oa#>.
 <https://alice.uthsc.edu/comments/36756>
   wao:hasTarget
      <https://bob.uthsc.edu/messages/diabetes/1234>
```

Annotations, Rich Embedding, and Social Interaction

The PHL can be used as a decentralized authoring, annotation, and social interaction framework that can be accessed from several platforms (eg, Web browsers and mobile apps). By implementing the Web Annotation specification [29], the PHL enables physicians to annotate content within resources stored under their PHLs. As a physician-researcher, Mary can benefit from the content generated through her PHL about diabetes. To this end, she can announce her inbox so she can receive LDN notifications of scholarly activities related to diabetes (eg, published articles, annotations in peer reviews [29], scientific observations). She can add identifiers to important concepts within her received content (eg, article) and add descriptive markup to those identified concepts. Her PHL uses the identifiers to automatically generate URIs for every article section to make it easy for others to refer to them or link them to external knowledge resources. Using Mary's descriptions, the PHL automatically generates RDFa markup documents, and by implementing the protocol, it enables her to expose those generated documents as Linked Data for other researchers to consume.

The PHL also supports rich embedding, whereby a researcher or a physician can embed raw data within a document in several formats and add provenance links (eg, nanopublications). For example, having Mary as a physician in the chat channel can add credibility to the discussion and enrich the knowledge exchanged. She can highlight certain statements exchanged in a message and correct a misconception. She can also comment on shared knowledge in a particular message and probably refer patients to more trusted sources of knowledge (eg, Centers for Disease Control and Prevention, World Health Organization). Similarly, when Bob receives information about treatment options from Mary as his physician, he can integrate it with information obtained from the blogging app. For example, he can highlight concepts in a blog post relating to diabetes and related interventions and link them to the same concepts in Mary's shared notepad or concepts shared by other users in his diabetes channel.

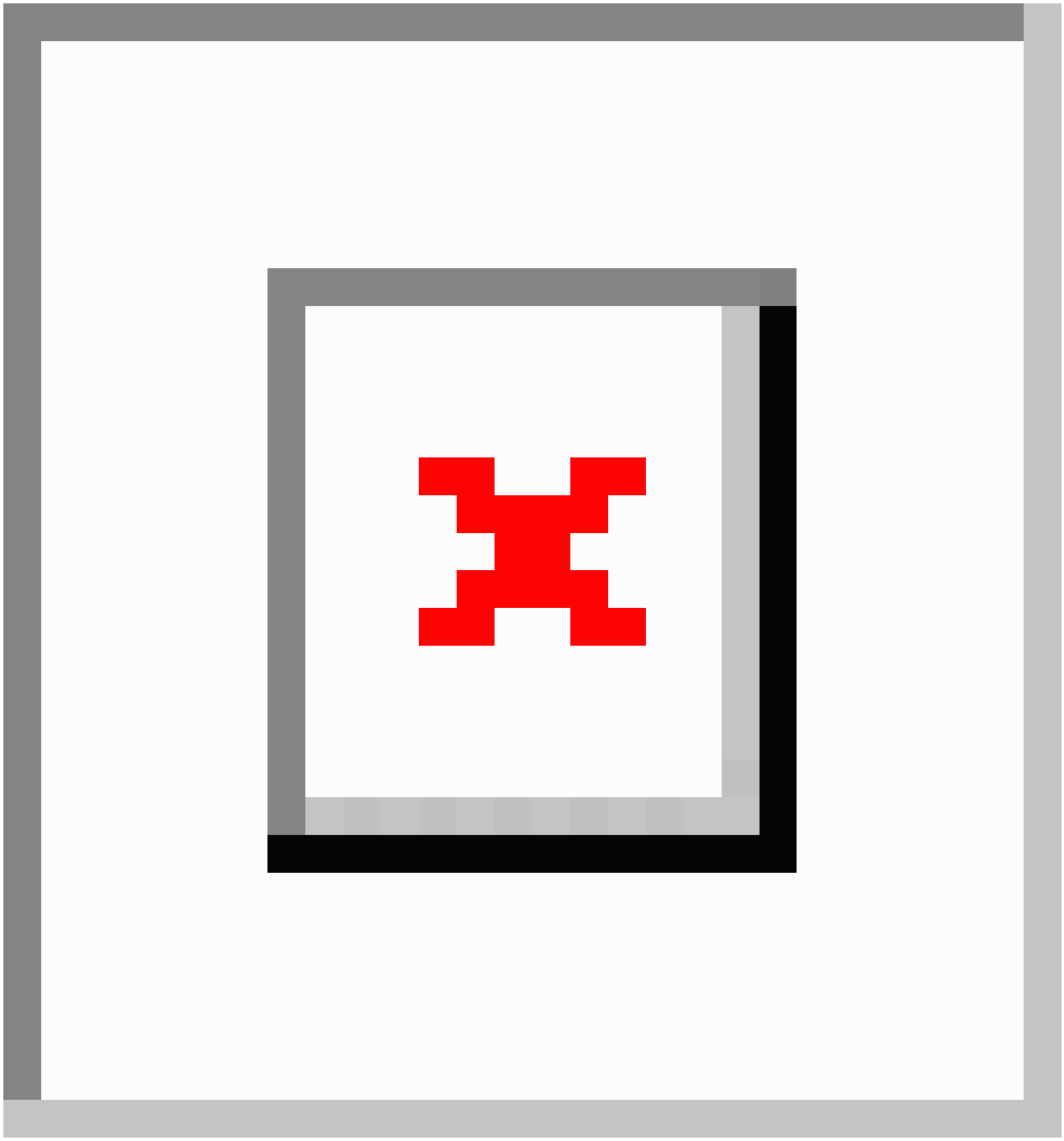
Mary, Bob, and Alice can perform social interactions within the PHL ecosystem in different ways (Figure 2). For example, (1) Alice and Mary can subscribe to Bob's channel using their WebIDs; (2) Alice can share the results of her lab tests by pushing them to Mary's inbox; (3.1) Alice can share a notepad with Mary to discuss digital health knowledge that she obtained from other sources, and Mary can interact with the content by adding annotations, refining provenance links, or linking words to scientific concepts; (3.2) Alice can also add annotations or comments to message content in Bob's diabetes channel; and (4) software from a clinic or other provider can share test results with Alice. Users get notifications for every activity performed on content under their PHLs, including annotations, replies, shares, reviews, citations, links, bookmarks, and even likes.

Data Access Through RESTful HTTP Operations

Data stored in the PHL is managed in a RESTful way; new resources are created under a container by sending them to the container's unique URI via an HTTP POST operation. The PHL supports two-way server-to-server and client-to-server communication. Either way, the requesting agent performs an HTTP operation on a given resource URI under a given POD server. Sending lab results from the clinic to Alice's inbox is an example of a server-to-server communication. The software agent (eg, Cerner) hosted on the clinic's server performs a POST operation on the /inbox/lab-tests resource on the POD server hosting Alice's PHL (Figure 2, step 4). A self-reported outcome is an example of a client-to-server communication that originates from a patient's PHL to external software. For example, Bob can report an allergy, which generates a POST request on a URI under the server hosting his EHR.

Agents do not need to know the internal structure of a patient's PHL. Each resource in PHL is its own SPARQL endpoint, which can be advertised through Link headers that can be discovered by sending HTTP GET/HEAD operations on the main PHL URI. Performing a GET operation on any URI that ends with a * returns an aggregate of all the resources that match the indicated pattern. For example, to fetch all diabetes self-management recommendations under Bob's diabetes container, an agent can send a single GET request to the /data/diabetes/* resource on the server hosting his PHL profile (Figure 11).

Figure 11. A GET request on Bob's diabetes messages folder under his POD.



Application-to-server communication happens when an application pushes content to a user's PHL. Users can interact with their PHLs through different apps installed on their phones after signing in using their WebIDs. For instance, when Bob installs the diabetes mHealth app on his phone, recommendations are pushed to the corresponding container within his messaging POD (Figure 1). By signing up, his PHL sends periodic GET requests to the app to get the latest recommendations. Similarly, Mary can add follow-up events using the calendar app of her choice, and that app can share those events by pushing them as resources under Bob's calendar container by sending POST requests on that container's URI.

Prototype Design of the PHL and mHealth App

Figure 12 shows the main interface design of the mobile app and Figure 13 shows the PHL. Through the PHL, patients can set their preferences to tailor the *content* in the desired recommendations in terms of focus (eg, diet, exercise, and medication adherence), frame (eg, educational, motivational, goal-based), and frequency (daily, weekly, etc).

Figure 12. Mobile app for chronic disease self-management.

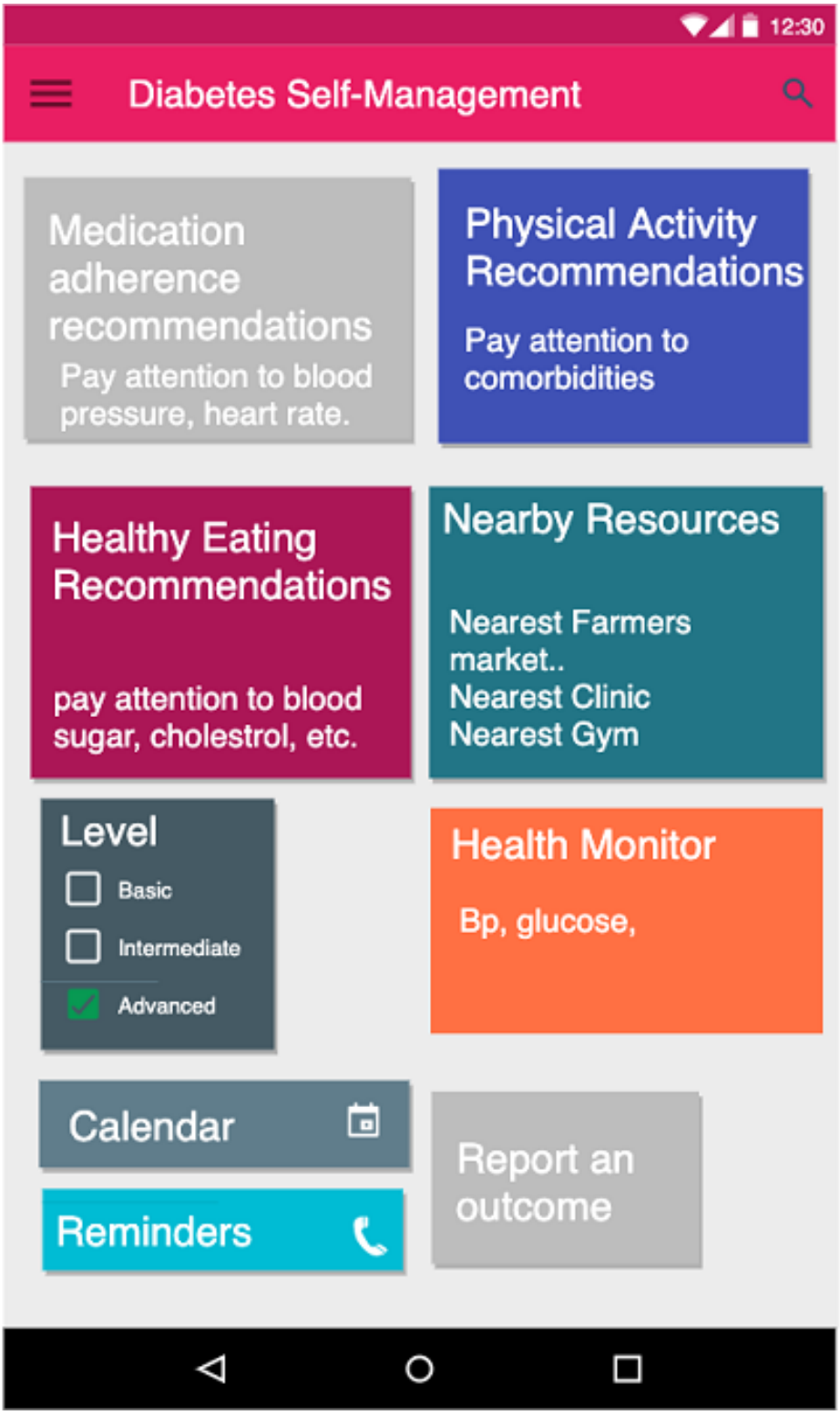
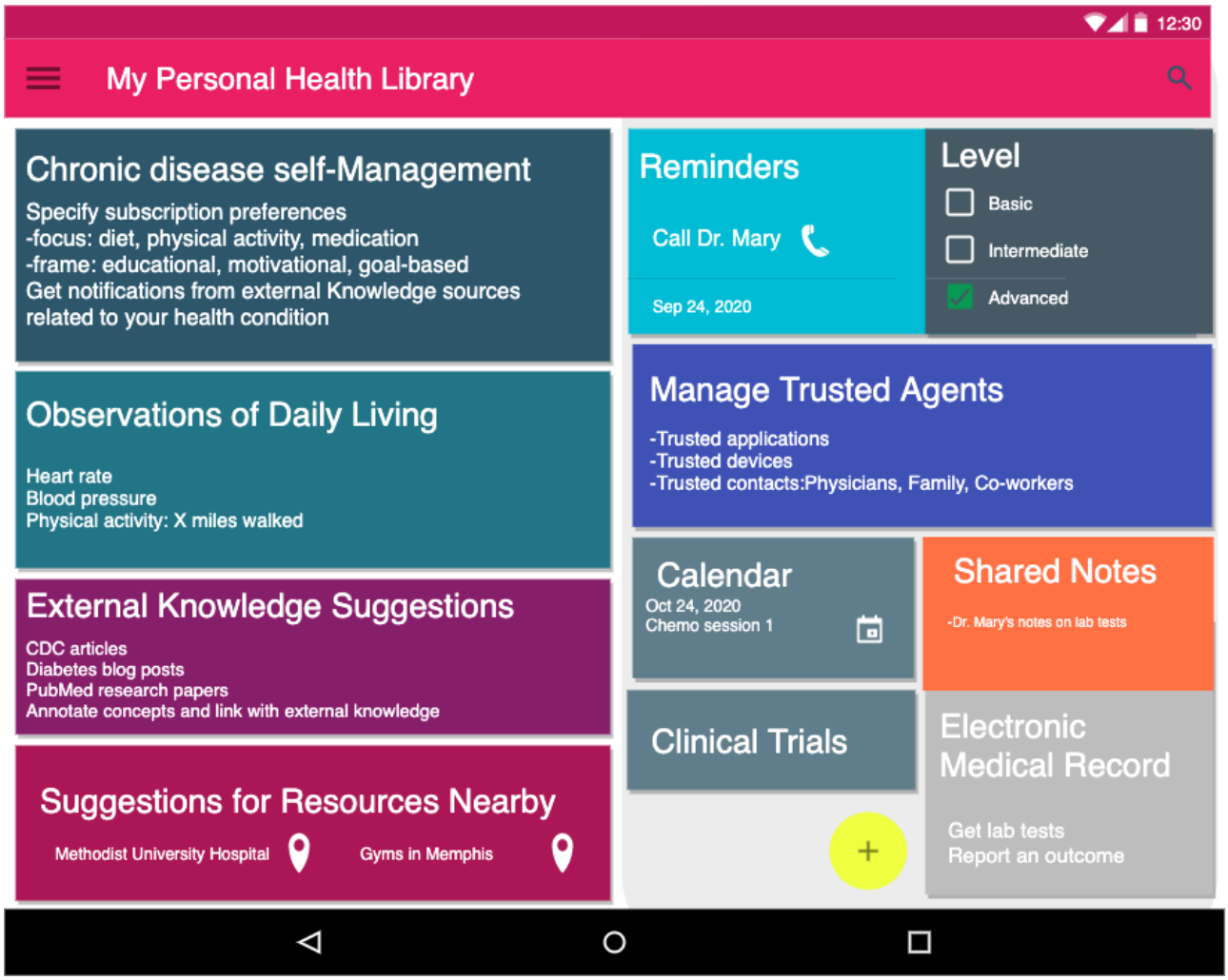


Figure 13. The PHL that enables the app in Figure 12. Through the PHL patients can perform (a) Chronic disease self-management, obtain (b) External knowledge and resource suggestions, and (c) manage trusted agents. The PHL can personalize recommendations by utilizing knowledge about monitored ODL readings, location-based detection of SDoH, external knowledge suggestions, and EHRs. Other features include reminders of medications and shared notepads with physicians.



The app, on the other hand, sends enriched message recommendations based on patients' preferences. By accessing dynamic knowledge in the PHL, the app can provide real-time hybrid recommendations that are both *content*- and *context*-based. To capture context, the PHL collects both ODL and SDoH data from activity trackers and population-level neighborhood characteristics. It utilizes both the data and the semantic conceptual hierarchies of knowledge types stored in the library to infer new knowledge and collect more evidence. For example, if a patient lives in a zip code that has a low walkability score, the app avoids sending recommendations that encourage walking in the neighborhood. In addition, if their device reading shows a high heart rate or if their EHR shows they have asthma, the app avoids recommendations of physical activities that would severely affect their health conditions. In addition to textual message recommendations, the app provides resource suggestions within the patients' zip code area or provides language-speaking services that respect their race or ethnicity.

Sources of Data and Knowledge

The PHL utilizes and integrates data and knowledge from several sources:

1. Historical data collected systematically through regional registries—in the case of the mobile app proposed in this paper, we utilize the Diabetes Wellness and Prevention Coalition (DWPC) regional registry, which aims to improve care for people with obesity, diabetes, and other obesity-associated chronic conditions. It provides clinical information, including hospital and clinic visits, as well as labs, diagnoses, and medications.

2. ODL data collected through mobile devices and activity trackers, such as physical exercises, heart rate, etc

3. Population-level SDoH data, such as walkability scores in a neighborhood and access to public transport, among others

4. Public knowledge obtained from research and news articles, blogs, and social networks

Enabling Technologies for Building the mHealth App

To implement PHL-enabled mobile apps, we will utilize a RESTful API that will expose the RDF-based PKG data model stored in a patient's PHL. We will leverage the Solid technology stack from within JavaScript-based environments, which will enable us to integrate data by invoking APIs exposed by different POD servers. For Linked Data manipulation and

querying, we will utilize RDFlib and LDflex. We will use the Solid-authenticated RDFlib API and query engine for advanced parsing and querying of the patient's RDF-based PKG data model stored in a patient's PHL.

In addition to querying the patient's local PKG data model, we will utilize the LDflex domain-specific language to query any Linked Data resource on the Web, which enables dynamic external knowledge discovery and integration. LDflex provides concise expressions that allow us to perform complex federated query execution without having to craft all HTTP requests required in a query and without hardcoding resource URLs. We provided the query engine with the root node (eg, Bob's WebID), an entry point (eg, Bob's inbox/lab_tests/test1), and a property (eg, testResult) to query Bob's health knowledge graph data model. The engine uses the entry point to recognize the context in which the query is executed and resolves the expression into an actual path on Bob's graph. Expression execution involves several steps: obtaining Bob's WebID URL, resolving the terms included in the expressions (eg, lab_test) to their URLs, creating the SPARQL query that represents the expression, fetching the document of the root user (Bob) through an HTTP request, and executing the SPARQL query on the document and returning the result.

Evaluation Plan

Before fully implementing the PHL and the mHealth app, we will conduct two evaluation studies. In this section, we provide a protocol synopsis for each study.

Study 1: User-Centered Design and Formative Evaluation of the Prototype

The so-called "digital divide" may hinder the adoption of digital interventions, but the use of human factors engineering can overcome some of the challenges in that regard [40,41]. We will conduct a descriptive, iterative, user-centered design and formative evaluation study by seeking feedback from pre-development focus groups, specifically participants from regional hospitals, including patients, caregivers, health care professionals (clinicians, residents, physicians), and health education professionals. The goal is to reveal issues related to the (1) usability, (2) clinical content, and (3) educational content of both the (a) PHL platform and (b) recommendations produced by the app. We will organize a workshop to both verify initial requirements, scenarios, and storyboards and identify new ones. We will use our findings to refine the initial requirements and develop new scenarios and generate new storyboard simulations. We will assign participants to focus groups and run different scenarios and dashboard simulations by them. We will use the System Usability Scale and the EHR Usability Scale to gather feedback [42,43]. Finally, we will apply thematic assessment to the resulting transcripts and use the results to design the final PHL and mHealth app following the refined storyboard simulations.

Study 2: Pragmatic Clinical Trial for Evaluation of the Digital Intervention

To evaluate the effectiveness of this intelligent digital intervention compared to existing interventions on the population of patients with diabetes, we will replicate the pragmatic randomized controlled trial design described by Bailey et al [44], with variations. The three primary *outcomes* that we will test are the number of days in the previous week that participants (1) ate healthy meals, (2) participated in at least 30 minutes of physical activity, and (3) took medications as prescribed. Participants will be adults age ≥18 years who have uncontrolled diabetes (A1c≥8), exhibit one or more additional chronic conditions seen in clinics in medically underserved areas of the mid-South, and possess a phone with texting and voicemail capability. Intervention arms will be standard motivational SMS text messaging (TM) and intelligent recommendation–enhanced TM (IR-TM), and the control arm will be usual care. We will capture the patients' perceptions of diabetes self-care activities using subscales of the revised Summary of Diabetes Self-Care Activities questionnaire administered over the study follow-up period. We will also use the DWPC registry to obtain clinical data, including A1c, body mass index, and blood pressure. We will test multiple hypotheses to determine the comparative effectiveness of the control and intervention arms for each primary outcome. Sample size and power estimates for these types of digital interventions will follow the approach described by Bailey et al [44], with the assumption that effect sizes will range from small (standardized difference=0.375) to medium (=0.50). To obtain power estimates, projected mean changes over 12-month follow-up from baseline (mean for 12-month follow-up minus mean for baseline) of the control and intervention arms for each primary outcome will be obtained using results reported in the literature. We expect the TM and IR-TM arms to have adequate power to detect meaningful changes from baseline with respect to all three primary outcome variables compared to the usual care arm will.

Discussion

Our previous research has identified that minorities and low-income, underserved communities are disproportionately affected by chronic diseases [45] such as obesity [46], a risk factor for diabetes, heart disease, and cancer. In this paper, we describe a personal health library–enabled mHealth app that provides hybrid recommendations by incorporating SDoH and ODLs in addition to digital health information to provide insights for informing preventive digital interventions in chronic disease management.

The PHL gathers different types of knowledge into a single searchable resource. While there has been some effort to build similar systems, the novelty of the proposed approach lies in (1) providing a decentralized yet linked architecture; (2) supporting interoperability, portability, knowledge mapping, and reasoning by following protocol, format, and vocabulary standards; (3) building trust with patients by facilitating true ownership over their data and appropriate reporting; and (4) giving those patients fine-grained access control mechanisms.

The PHL will not only help patients and their caregivers to assume a central role in making decisions regarding their health but also equip health care providers with informatics tools that will support the collection, interpretation, and dissemination of the collected knowledge. By moving health care beyond clinical

settings, clinicians can benefit from the PHL in leading new treatment regimens and keeping in touch with their patients between office visits.

Future work will focus on further implementation of an end-to-end framework of an intelligent recommender and digital librarian, including text summarization, knowledge mapping, and personalized resource suggestions. In achieving this goal, we will incorporate artificial intelligence techniques and knowledge representation methods that have been successfully used in our previous works [47,48]. Other ongoing tasks will include establishing a clinical trial of the app and recruiting participants to fully evaluate the app. Future work will also focus on the enrichment of patients' health knowledge graphs to improve the reasoning capabilities of the knowledge layer.

Finally, we plan to expose parts of the PHL functionality as an open service for fostering the development of third-party applications that may provide motivational technological support in several national and international projects crossing different domains of interest. To achieve this, the library will serve as an API for querying, managing, and using a patient's health RDF-based knowledge graph. This will give the community access to the infrastructure of the library to enable building applications that benefit from the library for other phenotypes.

Conflicts of Interest

None declared.

References

1. Baird A, North F, Raghu TS. Personal Health Records (PHR) and the future of the physician-patient relationship. 2011 Presented at: 2011 iConference; February 2011; Seattle, WA p. 281-288. [doi: 10.1145/1940761.1940800]
2. Laine C, Davidoff F. Patient-centered medicine. A professional evolution. JAMA 1996 Jan 10;275(2):152-156. [Medline: 8531314]
3. People-Centred Health Care: A policy framework. Geneva: World Health Organization; 2007.
4. Diez Roux AV, Katz M, Crews DC, Ross D, Adler N. Social and Behavioral Information in Electronic Health Records: New Opportunities for Medicine and Public Health. Am J Prev Med 2015 Dec;49(6):980-983. [doi: 10.1016/j.amepre.2015.08.027] [Medline: 26590943]
5. Backonja U, Kim K, Casper GR, Patton T, Ramly E, Brennan PF. Observations of daily living: putting the "personal" in personal health records. NI 2012 (2012) 2012;2012:6 [FREE Full text] [Medline: 24199037]
6. McWilliam CL. Patients, persons or partners? Involving those with chronic disease in their care. Chronic Illn 2009 Dec;5(4):277-292. [doi: 10.1177/1742395309349315] [Medline: 19933246]
7. Auerbach SM. Should patients have control over their own health care?: empirical evidence and research issues. Ann Behav Med 2000;22(3):246-259. [doi: 10.1007/BF02895120] [Medline: 11126470]
8. Ross SE, Todd J, Moore LA, Beaty BL, Wittevrongel L, Lin C. Expectations of patients and physicians regarding patient-accessible medical records. J Med Internet Res 2005 May 24;7(2):e13 [FREE Full text] [doi: 10.2196/jmir.7.2.e13] [Medline: 15914460]
9. Wildemuth B, Blake CL, Spurgin K, Oh S, Zhang Y. Patient's perspectives on personal health records: An assessments of needs and concerns. In: Proceedings of Critical Issues in eHealth Research 2006. 2006 Presented at: Critical Issues in eHealth Research 2006; 2006; National Institutes of Health.
10. Hassol A, Walker JM, Kidder D, Rokita K, Young D, Pierdon S, et al. Patient experiences and attitudes about access to a patient electronic health care record and linked web messaging. J Am Med Inform Assoc 2004;11(6):505-513 [FREE Full text] [doi: 10.1197/jamia.M1593] [Medline: 15299001]
11. Arora NK, Johnson P, Gustafson DH, McTavish F, Hawkins RP, Pingree S. Barriers to information access, perceived health competence, and psychosocial health outcomes: test of a mediation model in a breast cancer sample. Patient Educ Couns 2002 May;47(1):37-46. [doi: 10.1016/s0738-3991(01)00170-7] [Medline: 12023099]
12. Pratt W, Unruh K, Civan A, Skeels MM. Personal health information management. Commun ACM 2006 Jan;49(1):51-55. [doi: 10.1145/1107458.1107490]
13. Kay M, Santos J, Takane M. mhealth: New horizons for health through mobile technologies.: World Health Organization; 2011. URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2021-02-12]
14. Ammar N, Bailey JE, Davis RL, Shaban-Nejad A. The Personal Health Library: A Single Point of Secure Access to Patient Digital Health Information. Stud Health Technol Inform 2020 Jun 16;270:448-452. [doi: 10.3233/SHTI200200] [Medline: 32570424]
15. Ammar N, Bailey J, Davis RL, Shaban-Nejad A. Implementation of a Personal Health Library (PHL) To Support Self-Management of Chronic Diseases. In: Shaban-Nejad A, Michalowski M, Buckeridge DL, editors. Explainable AI in Healthcare and Medicine: Building a Culture of Transparency and Accountability. New York City: Springer; 2020.
16. World Health Organization. Social determinants of health. 2019. URL: https://www.who.int/social_determinants/sdh_definition/en/ [accessed 2020-03-02]

17. Barr PJ, Hassanpour S, Haslett W, Dannenberg MD, Ganoe CH, Faill R, et al. The Development of a Personal Audio Health Library (Audio Pahl) - Preliminary Usability Testing & Annotation Guide Development. 2019 Presented at: American Medical Informatics Association (AMIA) Summit on Informatics; 2019; San Francisco, CA.
18. Barr PJ, Hassanpour S, Haslett W, Dannenberg MD, Ganoe CH, Faill R, et al. Improving Patient and Caregiver Engagement through the Application of Data Science Methods to Audio Recorded Clinic Visits Stored in Personal Health Libraries Using ORALS. 2018 Presented at: American Medical Informatics Association Summit on Informatics; 2018; San Francisco, CA.
19. Brennan PF, Downs S, Casper G. Project HealthDesign: rethinking the power and potential of personal health records. J Biomed Inform 2010 Oct;43(5 Suppl):S3-S5 [FREE Full text] [doi: 10.1016/j.jbi.2010.09.001] [Medline: 20937482]
20. Cohen DJ, Keller SR, Hayes GR, Dorr DA, Ash JS, Sittig DF. Developing a model for understanding patient collection of observations of daily living: A qualitative meta-synthesis of the Project HealthDesign Program. Pers Ubiquitous Comput 2015 Jan 01;19(1):91-102 [FREE Full text] [doi: 10.1007/s00779-014-0804-1] [Medline: 26949381]
21. Ralston JD, Revere D, Robins LS, Goldberg HI. Patients' experience with a diabetes support programme based on an interactive electronic medical record: qualitative study. BMJ 2004 May 15;328(7449):1159 [FREE Full text] [doi: 10.1136/bmj.328.7449.1159] [Medline: 15142919]
22. Maimone R, Guerini M, Dragoni M, Bailoni T, Eccher C. PerKApp: A general purpose persuasion architecture for healthy lifestyles. J Biomed Inform 2018 Jun;82:70-87 [FREE Full text] [doi: 10.1016/j.jbi.2018.04.010] [Medline: 29729482]
23. Amelie G, Manas G, Saeedeh S, Krishnaprasad T, Amit S. Personalized health knowledge graph. In: Proceedings of ISWC 2018 Contextualized Knowledge Graph Workshop. 2018 Presented at: ISWC 2018 Contextualized Knowledge Graph Workshop; 2018; Monterey, CA p. 29.
24. Rotmensch M, Halpern Y, Tlimat A, Horng S, Sontag D. Learning a Health Knowledge Graph from Electronic Medical Records. Sci Rep 2017 Jul 20;7(1):5994 [FREE Full text] [doi: 10.1038/s41598-017-05778-z] [Medline: 28729710]
25. Sheth, Amit, Pramod A, Krishnaprasad T. khealth: Proactive personalized actionable information for better healthcare. 2014 Presented at: First International Workshop on Personal Data Analytics in the Internet of Things; 2014; France URL: https://pdfs.semanticscholar.org/9722/6061407b61d9eb631698f315a3d704b55503.pdf
26. Seneviratne O, Rashid SM, Chari S, McCusker JP, Bennett KP, Hendler JA, et al. Knowledge integration for disease characterization: A breast cancer example. 2018 Presented at: International Semantic Web Conference; October 2018; Monterey, CA p. 223-238. [doi: 10.1007/978-3-030-00668-6_14]
27. Chari S, Qi M, Agu NN, Seneviratne O, McCusker JP, Bennett KP, et al. Making study populations visible through knowledge graphs. 2019 Presented at: International Semantic Web Conference; October 2019; Auckland, New Zealand p. 53-68. [doi: 10.1007/978-3-030-30796-7_4]
28. Linked Data. W3C. 2006. URL: https://www.w3.org/DesignIssues/LinkedData [accessed 2021-03-02]
29. The Web Annotation specification. W3C. URL: https://www.w3.org/TR/annotation-model/ [accessed 2021-03-02]
30. Zeng J, Shufean MA, Khotskaya Y, Yang D, Kahle M, Johnson A, et al. OCTANE: Oncology Clinical Trial Annotation Engine. JCO Clin Cancer Inform 2019 Jul;3:1-11 [FREE Full text] [doi: 10.1200/CCI.18.00145] [Medline: 31265323]
31. Fielding RT, Taylor RN. Architectural Styles and the Design of Network-Based Software Architectures (Vol. 7). Irvine: University of California, Irvine; 2000.
32. Vander Sande M, Verborgh R, Dimou A, Colpaert P, Mannens E. Hypermedia-Based Discovery for Source Selection Using Low-Cost Linked Data Interfaces. International Journal on Semantic Web and Information Systems 2016;12(3):79-110. [doi: 10.4018/ijswis.2016070103]
33. Verborgh R, Dumontier M. A Web API Ecosystem through Feature-Based Reuse. IEEE Internet Comput 2018 May;22(3):29-37. [doi: 10.1109/mic.2018.032501515]
34. Verborgh R, Vander Sande M, Hartig O, Van Herwegen J, De Vocht L, De Meester B, et al. Triple Pattern Fragments: A Low-Cost Knowledge Graph Interface for the Web. Journal of Web Semantics 2016;37-38:184-206. [doi: 10.2139/ssrn.3199223]
35. Yeung CMA, Liccardi I, Lu K, Seneviratne O, Berners-Lee T. The future of online social networking. In: W3C Workshop on the Future of Social Networking Position Papers (Vol. 2). 2009 Presented at: W3C Workshop on the Future of Social Networking Position Papers; January 2009; Barcelona, Spain p. 2-7.
36. Sambra AV, Mansour E, Hawke S, Zareba M, Greco N, Ghanem A, et al. Solid: A Platform for Decentralized Social Applications Based on Linked Data. Technical Report. 2017. URL: http://www.emansour.com/publications/paper/solid_protocols.pdf [accessed 2021-02-19]
37. Sambra AV, Story H, Berners-Lee T. WebID Specification. 2014. URL: http://www.w3.org/2005/Incubator/webid/spec/identity/ [accessed 2021-02-19]
38. Krisztian BK, Kenter T. Personal Knowledge Graphs: A Research Agenda. : ACM; 2019 Presented at: ACM SIGIR International Conference on Theory of Information Retrieval; 2019; Paris, France p. 217-220. [doi: 10.1145/3341981.3344241]
39. Capadisli S, Guy A, Lange C, Auer S, Sambra A, Berners-Lee T. Linked data notifications: a resource-centric communication protocol. 2017 Presented at: European Semantic Web Conference; May 2017; Slovenia p. 537-553. [doi: 10.1007/978-3-319-58068-5_33]

40. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, et al. The digital divide in adoption and use of a personal health record. Arch Intern Med 2011 Mar 28;171(6):568-574. [doi: 10.1001/archinternmed.2011.34] [Medline: 21444847]
41. Sittig DF, Wright A, Osheroff JA, Middleton B, Teich JM, Ash JS, et al. Grand challenges in clinical decision support. J Biomed Inform 2008 Apr;41(2):387-392 [FREE Full text] [doi: 10.1016/j.jbi.2007.09.003] [Medline: 18029232]
42. System usability scale (SUS). Usability.gov. 2018. URL: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html [accessed 2021-03-02]
43. Wiklund M, Kendler J, Hochberg L, Weinger M. Technical basis for user interface design of health IT (NIST GCR 15-996). Washington: National Institute of Standards and Technology, US Department of Commerce; 2015.
44. Bailey JE, Surbhi S, Gatwood J, Butterworth S, Coday M, Shuvo SA, et al. The management of diabetes in everyday life study: Design and methods for a pragmatic randomized controlled trial comparing the effectiveness of text messaging versus health coaching. Contemp Clin Trials 2020 Sep;96:106080. [doi: 10.1016/j.cct.2020.106080] [Medline: 32653539]
45. Shin EK, Kwon Y, Shaban-Nejad A. Geo-clustered chronic affinity: pathways from socio-economic disadvantages to health disparities. JAMIA Open 2019 Oct;2(3):317-322 [FREE Full text] [doi: 10.1093/jamiaopen/ooz029] [Medline: 31984364]
46. Addy NA, Shaban-Nejad A, Buckeridge DL, Dubé L. An innovative approach to addressing childhood obesity: a knowledge-based infrastructure for supporting multi-stakeholder partnership decision-making in Quebec, Canada. Int J Environ Res Public Health 2015 Jan 23;12(2):1314-1333 [FREE Full text] [doi: 10.3390/ijerph120201314] [Medline: 25625409]
47. Ammar N, Shaban-Nejad A. Explainable Artificial Intelligence Recommendation System by Leveraging the Semantics of Adverse Childhood Experiences: Proof-of-Concept Prototype Development. JMIR Med Inform 2020 Nov 4;8(11):e18752. [doi: 10.2196/18752] [Medline: 33146623]
48. Brenas JH, Shin EK, Shaban-Nejad A. A Hybrid Recommender System to Guide Assessment and Surveillance of Adverse Childhood Experiences. Stud Health Technol Inform 2019 Jul 04;262:332-335. [doi: 10.3233/SHTI190086] [Medline: 31349335]

Abbreviations

ACL: access control list
API: application programming interface
DWPC: Diabetes Wellness and Prevention Coalition
EHR: electronic health record
IR-TM: intelligent recommendation–enhanced TM
LDP: Linked Data Platform
LDN: Linked Data Notifications
LOD: Linked Open Data
LOR: Linked Open Research
ODLs: observations of daily living
PHKG: personal health knowledge graph
PHL: personal health library
POD: personal online data store
REST: representational state transfer
RDF: Resource Description Framework
RDFa: RDF in Attributes
SDoH: social determinants of health
Solid: Social Linked Data
SPARQL: SPARQL Protocol and RDF Query Language
TM: text messaging
W3C: World Wide Web Consortium
WAC: Web Access Control
WAO: Web Annotation Ontology

Edited by G Eysenbach; submitted 02.10.20; peer-reviewed by Y Chu; comments to author 26.10.20; revised version received 08.12.20; accepted 12.02.21; published 16.03.21.

Please cite as:
Ammar N, Bailey JE, Davis RL, Shaban-Nejad A
Using a Personal Health Library–Enabled mHealth Recommender System for Self-Management of Diabetes Among Underserved Populations: Use Case for Knowledge Graphs and Linked Data
JMIR Form Res 2021;5(3):e24738
URL: https://formative.jmir.org/2021/3/e24738
doi:10.2196/24738
PMID:33724197

Original Paper

Combining a Hudl App With Telehealth to Increase Home Exercise Program Adherence in People With Chronic Diseases Experiencing Financial Distress: Randomized Controlled Trial

Ann Van de Winckel[1], MSPT, PT, PhD; Tanjila Nawshin[2], MBBS, MS; Casey Byron[3], BS, MPT, CCCE

[1]Division of Physical Therapy, Division of Rehabilitation Science, Department of Rehabilitation Medicine, Medical School, University of Minnesota, Minneapolis, MN, United States

[2]Division of Rehabilitation Science, Department of Rehabilitation Medicine, Medical School, University of Minnesota, Minneapolis, MN, United States

[3]Hennepin County Medical Center, Minneapolis, MN, United States

Corresponding Author:
Ann Van de Winckel, MSPT, PT, PhD
Division of Physical Therapy, Division of Rehabilitation Science
Department of Rehabilitation Medicine, Medical School
University of Minnesota
420 Delaware St SE (MMC 388)
Minneapolis, MN, 55455
United States
Phone: 1 612 625 1191
Fax: 1 612 625 4274
Email: avandewi@umn.edu

Abstract

Background: Patients with chronic diseases often need to adhere to long-term individualized home exercise programs (HEPs). Limited adherence to long-term exercise given during physical therapy (PT) visits reduces the capacity of exercise to manage or improve symptoms related to chronic disease. In addition, a lower socioeconomic status negatively impacts exercise adherence. To mitigate this, apps that motivate people to exercise could be a viable option. Using an app through telehealth may help adults with chronic diseases to achieve long-term HEP adherence. However, because apps for rehabilitation are an emerging field, the feasibility of the app needs to be evaluated.

Objective: To address HEP adherence in participants with chronic diseases who are experiencing financial distress, we aim to evaluate the feasibility of and satisfaction with the Hudl Technique app and telehealth and satisfaction with PT care and to monitor HEP adherence and compliance (ie, percentage of participant-recorded videos sent) in participants using the app with telehealth compared with those using standard HEPs on paper.

Methods: We recruited patients scheduled for outpatient PT. We performed a randomized controlled trial in which the experimental group received weekly HEP demonstrations through app videos on a tablet with feedback on their self-recorded HEP video performance from the telehealth physical therapist. The control group received HEPs on paper without feedback, as is customary in PT practice. Demographic, clinical, and health coverage information was collected for screening and baseline measurements. Adherence and compliance were evaluated. Both groups completed surveys at 8 and 24 weeks on their satisfaction with PT care, and the experimental group also completed a survey on their satisfaction with the app with telehealth use. Descriptive and nonparametric statistics were used for within-group and between-group comparisons and analyzed with JMP, version 13.

Results: Overall, 45 adults with chronic diseases who were experiencing financial distress were randomized into experimental (23/45, 51%) and control (22/45, 49%) groups, with 74% (17/23) and 86% (19/22) participants completing the 24-week HEP, respectively. The experimental group had an HEP adherence frequency of 4 (SD 2) to 5 (SD 2) times per week at 8 and 24 weeks (P=.14), whereas HEP adherence decreased in the control group from 4 (SD 2) to 3 (SD 2) times per week (P=.07), with a significant difference (P=.01) between groups at 24 weeks. Of the total participants, 68% (15/22) sent videos. They sent 68% (16/24) of the requested number of videos on average. The average score for PT care satisfaction was maintained at 87% in the experimental group (P=.99), whereas it decreased from 89% at 8 weeks to 74% at 24 weeks (P=.008) in the control group. App-related adverse events were not observed.

Conclusions: The Hudl app/telehealth platform is feasible for delivering HEPs and maintaining HEP adherence in participants with chronic diseases who are experiencing financial distress.

Trial Registration: ClinicalTrials.gov NCT02659280; https://clinicaltrials.gov/ct2/show/NCT02659280

(JMIR Form Res 2021;5(3):e22659) doi:10.2196/22659

KEYWORDS

chronic disease; spinal cord injury; stroke; telehealth; telemedicine; traumatic brain injury

Introduction

Background

Approximately 200 million Americans have at least one chronic disease, and 80 million have multiple chronic diseases [1-3]. Patients with chronic health conditions often must adhere to a long-term individualized home exercise program (HEP) to manage their symptoms and improve or maintain their cardiovascular health, flexibility, and/or strength [4]. Most HEPs are delivered on paper to be practiced at home, and accountability is checked at the next physical therapy (PT) visit [5]. However, low HEP adherence is common, especially when exercising is required over a longer period [4-6]. A recent meta-analysis in people with chronic health conditions supports this observation, as the authors reported only 33% full HEP adherence and 37% partial adherence, with no difference in adherence when follow-up was done face-to-face or over the phone [7]. Low HEP adherence is not entirely surprising as it is challenging for patients to stay on an HEP for an extended period. In fact, a study showed that patients start to slack in HEP adherence as early as after 2 weeks [8]. Low HEP adherence is a serious problem because it will compromise the improvement or stabilization of chronic health issues.

Exercise adherence is even more negatively impacted when patients have a lower socioeconomic status [9]. Patients experiencing financial distress do not have as many options to manage their chronic disease as do people with a higher socioeconomic status, who can visit a health club or hire a personal trainer. Consequently, the health situation of adults experiencing financial distress can worsen and eventually they may have to lean on resources from the health care system more often. This is a serious problem for the patients and for the health care system. Thus, it is critical for clinics and hospitals to find alternative options to reach and treat underserved patients with a challenging socioeconomic status, especially regarding the long-term management of chronic diseases. More specifically, there is a need to find an effective delivery method for HEP that (1) is more inspiring and satisfying to these patients so that they are motivated to keep practicing at home for an extended period [10] and (2) can be used as a tool to provide professional feedback on exercise correctness [5].

To achieve this goal in the long run, Medtronic Philanthropy funded our exploratory study, titled *Expanding Access to Physical Therapy for Underserved Patients with Chronic Disease*. In this study, we looked into apps on mobile phones or on touchscreen tablets to determine whether they could potentially be an effective tool to provide such feedback and motivation [5,11,12]. Regular feedback from a PT has shown to be essential to correctly execute the HEP [8,13]. However, to date, studies in which apps were used for rehabilitation provided only occasional phone support from a PT or used a virtual coach to motivate participants to perform the app-based exercises. These methods limit the extent to which individualized feedback can be provided [5,10,14]. To overcome this problem, we have chosen to use the free Hudl Technique app aiming to help patients with HEP adherence. This app, originally designed for athletes' coaches, enables patients and the PT to record and share exercises on video; most importantly, it enables the PT to provide individualized feedback on the patients' self-recorded videos by using overlaying audio and visual feedback cues [15]. An additional benefit of using the Hudl app as a part of telehealth (ie, providing and sending individualized feedback from a PT to a patient via the Hudl app) is that this part is reimbursable through Medicaid or similar health care coverage [16,17], thereby making it available to patients with lower socioeconomic status who are in need of follow-up and feedback on their long-term HEPs.

To date, the use of the Hudl app in health care has been limited because app-based rehabilitation approaches are relatively new. Only 3 studies have used the Hudl app in sports medicine such that the app was (1) used to demonstrate features for potential use in sports medicine [18]; (2) used as a training tool for athletes to avoid sport-related concussions [15]; or (3) found to be unreliable for 2D kinematic analysis to evaluate patellofemoral pain during running compared with a laboratory-based 3D analysis [19]. In fact, to the best of our knowledge, no studies have reported the use of the Hudl app for rehabilitation.

Objectives

Therefore, we aimed to conduct an exploratory study in adults with chronic diseases who were experiencing financial distress to investigate whether the Hudl app combined with telehealth (experimental group) results in better HEP adherence than when HEP is delivered through the standard paper format (control group). We defined adults experiencing financial distress as those receiving Medicaid or similar health care assistance. Over 7 months, we sampled adults with chronic diseases who were experiencing financial distress by visiting a general outpatient PT clinic and recruited a total of 45 participants, 38 of whom were classified as having a neurological chronic disease (stroke, spinal cord injury, traumatic brain injury, Guillain-Barre syndrome, or Parkinson disease) and 7 with other types of chronic diseases, including heart disease, diabetes mellitus, low back pain, autoimmune disease, and chronic obstructive pulmonary disease.

Owing to the limited funding and recruitment setting, the recruitment results did not allow us to recruit sufficiently large sample sizes per specific diagnostic category to allow for separate subanalyses per diagnosis. Nonetheless, we decided to proceed with this study on the Hudl app with this patient pool for four reasons: first, as the Hudl app has not been used in research for rehabilitation purposes, it was essential to determine whether patients, especially those with neurological chronic diseases, were able to use the app. Second, it was important to determine whether patients with financial hardship could use the app when they may or may not have sufficient Wi-Fi access at home or may not always have a home. Third, it was important to determine whether patients would demonstrate long-term HEP adherence by using the app, beyond the initial period of enthusiasm for using a new tool. It should be noted that we followed up the participants for an extended period (ie, 24 weeks). Finally, if those aspects of feasibility were met, we needed to determine whether our patient pool collectively achieved better HEP adherence with the use of the app, compared with when exercises were provided in paper format. To address these four key components, we incorporated questions in our survey on the ease of use of the equipment, Wi-Fi access, HEP adherence, and overall telehealth and PT satisfaction as patient satisfaction has been linked to the success of treatment adherence [20-23]. We measured HEP adherence (ie, how often they practice HEP per week) and compliance in returning self-recorded videos weekly and sent out a survey at 8 and 24 weeks to inquire about their satisfaction with PT care and, for the experimental group, their satisfaction with the Hudl app/telehealth platform in terms of feasibility. The features of the Hudl app included videos of the physical therapist demonstrating the exercises and individualized feedback given by the physical therapist through digital features of the Hudl app projected on the self-recorded patient videos. We hypothesize that these unique features of the Hudl app would likely motivate the participants to stay in the program.

Methods

Recruitment

Between March 10, 2016, and November 5, 2017, we recruited participants through consecutive sampling who were discharged from the Hennepin County Medical Center and set to receive outpatient PT. Inclusion criteria were adults (≥18 years) who had a chronic disease and financial hardship defined as having Emergency Medical Assistance, Medical Assistance, Medicaid, Medicaid MCO (Medicare Contracted Organizations), hospital discount funding (*Hennepin Care*), or no insurance as those health care coverages are assigned only to adults experiencing financial distress. We assessed their ability to use a tablet device, internet features, and the Hudl app. Participants provided written informed consent. The study was approved by the institution's internal review board (HSR#15-4040) and was performed in accordance with the Declaration of Helsinki.

Study Design

Using a randomized controlled design, 60 sealed envelopes were used for 1:1 participant allocation. The treating therapist and data analyst were blinded to the allocation. Through electronic medical record review, we collected baseline data on age, sex, race, spoken language, health coverage status reflecting their financial situation, and diagnosis. The participant's treating physical therapist designed and demonstrated the HEP during the first PT visit. All patients were instructed to perform exercises daily at home without any equipment.

The control group received individualized HEPs on paper and did not receive feedback. To maintain the same incentive for both groups, the control group also received a tablet but without the Hudl app installed. The participants in the experimental group received a tablet and were trained on using the Hudl app. The experimental group received weekly videos with individualized HEP (ranging between 4 minutes 37 seconds and 9 minutes 23 seconds), as demonstrated by the telehealth physical therapist. After reviewing the video of the telehealth physical therapist, the participants video recorded their HEP performance, showing at least one set of repetitions of all required exercises. In addition, they reported whether they had completed all the required sets. The required sets included 3 trials per set for all strengthening exercises (maximum 10 repetitions each) and 3 trials per set for all balance exercises (maximum 30 seconds each). The telehealth physical therapist reviewed the content and compliance to the HEPs and created a critiqued video review using the Hudl app tools (Figure 1). The most frequently used tools were (1) the *voice-over* to offer feedback; (2) the *chronometer* to record the time required to perform, for example, a single-limb stance test; (3) the *goniometer* to measure joint angles; (4) the *visual tool* displaying 2 videos side by side to show weekly progress or to compare the telehealth physical therapist's video with the participant's video, highlighting the differences in execution; (5) the *drawing tool* to review the participant's postural symmetry during exercises; and (6) the *slow-motion video* to analyze dynamic balance and exercise sequences. The participants reviewed the critiqued video and, within a few days, received a new video from the telehealth physical therapist with a HEP for the following week. The participants only received a new video if they returned a video to a telehealth physical therapist.

Figure 1. Examples of Hudl app video captures of the telehealth physical therapist's home exercise program. The telehealth physical therapist used the tools displayed on the right-hand side of the screen (arrows, plumb line, chronometer, etc) to highlight the parts that participants needed to pay attention to.

At 8 and 24 weeks, participants completed a questionnaire anonymously on the web (17/45, 38%) or in paper format (28/45, 62%) if they had trouble completing the web-based survey. The survey asked the participants to quantify their average HEP adherence per week; overall PT satisfaction; and, for the experimental group, satisfaction with the app and telehealth use. Scoring for questions ranged from 1 (very dissatisfied or difficult) to 7 (very satisfied or easy). Overall PT satisfaction encompassed both the Hudl app with telehealth platform use and care or feedback from the telehealth physical therapist and treating PT service for the experimental group and HEP on paper and the treating PT service for the control group.

Statistical Analysis

Power

The sample size was determined through a priori power analysis using G.Power 3.1 and based on data from a study with a similar setup, which reported 71% (SD 25%) training adherence in the group using a tablet versus 48% (SD 42%) in the group using a paper version [24]. With 85% power, α=.05, and a 1:1 enrollment ratio, we needed 16 participants per group. We aimed to recruit at least 5 additional persons per group to account for a 31% attrition rate, as reported in a previous study [24].

Data Analysis

Data were analyzed using JMP, version 13 (SAS Institute Inc, 1989-2007). The Shapiro-Wilk test informed the decision to conduct *t* tests versus Mann-Whitney *U* tests for between-group comparisons at 24 weeks or a paired *t* test versus Wilcoxon tests for within-group comparisons at 8 and 24 weeks. Nominal data were calculated using chi-square test for between-group comparisons [25] or with the McNemar test for within-group comparisons [26].

For the 7 questions on app with telemedicine use, we reported descriptively on the percentage of people who were unsatisfied (scores 1-3), neutral (score 4), or satisfied (scores 5-7). We used the McNemar test to evaluate differences in the ratio of people who were unsatisfied or neutral versus satisfied between 8 and 24 weeks. We reported on the proportion of people filling in the survey and the percentage of people satisfied with the PT treatment (scores 5-7). The average score (in percentage) given on overall PT satisfaction at 8 and 24 weeks was compared between and within groups.

We compared between- and within-group differences in percentage HEP adherence (based on the patient-reported number of sessions completed per week). We calculated the effect size for between-group differences in HEP adherence at 24 weeks (Hedges *g* for unequal sample sizes) [27] and repeated-measures effect size ($d_{repmeas}$) [28] for within-group differences in HEP adherence [29].

We calculated the percentage of people returning videos and, of those participants, how many videos they returned compared with the requested weekly videos (reported in percentage).

Results

Demographics and Patient Characteristics

The CONSORT flowchart (Figure 2) illustrates the steps in the study. Of the 51 participants approached, 45 were randomized into experimental (n=23) and control (n=22) groups, and 17 and 19 participants completed the study, respectively. As stated earlier, of the 45 participants, 38 were classified as having a chronic neurological disease (stroke, spinal cord injury, traumatic brain injury, Guillain-Barre syndrome, or Parkinson disease) and 7 with other types of chronic diseases, including

heart disease, diabetes mellitus, low back pain, autoimmune disease, and chronic obstructive pulmonary disease. Demographic (Table 1) and clinical (Table 2) data were reported for each group. There were no significant differences between the groups with respect to age (U=−0.72; P=.47), sex (χ^2_1=1.1; P=.30), race (χ^2_1=6.2; P=.29), or language (χ^2_1=5.3; P=.15). Table 2 details the type of chronic diseases the participants had and the duration of their chronic disease. Table 2 also mentions the type of individualized HEPs the participants received to manage their symptoms or impairments related to the chronic disease.

Figure 2. The CONSORT (Consolidated Standards of Reporting Trials) flowchart.

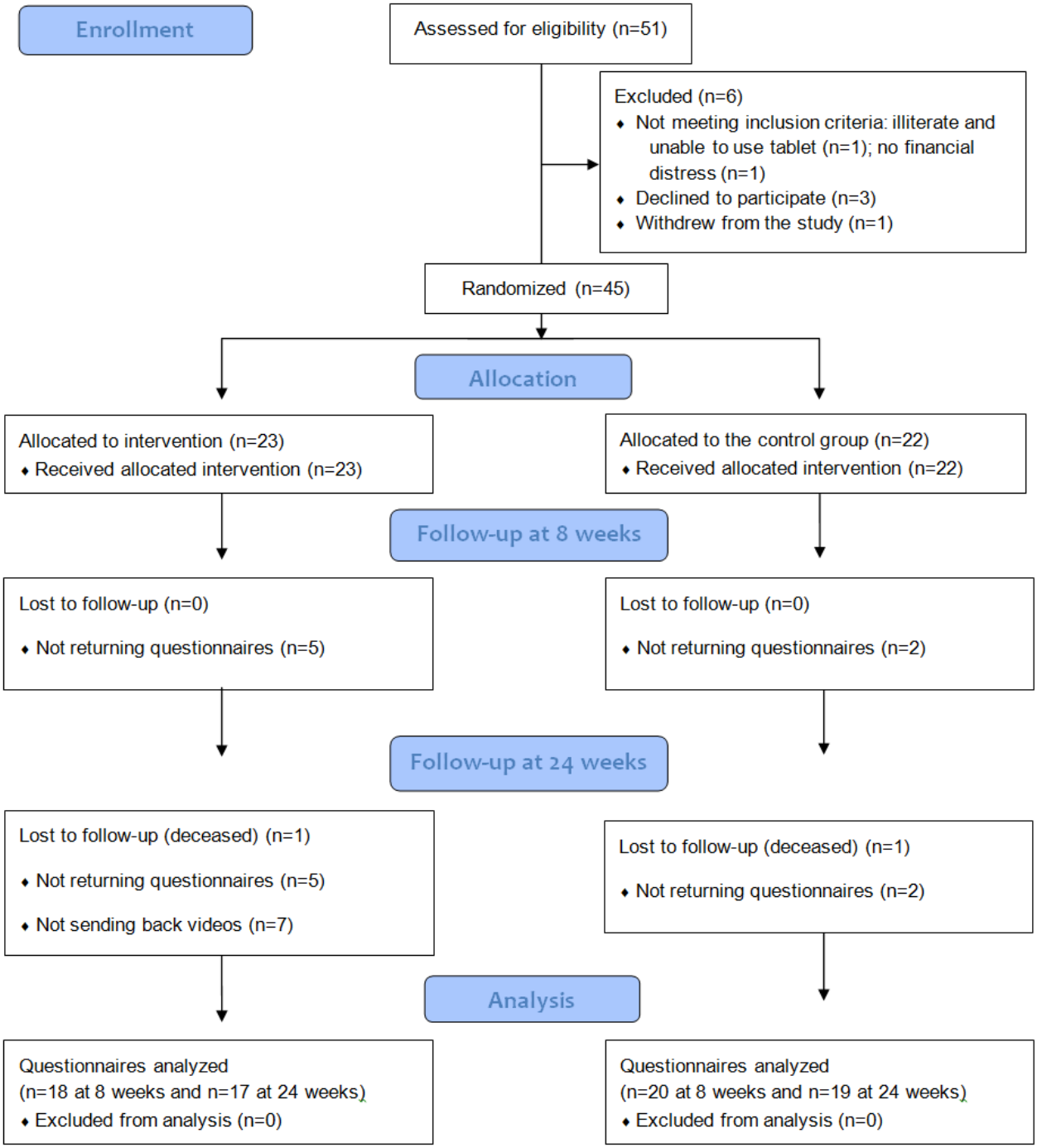


Table 1. Demographic data of the participants per allocated group (N=45).

Categories	Experimental group (n=23)	Control group (n=22)	Statistical tests[a]	*P* value
Age (years), mean (SD)	47.26 (13.33)	44.59 (13.79)	U=−0.72	.47
Sex, n (%)			χ^2_1=1.1	.30
Male	16 (70)	12 (55)		
Female	7 (30)	10 (45)		
Race, n (%)			χ^2_1=6.2	.29
African	2 (9)	2 (9)		
American Indian	0 (0)	1 (4)		
Asian	1 (4)	0 (0)		
Black	9 (39)	7 (32)		
White	9 (39)	5 (23)		
Latino	2 (9)	7 (32)		
Language, n (%)			χ^2_1=5.3	.15
Amharic	1 (4)	0 (0)		
English	21 (92)	16 (73)		
Igbo	0 (0)	1 (4)		
Spanish	1 (4)	5 (23)		

[a]Statistical tests: the chi-square test was used for sex, race, and language and the Mann-Whitney *U* test was used for age.

Table 2. Clinical data of the participants per allocated group (n=45).

Diagnosis	Experimental group (n=23)		Control group (n=22)		Type of home exercise program
	n (%)	Duration of symptoms[a]	n (%)	Duration of symptoms[a]	
Autoimmune disease	0 (0)	N/A[b]	1 (5)	3.5 months	Stretching, strengthening, and balance
Chronic obstructive pulmonary disease	1 (4)	6 years and 2 months	0 (0)	N/A	Stretching and strengthening
DM[c]	1 (4)	13 years and 8 months	1 (5)	20.5 years	Strengthening and balance
DM and chronic low back pain	1 (4)	16 years and 9 months	1 (5)	5 years and 2 months for DM, 3 months for low back pain	Strengthening
Guillain-Barre syndrome	0 (0)	N/A	1 (5)	3.5 months	Strengthening and assisted mobility
Heart disease	1 (4)	7 years and 7 months	0 (0)	N/A	Stretching, strengthening, balance, and cardiorespiratory interval training
Parkinson disease	1 (4)	4 months	0 (0)	N/A	Stretching, strengthening, and assisted mobility
Spinal cord injury	3 (13)	Between 3.5 months and 20 years	2 (9)	Between 5 months and 17 years	Strengthening, cardiorespiratory training, assisted stretching, and balance
Stroke	6 (27)	Between 3 and 5 months	9 (40)	Between 3 and 5.5 months	Strengthening, cardiorespiratory training, stretching, balance, assisted mobility, and functional activity
Traumatic brain injury	9 (40)	Between 3 and 4 months	7 (31)	Between 3 and 5 months	Balance, strengthening, cardiorespiratory training, assisted stretching, and assisted mobility

[a]For n=1, the duration is the exact duration of that person's symptom. For n>1, the duration is the range of durations of the symptom.

[b]N/A: not applicable.

[c]DM: diabetes mellitus.

There were no app-related adverse events. Two participants died because of unrelated causes before the 24-week questionnaires were distributed. The electronic medical record stated that 1 participant (experimental group) died from natural causes, and the other participant (control group) had a fatal myocardial infarction.

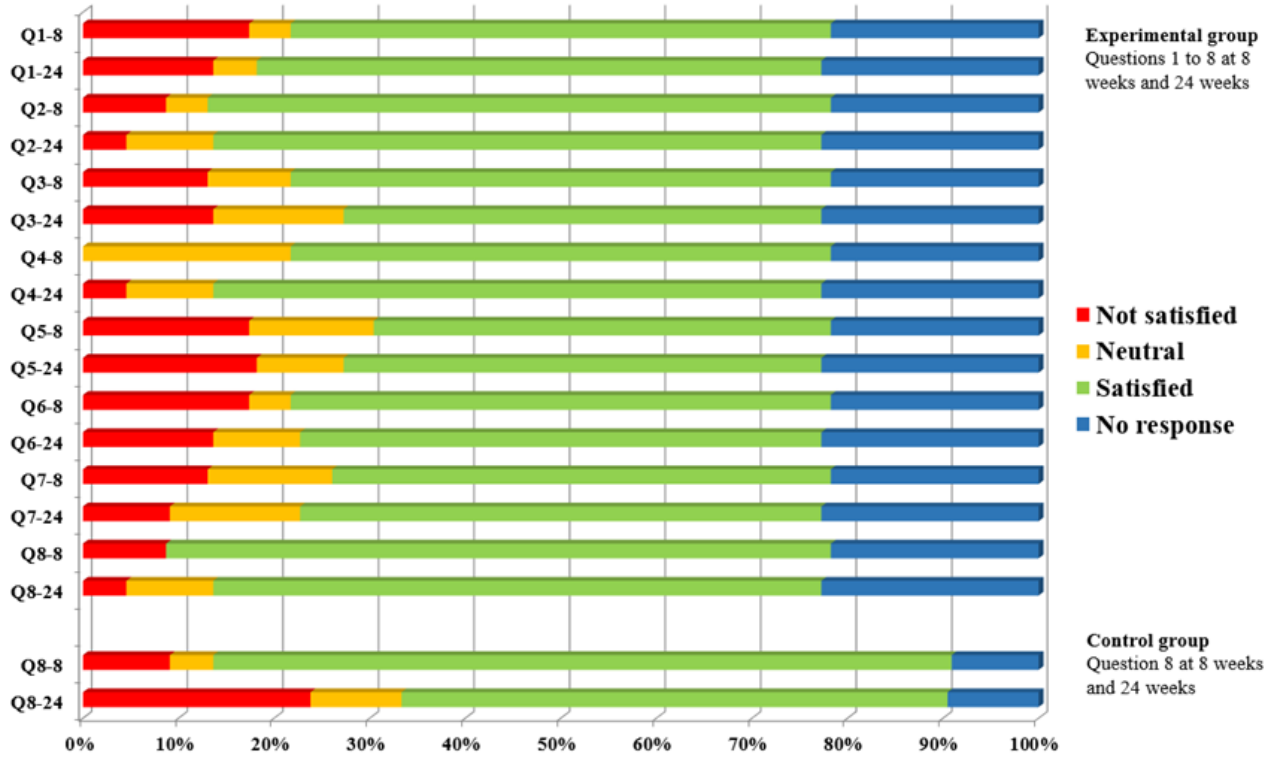
Primary Outcome

Satisfaction With PT Care and Use of the App and Telehealth in Terms of Its Feasibility

The average scores (in %) given on the overall PT experience in both groups were determined. There was no significant between-group difference at 24 weeks (U=342.5; P=.34). The within-group average score for PT satisfaction was maintained in the experimental group at 8 and 24 weeks (87%, SD 21%, to 87%, SD 19%; Wilcoxon S=−5.50; P=.99) but decreased significantly over time in the control group (89%, SD 24% to 74%, SD 31%; S=−62.00; P=.008).

The percentage of people who were satisfied to very satisfied (scores 5-7) with the app and telemedicine use in terms of its feasibility is shown in Multimedia Appendix 1. There was no significant change in the ratio of people satisfied versus neutral or not satisfied between 8 and 24 weeks or between groups on the overall PT satisfaction question at 8 or 24 weeks. The percentage of people who were not satisfied (scoring 1-3, which meant *very unsatisfied* to *a little unsatisfied*), neutral (score 4), or satisfied (*a little satisfied* to *very satisfied*, scores 5-7) on all questions, including the overall PT satisfaction for both groups are shown in Figure 3.

Figure 3. Percentage responses in categories unsatisfied (red), neutral (orange), satisfied (green), and no response (blue) from the survey completed by the experimental and control groups. The experimental group answered 7 questions related to app with telemedicine use, feasibility, and ease of interaction with the telehealth physical therapist. The eighth question was answered by both groups and pertained to their rating of satisfaction with the overall physical therapy experience. The answer options ranged from scores 1 to 3 (being not satisfied [scoring very unsatisfied to a little unsatisfied]), score 4 (neutral), or scores 5 to 7 (satisfied [scoring a little satisfied to very satisfied]) or no response. The question and answer options for each question are detailed at the bottom of the figure.



Questions on the survey

1.	How easy was it to interact with the telehealth physical therapist?	[Range 1-7: very difficult-very easy]
2.	How much did the telehealth physical therapist seem to care about you as a person?	[Range 1-7: very little-very much]
3.	Did you feel relaxed or tense during the session?	[Range 1-7: very tense-very relaxed]
4.	Do you think telehealth improved your care?	[Range 1-7: not at all-very much]
5.	Do you think telehealth sessions were as good as regular in-person visits?	[Range 1-7: not as good-much better]
6.	How well did the app with telehealth equipment work overall?	[Range 1-7: very badly-very well]
7.	Would you want to use telehealth again?	[Range 1-7: not at all-very much]
8.	Overall, how satisfied were you with your physical therapy experience?	[Range 1-7: very unsatisfied-very satisfied]

HEP Adherence

Owing to the limited funding and recruitment setting, the recruitment results did not allow us to recruit sufficiently large sample sizes per specific diagnostic category to allow for separate subanalyses per diagnosis. Therefore, the experimental and control groups combined adults who all experienced financial distress but may have varied in terms of the diagnosis of chronic disease. The experimental group had an average HEP adherence of 4 (SD 2) times per week at 8 weeks to 5 (SD 2) times per week at 24 weeks ($d_{repmeas}$=0.48). The control group decreased on average from 4 (SD 2) times per week to 3 (SD 2) times per week. The within-group differences in HEP adherence were not significant (experimental group S=27.5, P=.14; control group S=−46.5, P=.07), the latter reflecting a small effect size $d_{repmeas}$=0.42. The between-group difference in HEP adherence at 24 weeks was statistically significant (U=2.49; P=.01), with a moderate effect size of Hedges g=0.76.

In total, 67% (12/18) and 82% (14/17) of participants in the experimental group at 8 and 24 weeks, respectively, were at least partially HEP adherent (ie, at least 4 times per week) compared with 70% (14/20) and 42% (8/19) at 8 and 24 weeks, respectively, in the control group. In the experimental group, 68% (15/22) of participants sent videos back, returning on average 68% of the requested number of videos (range 25%-100%). Participants did not report problems with Wi-Fi access at home; however, 43% (10/23) of the participants who used community-based Wi-Fi reported difficulty in returning videos because of occasional unreliable Wi-Fi. Participants who completed HEPs were 100% compliant with the requested number of trials per exercise and the requested number of exercises (verified through video recording).

Discussion

Overview

Given our patient pool of 45 adults experiencing financial distress, among which 38 had neurological chronic diseases and 7 had other types of chronic diseases described above, we explored whether the Hudl app with telehealth platform would be feasible in patients with physical and financial difficulties. If it proved to be feasible, we then asked if participants who were using the Hudl app with telehealth platform were able to maintain their HEP adherence over 24 weeks and whether the HEP adherence in that group would be higher than that in a control group that used HEP on paper. We specifically selected participants who were experiencing financial distress because our long-term goal is to provide a cost-effective solution to patients in a lower socioeconomic status to manage their health effectively, which, in turn, will lead to significant cost savings for the health care system. This is the first critical step for achieving such a goal.

Principal Findings

We obtained the following answers to whether the Hudl app with telehealth platform is feasible: (1) participants were able to use the app. (2) None of the participants reported problems with Wi-Fi at home; however, 43% (10/23) of the participants who used community-based Wi-Fi reported some difficulty returning videos because of occasional unreliable Wi-Fi. (3) Participants who used the app with telehealth platform maintained their HEP adherence over 24 weeks. This period of HEP adherence was long enough so that any initial excitement of the novelty should have worn off over time, revealing the real feasibility of the app/telehealth combination approach over this prolonged period. (4) There was a statistically significant difference in HEP adherence between both groups at 24 weeks because participants who used the app with telehealth maintained their HEP adherence whereas participants in the control group decreased in HEP adherence. These results suggested the overall feasibility of the app/telehealth combination approach.

In addition, as mentioned in the Introduction section, patient satisfaction is an important aspect of feasibility, linked to the success of treatment adherence, and is an important outcome for quality of health care [20-23]. Care and ease of interaction with the physical therapist have been shown to be rated of the highest importance for patient satisfaction [30]. The satisfaction with PT care dropped significantly in the control group from 8 to 24 weeks. In contrast, we observed that patient satisfaction was rated high for the group using the app and telerehabilitation, with one of the highest satisfactions being the *care of the telehealth-PT* through 8 to 24 weeks. We believe that the high satisfaction score was attributed to the specific feedback given through the various app tools, which were superimposed on the patient's self-recorded video.

In summary, this study indicated the feasibility of the combined app with telehealth approach in a specific group of adults who experienced financial hardship. The study showed promising results in terms of HEP adherence for symptoms of participants who had mostly neurological chronic diseases. These promising results justify the validation of this study in a larger cohort to demonstrate the cost-effectiveness and efficacy of the app with telehealth in terms of health outcomes in populations with specific chronic diseases; therefore, these app with telehealth platforms can be implemented more widely. Implementation of such strategies in health care would benefit patients because of reduced costs in involved expenses, including transportation. In addition, during pandemic periods with requirements for social distancing or quarantine for high-risk adults, our combined app/telehealth approach might offer a valuable alternative to monitor the status of patients who need long-term follow-up care [31].

Limitations

As mentioned earlier, our patient pool does not have a sufficient sample size and spectrum of chronic diseases to allow us to perform subanalyses per diagnosis group. Therefore, we cannot make any assumptions whether adults with one specific type of chronic disease would have better HEP adherence with the use of the Hudl app with telerehabilitation than another type. This issue should be addressed in future studies.

Comparison With Previous Work

A total of 3 other studies have reported HEP adherence with an app in adults with a frozen shoulder, diseases involving the musculoskeletal system, and Parkinson disease [5,10,14]. The studies in adults with a frozen shoulder or Parkinson disease used custom software for their apps [5,14], whereas the study with the musculoskeletal system used an app connected to free exercises on the web [10]. Participants either did not receive feedback at all [5], received only periodic phone calls or motivational text messages [10], or interacted only with a virtual coach for 5 minutes per exercise session [14]. In all the 3 studies, the results were reported over 3 to 4 weeks only, which poses a potential concern in terms of ascertaining whether participants would keep practicing for the required long-term duration that goes with managing chronic symptoms. Thus, it is possible that the excitement of the novelty of the app might have motivated the participants to follow the program during that time, resulting in relatively high HEP adherence. In contrast, we monitored HEP adherence for an extended period (ie, 24 weeks).

The HEP adherence in our experimental group was 67% (12/18) and 82% (14/17) partial HEP adherence at 8 and 24 weeks, respectively. Although a direct comparison is not possible because of the difference in sample size, the adherence percentages were higher than those reported in a recent meta-analysis in which 37% partial HEP adherence was reported in 11 studies, with a total of 1231 adults with chronic diseases [7]. These studies used traditional approaches of delivering HEP on paper, and their results were similar even when provided in different settings—in a center, off-site, or at home—and irrespective of whether they received phone follow-up at home [7].

Conclusions

The results from this study indicate that the use of the Hudl app with telehealth platform is feasible for participants experiencing financial distress and demonstrate that participants are able to maintain a high level of HEP adherence when using the app

compared with when HEP is delivered on paper. Our long-term goal is to provide a cost-effective solution to adults with lower socioeconomic status to manage their health more effectively in the long run in such a way that they can benefit from reimbursement through Medicaid or a similar health care coverage. This project is the first critical step for such a goal, with promising results justifying the planning of a larger sample clinical trial.

Acknowledgments

The authors would like to thank Dr Yuichiro Takagi for critically reading the manuscript. The data sets generated and/or analyzed during this study are available from the corresponding author upon reasonable request. This work was supported by the Medtronic Philanthropy (Grant No. FY15-00261) to explore novel ways in *Expanding Access to Physical Therapy for Underserved Patients With Chronic Disease*.

Authors' Contributions

AV and CB significantly contributed to the concept of the work. CB was responsible for the random allocation sequence, enrollment of participants, and assignment of participants to interventions. CB oversaw the project and acquired the data, and AV performed the data analysis and interpretation. AV wrote the manuscript. CB and TN provided critical revisions. All authors approved the final version and agreed to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Percentages of satisfaction scores (scores 5-7) using the Hudl Technique coaching app in terms of its feasibility.
[DOCX File , 17 KB - formative_v5i3e22659_app1.docx]

Multimedia Appendix 2
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1742 KB - formative_v5i3e22659_app2.pdf]

References

1. About Chronic Diseases. National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP). 2019 Apr 26. URL: https://www.cdc.gov/chronicdisease/about/index.htm [accessed 2020-03-21]
2. Ward BW, Schiller JS, Goodman RA. Multiple chronic conditions among US adults: a 2012 update. Prev Chronic Dis 2014 Apr 17;11:E62 [FREE Full text] [doi: 10.5888/pcd11.130389] [Medline: 24742395]
3. Benjamin RM. Multiple chronic conditions: a public health challenge. Public Health Rep 2010;125(5):626-627 [FREE Full text] [Medline: 20873276]
4. Pentecost C, Taket A. Understanding exercise uptake and adherence for people with chronic conditions: a new model demonstrating the importance of exercise identity, benefits of attending and support. Health Educ Res 2011 Oct;26(5):908-922. [doi: 10.1093/her/cyr052] [Medline: 21734226]
5. Stütz T, Emsenhuber G, Huber D, Domhardt M, Tiefengrabner M, Oostingh GJ, et al. Mobile phone-supported physiotherapy for frozen shoulder: feasibility assessment based on a usability study. JMIR Rehabil Assist Technol 2017 Jul 20;4(2):e6 [FREE Full text] [doi: 10.2196/rehab.7085] [Medline: 28729234]
6. Beinart NA, Goodchild CE, Weinman JA, Ayis S, Godfrey EL. Individual and intervention-related factors associated with adherence to home exercise in chronic low back pain: a systematic review. Spine J 2013 Dec;13(12):1940-1950. [doi: 10.1016/j.spinee.2013.08.027] [Medline: 24169445]
7. Jansons PS, Haines TP, O'Brien L. Interventions to achieve ongoing exercise adherence for adults with chronic health conditions who have completed a supervised exercise program: systematic review and meta-analysis. Clin Rehabil 2017 Apr;31(4):465-477. [doi: 10.1177/0269215516653995] [Medline: 27301799]
8. Faber M, Andersen MH, Sevel C, Thorborg K, Bandholm T, Rathleff M. The majority are not performing home-exercises correctly two weeks after their initial instruction-an assessor-blinded study. Peer J 2015 Jul;3:e1102 [FREE Full text] [doi: 10.7717/peerj.1102] [Medline: 26244112]
9. Picorelli AM, Pereira LSM, Pereira DS, Felício D, Sherrington C. Adherence to exercise programs for older people is influenced by program characteristics and personal factors: a systematic review. J Physiother 2014 Sep;60(3):151-156 [FREE Full text] [doi: 10.1016/j.jphys.2014.06.012] [Medline: 25092418]
10. Lambert TE, Harvey LA, Avdalis C, Chen LW, Jeyalingam S, Pratt CA, et al. An app with remote support achieves better adherence to home exercise programs than paper handouts in people with musculoskeletal conditions: a randomised trial. J Physiother 2017 Jul;63(3):161-167 [FREE Full text] [doi: 10.1016/j.jphys.2017.05.015] [Medline: 28662834]

11. Blanquero J, Cortés-Vega MD, García-Frasquet M, Sánchez-Laulhé PR, Nieto Díaz de Los Bernardos MI, Suero-Pineda A. Exercises using a touchscreen tablet application improved functional ability more than an exercise program prescribed on paper in people after surgical carpal tunnel release: a randomised trial. J Physiother 2019 Apr;65(2):81-87 [FREE Full text] [doi: 10.1016/j.jphys.2019.02.008] [Medline: 30926400]
12. 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;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
13. Thiamwong L, Stout JR, Sole ML, Ng BP, Yan X, Talbert S. Physio-Feedback and Exercise Program (PEER) improves balance, muscle strength, and fall risk in older adults. Res Gerontol Nurs 2020 Nov 01;13(6):289-296. [doi: 10.3928/19404921-20200324-01] [Medline: 32286669]
14. Ellis T, Latham NK, DeAngelis TR, Thomas CA, Saint-Hilaire M, Bickmore TW. Feasibility of a virtual exercise coach to promote walking in community-dwelling persons with Parkinson disease. Am J Phys Med Rehabil 2013 Jun;92(6):472-81; quiz 482 [FREE Full text] [doi: 10.1097/PHM.0b013e31828cd466] [Medline: 23552335]
15. Quintero LM, Moore JW, Yeager MG, Rowsey K, Olmi DJ, Britton-Slater J, et al. Reducing risk of head injury in youth soccer: an extension of behavioral skills training for heading. J Appl Behav Anal 2020 Jan;53(1):237-248. [doi: 10.1002/jaba.557] [Medline: 30924148]
16. 2021 Minnesota Department of Human Services. Telemedicine. MN Department of Human Services website. 2019. URL: https://www.dhs.state.mn.us/main/idcplg?IdcService=GET_DYNAMIC_CONVERSION&RevisionSelectionMethod=LatestReleased&dDocName=DHS-296422 [accessed 2019-07-27]
17. Program SR. Center for Connected Health Policy website. 2020 Sep. URL: https://www.cchpca.org/telehealth-policy/current-state-laws-and-reimbursement-policies/minnesota-medicaid-store-and-forward-policy [accessed 2018-02-08]
18. Weiler R. Übersense Coach app for sport medicine? Slow motion video analysis (Mobile App User Guide). Br J Sports Med 2016 Feb;50(4):255-256. [doi: 10.1136/bjsports-2015-094876] [Medline: 26130696]
19. Neal BS, Lack SD, Barton CJ, Birn-Jeffery A, Miller S, Morrissey D. Is markerless, smart phone recorded two-dimensional video a clinically useful measure of relevant lower limb kinematics in runners with patellofemoral pain? A validity and reliability study. Phys Ther Sport 2020 May;43:36-42. [doi: 10.1016/j.ptsp.2020.02.004] [Medline: 32066107]
20. Bush T, Cherkin D, Barlow W. The impact of physician attitudes on patient satisfaction with care for low back pain. Arch Fam Med 1993 Mar;2(3):301-305. [doi: 10.1001/archfami.2.3.301] [Medline: 8252151]
21. Hudak PL, Wright JG. The characteristics of patient satisfaction measures. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3167-3177. [doi: 10.1097/00007632-200012150-00012] [Medline: 11124733]
22. Keith RA. Patient satisfaction and rehabilitation services. Arch Phys Med Rehabil 1998 Sep;79(9):1122-1128. [doi: 10.1016/s0003-9993(98)90182-4] [Medline: 9749695]
23. Nitse PS, Rushing V. Patient satisfaction: the new area of focus for the physician's office. Health Mark Q 1996;14(2):73-84. [doi: 10.1300/j026v14n02_07] [Medline: 10164449]
24. Silveira P, van de Langenberg R, van Het Reve E, Daniel F, Casati F, de Bruin ED. Tablet-based strength-balance training to motivate and improve adherence to exercise in independently living older people: a phase II preclinical exploratory trial. J Med Internet Res 2013 Aug 12;15(8):e159 [FREE Full text] [doi: 10.2196/jmir.2579] [Medline: 23939401]
25. Stangroom J. Chi Square Calculator. Social Science Statistics. 2021. URL: https://www.socscistatistics.com/tests/chisquare/default2.aspx [accessed 2020-05-03]
26. Motulsky H. McNemar Test Calculator. GraphPad. 2018. URL: https://www.graphpad.com/quickcalcs/McNemar1.cfm [accessed 2020-05-03]
27. Hedges LV, Okin I. Statistical Methods for Meta-Analysis. San Diego, CA: Academic Press; 1985.
28. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. Psychol Methods 2002 Mar;7(1):105-125. [doi: 10.1037/1082-989x.7.1.105] [Medline: 11928886]
29. Lenhard W, Lenhard A. Calculation of effect sizes. Psychometrica 2016 May 20 [FREE Full text] [doi: 10.13140/RG.2.2.17823.92329]
30. Beattie PF, Pinto MB, Nelson MK, Nelson R. Patient satisfaction with outpatient physical therapy: instrument validation. Phys Ther 2002 Jun;82(6):557-565. [Medline: 12036397]
31. Shanthanna H, Strand NH, Provenzano DA, Lobo CA, Eldabe S, Bhatia A, et al. Caring for patients with pain during the COVID-19 pandemic: consensus recommendations from an international expert panel. Anaesthesia 2020 Jul;75(7):935-944 [FREE Full text] [doi: 10.1111/anae.15076] [Medline: 32259288]

Abbreviations

HEP: home exercise program
PT: physical therapy

Edited by G Eysenbach, R Kukafka; submitted 20.07.20; peer-reviewed by J Li, N Mohammad Gholi Mezerji; comments to author 11.08.20; revised version received 25.09.20; accepted 27.02.21; published 18.03.21.

Please cite as:
Van de Winckel A, Nawshin T, Byron C
Combining a Hudl App With Telehealth to Increase Home Exercise Program Adherence in People With Chronic Diseases Experiencing Financial Distress: Randomized Controlled Trial
JMIR Form Res 2021;5(3):e22659
URL: https://formative.jmir.org/2021/3/e22659
doi:10.2196/22659
PMID:33640865

Original Paper

eHealth Program to Reduce Hospitalizations Due to Acute Exacerbation of Chronic Obstructive Pulmonary Disease: Retrospective Study

Amanda R van Buul[1], MD, PhD; Caroline Derksen[2]; Ouke Hoedemaker[2], MBA; Oscar van Dijk[3], IR, MBA; Niels H Chavannes[4,5], MD, Prof Dr; Marise J Kasteleyn[1,4,5], PhD

[1]Department of Pulmonology, Leiden University Medical Center, Leiden, Netherlands

[2]Bravis, Department of Pulmonology, Roosendaal, Netherlands

[3]Medicine Men, Utrecht, Netherlands

[4]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, Netherlands

[5]National eHealth Living Lab, Leiden, Netherlands

Corresponding Author:
Amanda R van Buul, MD, PhD
Department of Pulmonology
Leiden University Medical Center
Leiden
Netherlands
Phone: 31 715297550
Email: a.r.van_buul@lumc.nl

Abstract

Background: Hospitalization for acute exacerbation of chronic obstructive pulmonary disease (COPD) is associated with poor prognosis. eHealth interventions might improve outcomes and decrease costs.

Objective: This study aimed to evaluate the effect of an eHealth program on COPD hospitalizations and exacerbations.

Methods: This was a real-world study conducted from April 2018 to December 2019 in the Bravis Hospital, the Netherlands. An eHealth program (EmmaCOPD) was offered to COPD patients at risk of exacerbations. EmmaCOPD consisted of an app that used questionnaires (to monitor symptoms) and a step counter (to monitor the number of steps) to detect exacerbations. Patients and their buddies received feedback when their symptoms worsened or the number of steps declined. Generalized estimating equations were used to compare the number of days admitted to the hospital and the total number of exacerbations 12 months before and (max) 18 months after the start of EmmaCOPD. We additionally adjusted for the potential confounders of age, sex, COPD severity, and inhaled corticosteroid use.

Results: The 29 included patients had a mean forced expiratory volume in 1 second of 45.5 (SD 17.7) %predicted. In the year before the intervention, the median total number of exacerbations was 2.0 (IQR 2.0-3.0). The median number of hospitalized days was 8.0 days (IQR 6.0-16.5 days). Afterwards, there was a median 1.0 (IQR 0.0-2.0) exacerbation and 2.0 days (IQR 0.0-4.0 days) of hospitalization. After initiation of EmmaCOPD, both the number of hospitalized days and total number of exacerbations decreased significantly (incidence rate ratio 0.209, 95% CI 0.116-0.382; incidence rate ratio 0.310, 95% CI 0.219-0.438). Adjustment for confounders did not affect the results.

Conclusions: The eHealth program seems to reduce the number of total exacerbations and number of days of hospitalization due to exacerbations of COPD.

(JMIR Form Res 2021;5(3):e24726) doi:10.2196/24726

KEYWORDS

COPD; eHealth; exacerbations; hospitalizations; mHealth

Introduction

Chronic obstructive pulmonary disease (COPD) is a treatable, preventable, chronic lung disease that accounts for years lived with disability [1] and reduced life expectancy [2]. The prognosis of COPD depends on multiple factors [3]. From previous research, it is known that hospitalization for an acute exacerbation of COPD is associated with poor prognosis and increased risk of death [4]. A substantial proportion of patients dies within 1 year after being discharged from their first hospitalization for an exacerbation of COPD [5]. Patients with COPD Global Initiative for Chronic Obstructive Lung Disease (GOLD) [6] stages 3 or 4 (severe and very severe airway obstruction, respectively) have the highest risk for an exacerbation, although patients with COPD GOLD 2 (moderate airway obstruction) are also at risk [4]. The costs of COPD rise with increasing severity of exacerbations, with hospital admissions accounting for most of these costs [7]. Tools to prevent or shorten hospital admissions are necessary to slow down COPD progression and to limit health care costs.

eHealth interventions are promising for improving outcomes and decreasing costs in chronic diseases, including COPD [8,9]. Different types of eHealth interventions for COPD exist, ranging from apps to support self-management to telemonitoring programs in which patients are followed extensively [10]. Diverse outcomes in various settings with a variety of eHealth interventions have been studied. Previous studies have shown that eHealth could decrease exacerbations and hospital admissions in COPD patients [11,12]. Of those studies, one offered patients who were discharged from the hospital (admission due to an exacerbation of COPD) an intervention that included a comprehensive assessment at discharge, education, an individually tailored care plan, weekly phone calls, and access to a specialized nurse at the hospital through an online platform. The intervention resulted in a reduction in hospital admissions [12]. Another study included patients with COPD GOLD stages 3 and 4 who were seen by a pulmonologist. Patients were monitored via home-based telemonitoring that consisted of a device with a large screen and 4 buttons that patients used to fill out a daily questionnaire. Patients received feedback from their device, and the responses were also sent to a secure data center. The responses were categorized and prioritized, and respiratory nurses contacted the patients if values were alarming. After 6 months, there was a decrease in hospital admissions and exacerbations, and there was a tendency toward decreased number of days in the hospital and outpatient visits [13].

It is thought that patients with frequent exacerbations may benefit more from eHealth programs [14-18]. Despite the promising results from previous studies, no eHealth programs were incorporated in the latest COPD statement [6]. Based on previous research, for the current study, we hypothesized that giving patients the responsibility to act on signs of a COPD exacerbation and make them aware of changes in COPD symptoms and physical activity will influence self-management, which can lead to a reduction of exacerbations and hospitalizations. This was incorporated in the EmmaCOPD eHealth program. A new item in the intervention was the involvement of informal care givers (“buddies”). The Bravis Hospital (Roosendaal, The Netherlands) offered COPD patients who are at risk of exacerbations the possibility to use this EmmaCOPD program, consisting of an app that includes questionnaires and an activity coach. This program was designed to recognize signs of a COPD exacerbation and inform patients and buddies when symptoms worsened or the number of steps per day declined.The primary aim of this study was to determine the effect of EmmaCOPD on the number of days of hospitalization. The secondary outcome of this study was to assess the effect of this program on the number of total exacerbations.

Methods

Study Design

This was a retrospective study with a pre-post research design using real-world data that were retrieved from the electronic record system at the Bravis Hospital and from EmmaCOPD. Data were collected between April 2018 and December 2019 from patients who agreed to participate in EmmaCOPD. Analyses were performed between January 2020 and March 2020. Due to the retrospective nature of the study and the fact that this study does not fall under the Medical Research Involving Human Subjects Act (in Dutch, Wet medisch - wetenschappelijk onderzoek met mensen [WMO]), there was no need for ethical approval. Patients were aware that this intervention was new in clinical practice. All patients signed informed consent to use their data for research.

Study Population

Patients could be included if they were treated by a pulmonologist in the Bravis Hospital. Patients were eligible for the intervention if they had COPD and if they had at least 2 exacerbations of COPD in the previous 12 months. An exacerbation was defined as an increase in symptoms that was more than day-to-day variation combined with prescription of a course of oral corticosteroids or antibiotics. Patients could also be included if they were at increased risk of exacerbations according to their health care provider.

Patients were excluded if they used EmmaCOPD before April 2018, because these patients could already have experienced the beneficial effects of the intervention. Furthermore, patients were excluded if they did not own an Android-based smartphone since the app for the Activity coach was only compatible with Android-based smartphones.

EmmaCOPD Intervention

Starting in November 2016, the Bravis Hospital (Roosendaal, The Netherlands) offered patients with COPD who are at risk for hospitalization due to an exacerbation a new eHealth program: EmmaCOPD, an app [19] that was designed to recognize signs of an exacerbation of COPD. EmmaCOPD was developed by Medicine Men BV (Utrecht, The Netherlands), with input from patients with COPD and physicians. The app used questionnaires (to monitor symptoms) and a step counter (to monitor the number of steps) to detect exacerbations (Figure 1 and Figure 2). Patients received feedback when their symptoms worsened or the number of steps declined. A “buddy”

received this information too. A buddy was an informal caregiver who was close to the patient; this could be a relative, good friend, or neighbor. Health care professionals could be contacted by the patient or buddy if the app advised them to do so or if the patient or buddy was worried, but health care professionals were not involved in the fast response (Figure 3). Health care providers had access to the Emma account, and the account could be checked if it was needed.

Figure 1. Questionnaire app to monitor symptoms.

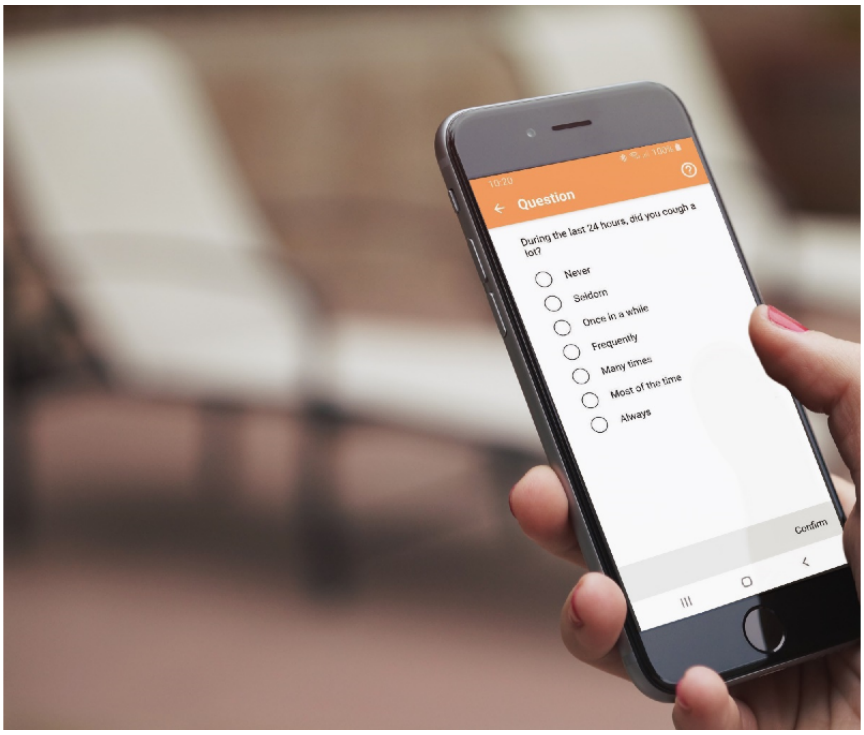
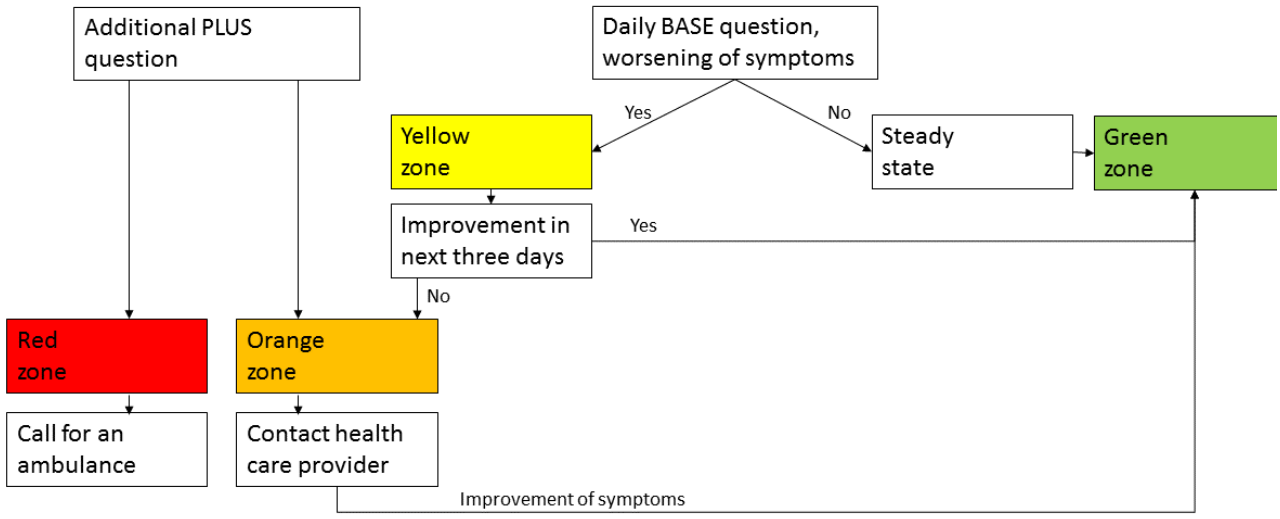


Figure 2. Smartwatch with built-in step counter.

Figure 3. Flow chart with zones, based on questions.



In the program, patients could be in several zones: green, yellow, orange, or red. At the start of app use, patients were in the "green zone" (steady state). Every day, patients filled out the BASE questionnaire of the COPD action plan [20]. In the BASE questionnaire, patients reported whether they experienced worsening of COPD symptoms (Figure 3). If symptoms worsened (eg, shortness of breath; viscous sputum; sputum color changed to green, brown, or grey; wheezing; cough; fatigue; activities are difficult due to symptoms; headache; dizziness while awake or concentration problems), patients entered a so-called "yellow zone," and they were advised to read their individual exacerbation plan and adapt their medication accordingly. If there was no improvement in 3 days, patients entered the "orange zone," and the app advised them to either take the emergency medication or to contact a health care professional. When there was an improvement, the patient went back into the "green zone." At any time, the buddy received a signal (SMS, email, app signal, dashboard signal, or a combination) when the status changed to another color. If the patient or buddy suspected something serious, they could fill out (at any time) a second question [20] (PLUS question, see Figure 3). In response to some answers on the PLUS question (eg, hemoptoe, fever or too sick to do activities), the patient entered the "orange zone" (see description earlier in the paragraph). For answers indicating a potentially life-threatening situation (eg, very dyspneic, chest pain, confused, forgetfulness, dizziness, tendency to collapse or loss of consciousness), the patients entered the "red zone," and the buddy was advised to call for an ambulance immediately. Furthermore, patients completed the Clinical COPD Questionnaire (CCQ) [21] 3 times on a weekly basis to give insight into their COPD-related health status.

The Pebble Time, a modern programmable smartwatch that includes an accelerometer and gyrometer (Bosch Sensortec BMI160, BOSCH, Germany) was used to signal a decline in the number of steps. During the first 3 weeks, baseline activity level was assessed. Thereafter, a physiotherapist set a step goal. When patients reached this goal, patients were in the "green zone." If there was a decline of 20% in the number of steps, patients received a signal that they entered the "orange zone"; if there was a decline of 40%, patients were in the "red zone."

During an onboarding session in the Bravis Hospital, patients and buddies were prepared for the use of EmmaCOPD. Their individual exacerbation plan was checked, and an Emma account was created. Furthermore, the Emma questionnaire app and the Emma activity coaching companion app were installed on the patient's smartphone. The buddy, health care providers, physiotherapists, and the support department of Medicine Men had access to the account.

Data Collection

Baseline characteristics of participants were collected from the electronic health records at the Bravis Hospital. At baseline and follow-up, exacerbations were collected from the electronic records. At baseline, the number of exacerbations in the previous 12 months was collected. A mild exacerbation was defined as a "flare up of COPD symptoms with a change in COPD medication," a moderate exacerbation as a "flare up of COPD symptoms that requires prescription of a course of corticosteroids or antibiotics," and a severe exacerbation as a "flare up of COPD symptoms that led to hospital admission." At baseline and follow-up, the number of hospitalizations and the number of days admitted to the hospital were collected. GOLD category (A, B, C, and D) at baseline was determined using the number of exacerbations and modified Medical Research Council score [6]. Data on how often patients filled out the questionnaires, results of the CCQ, and how often patients entered the orange and red zones were collected from EmmaCOPD.

Power Calculation

From previous research, it is known that of all patients with COPD GOLD stage 2, 7% was admitted to the hospital. Of all patients with COPD GOLD stage 3, 18% was admitted. Of all patients with COPD GOLD 4, 33% was admitted [22]. Since the target population was treated by secondary care pulmonology, we estimated that 25% of patients were admitted

to the hospital within 1 year. The mean number of days admitted to the Bravis Hospital due to exacerbation of COPD was 6.0 days. When exploring this number of admission days for all patients, the mean number of days admitted to the hospital within 1 year was 1.5 days per person per year. In 2014, one of the goals for the Long Alliantie Nederland was a 25% reduction in the number of days admitted due to exacerbation of COPD [23]. The new number of admission days was calculated to be 1.125 days per person per year. With an alpha of 0.05, power of 80%, mean of 1.5 days (SD 0.75 days) preintervention, mean of 1.125 days (SD 0.56 days) postintervention, correlation of 0.4, and drop-out rate of 20%, 40 patients were needed with a follow-up period of 1 year. We expected that we needed 8 months to include the patients.

Statistical Analysis

Descriptive statistics are presented as mean (SD) for continuous variables with a normal distribution, median (IQR) for continuous variables without a normal distribution, and percentages for categorical variables. We used a pre-post research design. For the postintervention period, the follow-up duration was calculated as the number of days between the date of data extraction and the date of inclusion. If a patient died, the follow-up duration was calculated as the number of days between the date of death and date of inclusion. We compared the postintervention period with a preintervention period of 365 days before inclusion in the study. We analyzed the difference between the first CCQ score and the last CCQ score using descriptive statistics. Generalized estimating equations were used to analyze CCQ change over time. For the analysis of the effect of the intervention on the number of hospital admissions and the number of hospital admission days, which can be conceptualized as count data with repeated intra-individual measurements before and after initiation of EmmaCOPD, we used generalized estimating equation models. The distribution of the data was tested first to check whether the data fitted best with a Poisson distribution or negative binomial model. Outcomes are expressed as the incidence rate ratio (IRR). As explanatory variables, we used the length of follow-up (log-transformed) and intervention (coded as 0 or 1 for the preintervention and postintervention periods, respectively). Additionally, we adjusted for potential confounders of sex and age (model 2). In a final model, we additionally adjusted for baseline severity expressed as GOLD category (model 3) and inhaled corticosteroid use (model 4).

All analyses were conducted in SPSS version 25.0.

Results

Baseline characteristics are presented in Table 1; 29 patients were included with a mean age of 67.4 years (SD 8.0 years), mean forced expiratory volume in 1 second % predicted of 45.5 (SD 17.7). In the 12 months before baseline, patients had a median of 2.0 (IQR 2.0-3.0) severe exacerbations and were admitted to the hospital for a median of 8.0 days (IQR 6.0-16.5 days).

Table 1. Baseline characteristics of patients with chronic obstructive pulmonary disease (COPD; n=29).

Patient characteristics	Results
Age (years), mean (SD)	67.4 (8.0)
Sex (women), n (%)	13 (45)
Ethnicity (Caucasian), n (%)	26 (90)
BMI (kg/m^2), mean (SD)	27.3 (5.0)
Comorbidity (CCI[a] score), mean (SD)	2.1 (1.2)
Asthma (yes), n (%)	3 (10)
Smoking status, n (%)	
Current	3 (10)
Ex-smoker	25 (86)
Never	1 (3)
Pulmonary medication, n (%)	
ICS[b] mono (yes)	2 (7)
LABA[c] mono (yes)	6 (21)
LAMA[d] mono (yes)	11 (38)
ICS/LABA in one device (yes)	14 (48)
LABA/LAMA in one device (yes)	3 (10)
Oral corticosteroids (yes)	13 (45)
Physical activity (steps a day), median (IQR)[e]	2482.5 (1394.3-4184.3)
Pulmonary function, mean (SD)[e]	
FEV_1[f] (L)[e]	1.3 (0.6)
FEV_1 (% predicted)[e]	45.5 (17.7)
FVC[g] (L)[e]	2.9 (0.8)
FVC (% predicted)[e]	82.0 (16.5)
FEV_1/FVC[e]	41.2 (14.3)
Exacerbations	
Number of mild exacerbations[h] in previous 12 months, median (IQR)	0.0 (0.0-0.0)
Number of moderate exacerbations[i] in previous 12 months, median (IQR)	0.0 (0.0-0.0)
Number of severe exacerbations[j] in previous 12 months, median (IQR)	2.0 (2.0-3.0)
Total number of exacerbations in the previous 12 months, median (IQR)	2.0 (2.0-3.0)
Number of patients with ≥2 exacerbations, n (%)	25 (86)
COPD-related symptom scores, mean (SD)	
mMRC[k] score[e]	3.0 (1.1)
CCQ[l] score[e]	3.0 (1.2)
GOLD[m]-stage, n (%)[n]	
A[o]	0 (0)
B[p]	1 (3)
C[q]	1 (3)
D[r]	21 (72)

Patient characteristics	Results
Days admitted to the hospital due to COPD exacerbations, median (IQR)	8.0 (6.0-16.5)

[a]CCI: Charlson comorbidity index [24].
[b]ICS: inhaled corticosteroids.
[c]LABA: long-acting beta2 agonist.
[d]LAMA: long-acting muscarinic antagonist.
[e]n=26, data missing for 3 participants.
[f]FEV_1: forced expiratory volume in 1 second.
[g]FVC: forced vital capacity.
[h]Mild exacerbation: change in COPD medication.
[i]Moderate exacerbation: course of corticosteroids and/or antibiotics.
[j]Severe exacerbation: hospital admission.
[k]mMRC: modified Medical Research Council.
[l]CCQ: clinical COPD questionnaire.
[m]GOLD: Global Initiative for Chronic Obstructive Lung Disease.
[n]n=23, data missing for 6 participants.
[o]A: low symptoms, low risk for exacerbation.
[p]B: high symptoms, low risk for exacerbation.
[q]C: low symptoms, high risk for exacerbation.
[r]D: high symptoms, high risk for exacerbation.

Outcomes at follow-up are found in Table 2. The median follow-up was 587.0 days (IQR 372-594 days), and 3 (3/29, 10%) patients died. A follow-up duration of at least 12 months was achieved by 23 of the 29 patients. The median numbers of mild, moderate, and severe exacerbations were 0.0 (IQR 0.0-0.0), 0.0 (IQR 0.0-0.0), and 1.0 (IQR 0.0-2.0), respectively. The median number of days admitted to the hospital was 2.0 days (IQR 0.0-4.0 days), with a maximum of 15.0 days. The median difference between the first and last CCQ scores was 0.3 points (IQR –0.4 to 0.9). The CCQ change over time was not statistically significant (*P*=.860).

The data for both the number of hospitalization days and total number of exacerbations fitted best within a Poisson distribution. Unadjusted analyses showed that, after initiation of the EmmaCOPD intervention, both the number of hospitalization days (IRR 0.210, 95% CI 0.116-0.382) and the total number of exacerbations (IRR 0.310, 95% CI 0.219-0.438) decreased significantly (Table 3). Analyses adjusted for age and sex showed comparable results, with a significant decrease in hospitalization days (IRR 0.209, 95% CI 0.114-0.382) and total number of exacerbations (IRR 0.310, 95% CI 0.217-0.435). Additional adjustment for GOLD category and inhaled corticosteroid use showed comparable results (Table 3).

Table 2. Follow-up at 12-18 months after initiation of EmmaCOPD (n=29).

Outcomes	Results
Follow-up duration (days), median (IQR)	587.0 (372.0-594.0)
Mortality (yes), n (%)	3 (10)
Number of mild exacerbations[a], median (IQR)	0.0 (0.0 to 0.0)
Number of moderate exacerbations[b], median (IQR)	0.0 (0.0 to 0.0)
Number of severe exacerbations[c], median (IQR)	1.0 (0.0 to 2.0)
Total number of exacerbations, median (IQR)	1.0 (0.0 to 2.0)
Hospital admission (days), median (IQR)	2.0 (0.0 to 4.0)
Change in CCQ[d], median (IQR)[e]	0.3 (–0.4 to 0.9)

[a]Mild exacerbation: change in COPD medication.
[b]Moderate exacerbation: course of corticosteroids and/or antibiotics.
[c]Severe exacerbation: hospital admission.
[d]CCQ: clinical chronic obstructive pulmonary disease questionnaire.
[e]n=28, data missing for 1 participant.

Table 3. Effect of EmmaCOPD on length of hospitalization and number of exacerbations, compared between 365 days before the initiation of EmmaCOPD and 12-18 months after the initiation of EmmaCOPD.

Analytic model	Hospitalization (days), IRR[a] (95% CI)	Total number of exacerbations, IRR (95% CI)
Crude analysis (model 1)	0.210 (0.116-0.382)	0.310 (0.219-0.438)
Adjusted analysis (model 2)[b]	0.209 (0.114-0.382)	0.308 (0.217-0.435)
Adjusted analysis (model 3)[c]	0.225 (0.111-0.456)	0.327 (0.211-0.506)
Adjusted analysis (model 4)[d]	0.225 (0.111-0.456)	0.325 (0.208-0.508)

[a]IRR: incidence rate ratio.

[b]Adjusted for sex and age.

[c]Model adjusted for sex, age, and Global Iniative for Chronic Obstructive Lung Disease (GOLD) stage (patients with missing GOLD stage were exluded).

[d]Model adjusted for sex, age, GOLD stage (patients with missing GOLD stage were excluded), and inhaled corticsteroid use.

Data derived from EmmaCOPD are presented in Table 4. During the follow-up, the median number of daily BASE questions answered was 252.0 (IQR 125.0-423.0). The median number of answers on the BASE questions in the "yellow zone" (worsening of symptoms) was 26.0 (IQR 7.0-91.0), with a range of 0-527. Of the median 13.0 (IQR 5.0-68.0) PLUS questions answered, a median of 1.0 (IQR 0.0-4.0) was answered in the "orange zone" and 1.0 (IQR 0.0-3.0) in the "red zone." The median numbers of days in the "orange zone" and "red zone" were 0.0 (IQR 0.0-25.0) and 3.0 (IQR 2.0-3.0), respectively. The median number of steps a day was 1710.0 (IQR 1144.0-3078.0).

Table 4. EmmaCOPD outcomes (n=29), with results categorized in zones (green, yellow, orange, or red), with each zone except green (steady state) requiring a different action.

Variables	Results
BASE questions[a]	
Number of BASE question answered, median (IQR)	252.0 (125.0-453.0)
Number of BASE questions answered as yes (yellow zone[b]), median (IQR)	26.0 (7.0-91.0)
PLUS questions[c]	
Number of PLUS questions answered, median (IQR)	13.0 (5.0-68.0)
Number of PLUS questions answered with an answer in the orange zone[d], median (IQR)	1.0 (0.0-4.0)
Number of PLUS questions answered with an answer in the red zone[e], median (IQR)	1.0 (0.0-3.0)
Zones	
Number of days in the orange zone[d], median (IQR)	0.0 (0.0-25.0)
Number of days in the red zone[e], median (IQR)	3.0 (2.0-3.0)
Physical activity (steps per day), median (IQR)[f]	1710.0 (1144.0-3078.0)

[a]BASE question: daily question about worsening of symptoms;

[b]Yellow zone: patient experienced worsening of symptoms (BASE question yes) and given advice to adjust medication.

[c]PLUS question: additional question when patient or the patient's buddy suspected something serious.

[d]Orange zone: no improvements in 3 days (yes answer to the BASE for 3 days) or an orange-rated answer to a PLUS question, for which the patient is given advice to take emergency medication or contact health care provider.

[e]Red zone: red answer on the PLUS question, potentially life-threatening clinical situation, buddy was advised to call an ambulance.

[f]n=27, data missing for 2 participants.

Discussion

Principal Findings and Comparison With Prior Work

The aim of the present study was to evaluate the effect of a new eHealth program (EmmaCOPD) on the number of hospitalized days and the total number of exacerbations in patients with COPD who are at risk for hospitalization. The present study, using real-world data, showed a significant decrease in the number of exacerbations and the number of days admitted to the hospital.

In line with the results of the present study, a Cochrane review [17] and a recent review [10] have shown that eHealth care programs and patient platforms were effective in reducing hospital admissions in COPD. Mostly, the effect of telemonitoring was studied. A previous study used home-based telemonitoring to monitor patients. The home-based

telemonitoring consisted of a device with a large screen and 4 buttons that patients used to fill out a daily questionnaire. Patients received feedback from their device, and the responses were also sent to a secure data center. The responses were categorized and prioritized, and respiratory nurses contacted the patients if values were alarming. After 6 months, there was a decrease in hospital admissions and exacerbations, and there was a tendency toward decreased number of days in the hospitals and outpatient visits [13]. Another study examined an intervention for patients that were discharged from the hospital. The intervention included a comprehensive assessment, an educational session, an individually tailored care plan, weekly phone calls, and access to a specialized nurse at the hospital through a digital platform. The intervention resulted in a reduced number of hospital admissions [12] and an increase in BMI [11], but there was no difference in dyspnea, lung function, and quality of life [11]. Other studies also reported no effect of eHealth on dyspnea and quality of life. One study evaluated internet-based dyspnea self-management support that included education, skills training, and coaching and found improvement in arm endurance exercises, but no differences in dyspnea and quality of life [11]. In line with the mentioned studies, this study showed a decrease in the number of total exacerbations. Patients were also monitored via an app and smartwatch, and they received feedback. These are elements that were found to be missing in previous apps, as found in a previous review [25]. Furthermore, a new element in the present intervention was the involvement of buddies who received alerts when the clinical situation changed. This could have resulted in a decline in anxiety due to frequent checks and involvement of buddies. However, we did not assess anxiety in this study and cannot test this hypothesis. Another element of the intervention that could have resulted in the reduction of admission days included improvement of self-management (fast and adequate medication adaptation according to the individual exacerbation plan, feedback, and maintenance of physical activity).

In the Netherlands, several eHealth projects aiming to prevent hospital admissions have been initiated. One project is focusing on education of health care providers and patients for early detection of mild and moderate exacerbations to prevent severe exacerbations [26]. Another project is using the Assessment of Burden of COPD tool [27] that was filled out during each visit to the health care provider to give insight into the burden of COPD and to increase quality of life and the quality of perceived care. The (preliminary) results of these studies show a decrease in the number of days between exacerbation onset and recognition [26], between recognition and action [26], and between cognition and general practitioner visits [26] as well as improvement in quality of life and perceived quality of care [27].

Strengths and Limitations

A strength of this study was the use of real-world data. EmmaCOPD was implemented as part of usual care in the Bravis Hospital, not as part of a study. Therefore, the risk of bias associated with participating in a study was minimized. Subgroup analysis in a systematic review has shown that telemonitoring is effective in patients with COPD, as are interventions that last more than 6 months [15]. In this study, the intervention was integrated into usual care, and there was no end date. The patients could have benefited from these characteristics. Another strength of the study was how self-management was organized. To minimize the risk of respiratory-related mortality that was reported in previous self-management interventions [28], patients were clearly instructed how to react when symptoms worsened, and there was a second question built in the app when patients or buddies suspected something serious. Furthermore, buddies received alerts when the clinical situation changed.

This study has limitations. First, there were fewer patients included in the study than were calculated in the power calculation. The number of patients included in the study could probably have been higher if the app had also been available for iPhones. Also, eHealth interventions often face implementation challenges, including costs, that might explain why the sample size was not met [29]. The number of patients that were eligible for EmmaCOPD but were not willing to use EmmaCOPD is unknown. A previous study showed that 15.9% of patients reject eHealth when it is offered [25]. Still, there was a significant difference in the number of days of hospitalizations. More patients with more follow-up data could have resulted in a more precise difference in the number of hospitalization days. A second limitation is selection bias; patients with the strongest motivation will accept such an intervention program, while those who are not motivated to improve their COPD condition will refuse the intervention. From pulmonary rehabilitation, it is known that the most frequently mentioned reason for refusal is lack of interest [30]. However, not all patients were motivated to fill out the questionnaire on a daily basis; during the median study period of 587.0 days, the BASE questions were answered on a median of 252.0 days. It could also be true that patients felt too sick to answer the questions. This could have resulted in an underestimation of the number of days in the yellow, orange, and red zones. Furthermore, it is known that interest in eHealth declines over time [31]. Third, the study design can be seen as a limitation since we used a pre-post research design, and there was no parallel control group. The data were collected from electronic health records, which possibly resulted in missing data. The total number of exacerbations is possibly higher, since not all mild or moderate exacerbations that were treated by the general practitioner were processed in the digital records of the Bravis Hospital. Furthermore, all respiratory-related hospitalizations were included, since pneumonia, dyspnea, and exacerbation frequently got mixed up. Finally, this study is not generalizable to all COPD patients since this study included just patients who were at risk for hospitalization due to exacerbation of COPD and this was a highly symptomatic group with a median 26.0 days with worsening symptoms, with one patient that was in the “yellow zone” nearly the whole study period (527 days). However, especially for the patient group that is at risk of hospitalizations [32], there is a need for intensive support to prevent future hospitalizations.

Future Research

For future studies, we recommend a study with a longer follow-up since it is known that interest in eHealth often declines over time, with fewer responses on alerts [33]. In this study, we

observed that patients did not fill out the daily questionnaire on a daily basis. However, as the goal of the intervention was to decrease exacerbations, it does not necessarily mean that the questionnaire should be completed every day. Nevertheless, the impact of usage on the effect of EmmaCOPD would be of interest. To strengthen the conclusions of this study, a case-control design can be considered to control for similar background. Furthermore, studies with a larger number of included patients are preferable, so small differences in outcomes can be detected as well.

Conclusion

EmmaCOPD, an eHealth program that includes an app that signalled symptoms, a smartwatch with step counter, and provision of feedback to the patient and buddies, seems to reduce the number of total exacerbations and the number of days of hospitalization due to exacerbation of COPD in this real-world study. The effects of long-term use of EmmaCOPD should be studied further in future studies.

Acknowledgments

We acknowledge Bart Mertsens, Leiden University Medical Center, Leiden, the Netherlands for statistical checks. Furthermore, we acknowledge Bregje de Jong-ten Berge, Harry van Looij, and Rini Voeten from the Bravis ziekenhuis, Roosendaal, the Netherlands, for making possible the data extraction from the Bravis ziekenhuis.

Conflicts of Interest

OvD is founder and CEO of MedicineMen. All other authors declare that they have no conflicts of interest.

References

1. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet 2012 Dec 15;380(9859):2163-2196 [FREE Full text] [doi: 10.1016/S0140-6736(12)61729-2] [Medline: 23245607]
2. Shavelle RM, Paculdo DR, Kush SJ, Mannino DM, Strauss DJ. Life expectancy and years of life lost in chronic obstructive pulmonary disease: findings from the NHANES III Follow-up Study. Int J Chron Obstruct Pulmon Dis 2009;4:137-148 [FREE Full text] [doi: 10.2147/copd.s5237] [Medline: 19436692]
3. Celli BR, Cote CG, Marin JM, Casanova C, Montes de Oca M, Mendez RA, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. N Engl J Med 2004 Mar 04;350(10):1005-1012. [doi: 10.1056/NEJMoa021322] [Medline: 14999112]
4. Soler-Cataluña JJ, Martínez-García MA, Román Sánchez P, Salcedo E, Navarro M, Ochando R. Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. Thorax 2005 Nov;60(11):925-931 [FREE Full text] [doi: 10.1136/thx.2005.040527] [Medline: 16055622]
5. Serra-Picamal X, Roman R, Escarrabill J, García-Altés A, Argimón JM, Soler N, et al. Hospitalizations due to exacerbations of COPD: A big data perspective. Respir Med 2018 Dec;145:219-225 [FREE Full text] [doi: 10.1016/j.rmed.2018.01.008] [Medline: 29398283]
6. 2021 Global Strategy for Prevention, Diagnosis and Management of COPD. Global Initiative for Chronic Obstructive Lung Disease. 2021. URL: http://www.goldcopd.org/ [accessed 2021-02-15]
7. Toy EL, Gallagher KF, Stanley EL, Swensen AR, Duh MS. The economic impact of exacerbations of chronic obstructive pulmonary disease and exacerbation definition: a review. COPD 2010 Jun;7(3):214-228. [doi: 10.3109/15412555.2010.481697] [Medline: 20486821]
8. Elbert NJ, van Os-Medendorp H, van Renselaar W, Ekeland AG, Hakkaart-van Roijen L, Raat H, et al. Effectiveness and cost-effectiveness of ehealth interventions in somatic diseases: a systematic review of systematic reviews and meta-analyses. J Med Internet Res 2014 Apr 16;16(4):e110 [FREE Full text] [doi: 10.2196/jmir.2790] [Medline: 24739471]
9. Lorig KR, Ritter PL, Laurent DD, Plant K. Internet-based chronic disease self-management: a randomized trial. Med Care 2006 Nov;44(11):964-971. [doi: 10.1097/01.mlr.0000233678.80203.c1] [Medline: 17063127]
10. Hallensleben C, van Luenen S, Rolink E, Ossebaard HC, Chavannes NH. eHealth for people with COPD in the Netherlands: a scoping review. Int J Chron Obstruct Pulmon Dis 2019;14:1681-1690 [FREE Full text] [doi: 10.2147/COPD.S207187] [Medline: 31440044]
11. Garcia-Aymerich J, Hernandez C, Alonso A, Casas A, Rodriguez-Roisin R, Anto JM, et al. Effects of an integrated care intervention on risk factors of COPD readmission. Respir Med 2007 Jul;101(7):1462-1469 [FREE Full text] [doi: 10.1016/j.rmed.2007.01.012] [Medline: 17339106]
12. Casas A, Troosters T, Garcia-Aymerich J, Roca J, Hernández C, Alonso A, members of the CHRONIC Project. Integrated care prevents hospitalisations for exacerbations in COPD patients. Eur Respir J 2006 Jul;28(1):123-130 [FREE Full text] [doi: 10.1183/09031936.06.00063205] [Medline: 16611656]
13. Trappenburg JCA, Niesink A, de Weert-van Oene GH, van der Zeijden H, van Snippenburg R, Peters A, et al. Effects of telemonitoring in patients with chronic obstructive pulmonary disease. Telemed J E Health 2008 Mar;14(2):138-146. [doi: 10.1089/tmj.2007.0037] [Medline: 18361703]

14. Hanlon P, Daines L, Campbell C, McKinstry B, Weller D, Pinnock H. Telehealth Interventions to Support Self-Management of Long-Term Conditions: A Systematic Metareview of Diabetes, Heart Failure, Asthma, Chronic Obstructive Pulmonary Disease, and Cancer. J Med Internet Res 2017 May 17;19(5):e172 [FREE Full text] [doi: 10.2196/jmir.6688] [Medline: 28526671]
15. Sul A, Lyu D, Park D. Effectiveness of telemonitoring versus usual care for chronic obstructive pulmonary disease: A systematic review and meta-analysis. J Telemed Telecare 2020 May;26(4):189-199. [doi: 10.1177/1357633X18811757] [Medline: 30541375]
16. Lundell S, Holmner A, Rehn B, Nyberg A, Wadell K. Telehealthcare in COPD: a systematic review and meta-analysis on physical outcomes and dyspnea. Respir Med 2015 Jan;109(1):11-26 [FREE Full text] [doi: 10.1016/j.rmed.2014.10.008] [Medline: 25464906]
17. McLean S, Nurmatov U, Liu JLY, Pagliari C, Car J, Sheikh A. Telehealthcare for chronic obstructive pulmonary disease: Cochrane Review and meta-analysis. Br J Gen Pract 2012 Nov;62(604):e739-e749 [FREE Full text] [doi: 10.3399/bjgp12X658269] [Medline: 23211177]
18. Gregersen TL, Green A, Frausing E, Ringbæk T, Brøndum E, Suppli Ulrik C. Do telemedical interventions improve quality of life in patients with COPD? A systematic review. Int J Chron Obstruct Pulmon Dis 2016;11:809-822 [FREE Full text] [doi: 10.2147/COPD.S96079] [Medline: 27143872]
19. EmmaMed. Medicine Men. 2021. URL: https://www.medicinemen.eu/nl/producten/emma-2/ [accessed 2021-02-15]
20. Longaanval actieplan COPD. Longfonds. 2021. URL: https://bestellen.longfonds.nl/product/longaanval-actieplan-copd/ [accessed 2021-02-15]
21. Kon SSC, Dilaver D, Mittal M, Nolan CM, Clark AL, Canavan JL, et al. The Clinical COPD Questionnaire: response to pulmonary rehabilitation and minimal clinically important difference. Thorax 2014 Sep;69(9):793-798. [doi: 10.1136/thoraxjnl-2013-204119] [Medline: 24149828]
22. Hurst JR, Vestbo J, Anzueto A, Locantore N, Müllerova H, Tal-Singer R, Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) Investigators. Susceptibility to exacerbation in chronic obstructive pulmonary disease. N Engl J Med 2010 Sep 16;363(12):1128-1138. [doi: 10.1056/NEJMoa0909883] [Medline: 20843247]
23. Betere en doelmatigere longzorg. Long Alliantie Nederland. 2014. URL: http://www.longalliantie.nl/nationaal-actieprogramma-chronische-longziekten/ [accessed 2021-02-15]
24. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [doi: 10.1016/0021-9681(87)90171-8] [Medline: 3558716]
25. Sobnath DD, Philip N, Kayyali R, Nabhani-Gebara S, Pierscionek B, Vaes AW, et al. Features of a Mobile Support App for Patients With Chronic Obstructive Pulmonary Disease: Literature Review and Current Applications. JMIR Mhealth Uhealth 2017 Feb 20;5(2):e17 [FREE Full text] [doi: 10.2196/mhealth.4951] [Medline: 28219878]
26. Hallensleben C, Biewenga J, Boven VJ, Chavannes NH. REducing Delay through edUcation on eXacerbations (REDUX) in patients with COPD. European Respiratory Journal 2018;52(Suppl. 62):PA919. [doi: 10.1183/13993003.congress-2018.PA919]
27. Slok AHM, In 't Veen JCCM, Chavannes NH, van der Molen T, Mölken MPR, Kerstjens HAM, et al. Effectiveness of the Assessment of Burden of Chronic Obstructive Pulmonary Disease (ABC) tool: study protocol of a cluster randomised trial in primary and secondary care. BMC Pulm Med 2014 Aug 07;14:131 [FREE Full text] [doi: 10.1186/1471-2466-14-131] [Medline: 25098313]
28. Lenferink A, Brusse-Keizer M, van der Valk PD, Frith PA, Zwerink M, Monninkhof EM, et al. Self-management interventions including action plans for exacerbations versus usual care in patients with chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2017 Aug 04;8:CD011682 [FREE Full text] [doi: 10.1002/14651858.CD011682.pub2] [Medline: 28777450]
29. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
30. Benzo R, Wetzstein M, Neuenfeldt P, McEvoy C. Implementation of physical activity programs after COPD hospitalizations: Lessons from a randomized study. Chron Respir Dis 2015 Feb;12(1):5-10 [FREE Full text] [doi: 10.1177/1479972314562208] [Medline: 25511306]
31. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
32. Alqahtani JS, Njoku CM, Bereznicki B, Wimmer BC, Peterson GM, Kinsman L, et al. Risk factors for all-cause hospital readmission following exacerbation of COPD: a systematic review and meta-analysis. Eur Respir Rev 2020 Jun 30;29(156) [FREE Full text] [doi: 10.1183/16000617.0166-2019] [Medline: 32499306]
33. Kargiannakis M, Fitzsimmons DA, Bentley CL, Mountain GA. Does Telehealth Monitoring Identify Exacerbations of Chronic Obstructive Pulmonary Disease and Reduce Hospitalisations? An Analysis of System Data. JMIR Med Inform 2017 Mar 22;5(1):e8 [FREE Full text] [doi: 10.2196/medinform.6359] [Medline: 28330829]

Abbreviations

CCI: Charlson comorbidity index
CCQ: clinical COPD questionnaire
COPD: chronic obstructive pulmonary disease
GOLD: Global Initiative for Chronic Obstructive Lung Disease
IRR: incidence rate ratio

Edited by G Eysenbach; submitted 03.10.20; peer-reviewed by E Metting, N Agarwal, Z Ren; comments to author 14.11.20; revised version received 23.12.20; accepted 14.01.21; published 18.03.21.

Please cite as:
van Buul AR, Derksen C, Hoedemaker O, van Dijk O, Chavannes NH, Kasteleyn MJ
eHealth Program to Reduce Hospitalizations Due to Acute Exacerbation of Chronic Obstructive Pulmonary Disease: Retrospective Study
JMIR Form Res 2021;5(3):e24726
URL: https://formative.jmir.org/2021/3/e24726
doi:10.2196/24726
PMID:33734091

Original Paper

e-Learning for Instruction and to Improve Reproducibility of Scoring Tumor-Stroma Ratio in Colon Carcinoma: Performance and Reproducibility Assessment in the UNITED Study

Marloes A Smit[1], MD; Gabi W van Pelt[1], BSc; Elisabeth MC Dequeker[2], PhD; Raed Al Dieri[3], PhD; Rob AEM Tollenaar[1], MD, PhD; J Han JM van Krieken[4], MD, PhD; Wilma E Mesker[1], PhD; UNITED Group[1]

[1]Department of Surgery, Leiden University Medical Center, Leiden, Netherlands

[2]Department of Public Health and Primary Care, Biomedical Quality Assurance Research Unit, University of Leuven, Leuven, Belgium

[3]European Society of Pathology, Brussels, Belgium

[4]Department of Pathology, Radboud University Medical Center, Nijmegen, Netherlands

Corresponding Author:
Wilma E Mesker, PhD
Department of Surgery
Leiden University Medical Center
Albinusdreef 2
Leiden, 2333 ZA
Netherlands
Phone: 31 715264005
Email: w.e.mesker@lumc.nl

Abstract

Background: The amount of stroma in the primary tumor is an important prognostic parameter. The tumor-stroma ratio (TSR) was previously validated by international research groups as a robust parameter with good interobserver agreement.

Objective: The Uniform Noting for International Application of the Tumor-Stroma Ratio as an Easy Diagnostic Tool (UNITED) study was developed to bring the TSR to clinical implementation. As part of the study, an e-Learning module was constructed to confirm the reproducibility of scoring the TSR after proper instruction.

Methods: The e-Learning module consists of an autoinstruction for TSR determination (instruction video or written protocol) and three sets of 40 cases (training, test, and repetition sets). Scoring the TSR is performed on hematoxylin and eosin–stained sections and takes only 1-2 minutes. Cases are considered stroma-low if the amount of stroma is ≤50%, whereas a stroma-high case is defined as >50% stroma. Inter- and intraobserver agreements were determined based on the Cohen κ score after each set to evaluate the reproducibility.

Results: Pathologists and pathology residents (N=63) with special interest in colorectal cancer participated in the e-Learning. Forty-nine participants started the e-Learning and 31 (63%) finished the whole cycle (3 sets). A significant improvement was observed from the training set to the test set; the median κ score improved from 0.72 to 0.77 (P=.002).

Conclusions: e-Learning is an effective method to instruct pathologists and pathology residents for scoring the TSR. The reliability of scoring improved from the training to the test set and did not fall back with the repetition set, confirming the reproducibility of the TSR scoring method.

Trial Registration: The Netherlands Trial Registry NTR7270; https://www.trialregister.nl/trial/7072

International Registered Report Identifier (IRRID): RR2-10.2196/13464

(JMIR Form Res 2021;5(3):e19408) doi:10.2196/19408

KEYWORDS

colon cancer; tumor-stroma ratio; validation; e-Learning; reproducibility study; cancer; tumor; colon; reproducibility; carcinoma; prognosis; diagnostic; implementation; online learning

Introduction

The prediction of disease outcome for an individual patient as part of personalized medicine is becoming routine practice in the management of cancer patient treatment. Staging of colon cancer by pathologists is based on hematoxylin and eosin (H&E)-stained sections of the primary tumor. The tumor-node-metastasis (TNM) classification is used as the main selection criterion for additional treatment, along with noting of characteristics such as depth of invasion and differentiation grade [1], according to the American Joint Committee staging algorithm. However, conventional H&E sections provide more information than previously recognized.

In the last decade, research has not only focused on the tumor and its characteristics but increasingly also on the tumor microenvironment. The tumor microenvironment consists of the stromal background with a variety of cells such as fibroblasts, endothelial cells, and lymphocytes. The tumor-stroma ratio (TSR) is the amount of tumor relative to the amount of stroma in the primary tumor [2-4]. Patients with a high amount of stroma (stroma-high) have a worse prognosis compared to those harboring tumors with a low amount of stroma (stroma-low) in multiple types of cancer [5-13].

Scoring the TSR is performed on H&E-stained sections in only 1-2 minutes, with good to excellent interobserver agreement [3]. This implies that TSR scoring is easy to learn and useful in daily practice.

The Uniform Noting for International Application of the Tumor-Stroma Ratio as an Easy Diagnostic Tool (UNITED) study was developed to prepare for implementation of the TSR as an additional high-risk indicator along with traditional TNM classification. As part of the UNITED study, an instruction protocol and reproducibility study were initiated using an e-Learning module as described in the published research protocol [14].

Digital pathology is increasingly being implemented in daily diagnostic practice as well as for teaching. Digital pathology for instruction, most commonly used for instruction of students, has multiple advantages: more students can be reached because it is web-based; all students look at exactly the same case; annotations can be shared with the teacher, resulting in direct feedback; and students can complete the course when and where it suits them [15-18].

The use of e-Learning for education has been adopted in different medical specialties worldwide. An example of e-Learning used in pathology is a module developed for Dutch pathologists [19,20]. The module focuses on decreasing the variation in grading dysplasia in adenomas and increasing the consistency of scoring serrated lesions. Two separate studies have shown that e-Learning is a good method to improve performance [19,20].

e-Learning to instruct professionals has also been confirmed in specialties other than pathology [21,22]. For instance, based on a systematic review for surgical training (students, residents, and surgeons), Maertens et al [23] concluded that e-Learning is as effective as other methods for training.

The aim of this study was to confirm the high reproducibility of scoring the TSR using an e-Learning module to train a variety of pathologists.

Methods

Case Selection

The e-Learning module was based on H&E-stained sections of stage II and III colon cancer resection specimens. The cases were randomly selected from the archives of the Pathology Department of Leiden University Medical Centre (LUMC). The number of cases with very low stroma (ie, 10% or 20% stroma area) were limited from the analysis to increase the number of cases that are generally more difficult to score for pathologists and are therefore more suitable for training purposes. In the e-Learning, 55% of the cases were stroma-low (≤50% stroma) and 45% were stroma-high (>50% stroma). None of the patients had received neoadjuvant treatment at the time of sample collection. The sample size was based on a workable amount of cases to maintain quality without the case load being too high.

Slides were scanned using a 20× objective with the Panoramic 250 scanner (3D Histech) or with the IntelliSite Digital pathology slide scanner (Philips).

Participants

Pathologists and pathology residents from all over the world could participate in the UNITED study and in the e-Learning. The UNITED study started in 2018, and pathologists were invited to start (and complete) the e-Learning in the period of December 2017 to April 2019. Data collection ended in April 2019. Multimedia Appendix 1 provides an overview of participating countries, and the numbers of participating pathologists and residents.

TSR Scoring Method

The previously published protocol for scoring the TSR was used in this study [3,4]. In brief, the section of the deepest part of the tumor, usually the section used for determining the T-stage, was chosen. The area with the highest amount of stroma was selected and scored at 100× magnification in increments of 10%. A field should contain tumor cells on four opposite edges of the field of evaluation.

e-Learning

The e-Learning module was developed in PathXL Tutor version 6.1.1.1. (Philips, Belfast, UK). This software uses digital images, and was developed to easily share and teach a network of pathology students, or in this case pathologists. The PathXL software allows the pathologist to analyze the slide in a manner comparable to using a microscope. The e-Learning was prepared to resemble real-life microscopy as far as possible by using round annotations with a fixed size of 3.4 mm^2. This represents the size of the field of vision of microscopes, even from different brands, when using 100× magnification.

Participants were blinded to clinical data of the sections and were only informed that the patients did not receive neoadjuvant treatment.

Before starting the first set of the e-Learning module, participants were asked to watch the instruction video on the study website [24]. The training set consisted of 40 cases. This set started with 5 multiple choice questions where annotations were placed upfront in different areas of the section (Multimedia Appendix 2). Participants were asked to select the correctly placed annotation and determine the percentage of stroma. In the other 35 cases (and in the other sets), the participant was asked to place the annotations themselves at the most optimal position and to determine the stroma percentage.

The test set also consisted of 40 cases, including 37 (93%) new cases. To determine the intraobserver agreement of scoring the TSR, the test set was repeated (repetition set) after a 2-month washout period with the same cases placed in a different order.

The answer model used for evaluating the results was established by two experienced observers (GP, MS) together with a pathologist (VS) of the LUMC. The coordinators of the UNITED study checked all finished e-Learning sets for stroma percentage and for correct placement of the annotation. The answers of the participants were compared with the answer model. Continuing with the second set was allowed when an interobserver agreement (κ) of ≥0.7 with predefined scores was reached. In the case in which a participant did not pass a set due to a κ score below 0.7, the same set had to be rescored (see Multimedia Appendix 3 for the flowchart of the e-Learning module [14]) after feedback was given.

Statistical Analysis

Data collected in this study comprised: (1) whether the participant is a pathologist or resident, (2) participating country, (3) the stroma percentage of the different questions, and (4) whether or not the participant considered a question to be difficult. In this study, a possible bias could be that participating pathologists/residents are generally more motivated for participation.

Stroma percentages were classified as stroma-low (≤50% stroma) or as stroma-high (>50% stroma) [3,4,25]. This dichotomous output and the outcome of whether or not the annotation was placed correctly were used for measuring observer agreement. Cohen κ coefficient was used to measure inter- and intraobserver agreement. This score is quantitative and was used as a noncontinuous variable. Histograms were used to visualize the distribution of the κ scores for each set.

Nonnormally distributed continuous variables are described with the median and range (minimum and maximum values). For the median κ scores of a set, the first κ score of a participant of each set was used. The Wilcoxon signed-rank test was used to compare paired nonnormally distributed continuous variables (eg, measuring the progress between different e-Learning sets by participants).

Ethical Considerations

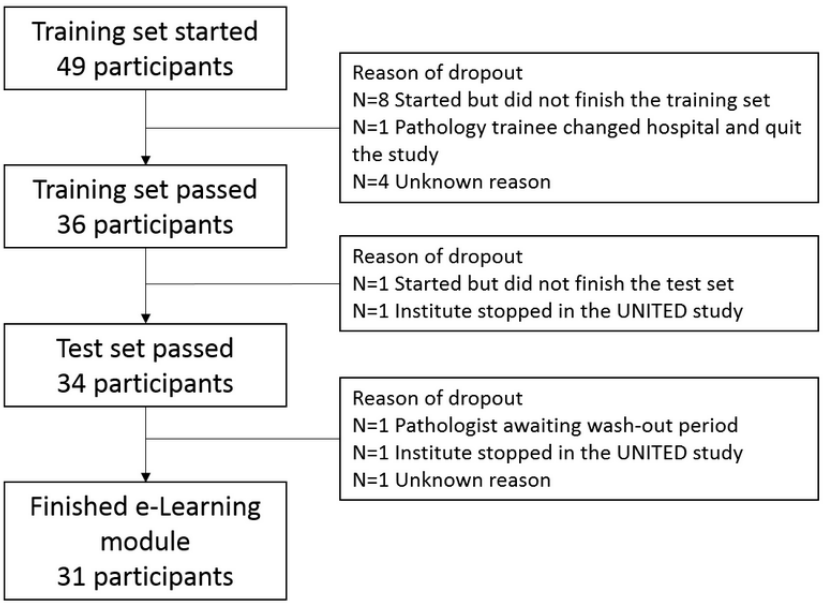
The UNITED study protocol has been approved by the Medical Research Ethics Committee of the LUMC (study number p17.302). All samples were handled in accordance with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Results

Participants

In total, 63 participants (49 pathologists and 14 residents) were registered for e-Learning. However, 14 participants were nonresponsive after registration, and thus 36 pathologists and 13 residents from 19 countries started the e-Learning; these 49 participants were used for analysis. All participating pathologists had gastrointestinal pathology as a subspecialty; however, most of them had more than one subspecialty. The residents had not yet chosen a pathology subspecialty. Thirty-six (73%) participants (28 pathologists and 8 residents) finished the training set and continued with the test set. In total, 31 (63%) participants finished the whole cycle (3 sets) of the e-Learning. A complete overview and reasons for participants to drop out are shown in Figure 1. The participants who quit the study were left out of the analysis.

Figure 1. Overview of the number of participants of the e-Learning and the reasons for not finishing the e-Learning cycle.



Observer Agreement

After finishing the training set of the e-Learning, the observer agreement was determined. Twenty-four (67%) participants (19 pathologists and 5 residents) passed the training set at the first attempt (ie, κ≥0.7). Two participants (both residents) needed a third chance to pass the training set. The median κ score for the training set was 0.72. The test set was passed the first time by 31 (91%) participants (23 pathologists and 8 residents) and 3 (9%) pathologists had to repeat the test set. The median κ score for the test set was 0.77. After a 2-month washout period, 28 (90%) participants (21 pathologists and 7 residents) directly passed the repetition set, with a median κ score of 0.76 (Table 1, Figure 2). A significant improvement was observed from the training set to the test set (P=.002). No significant changes of the κ scores were observed between the test set and the repetition set (P=.30, Figure 3). Intraobserver agreement was measured for the 31 participants who finished the repetition set. The median κ score of the intraobserver agreement was 0.77 (Table 1). Scoring results from pathologists showed significant improvement from the training set to the test set (P=.006) and no fall back (P=.74) after a washout period of 2 months. The scoring results of the residents showed no significant changes between sets (P=.26 and P=.13). Details are shown in Multimedia Appendix 4.

Table 1. Distribution of κ scores per set.

Set	Completed set, N	κ, median (range)
Training set	36	0.72 (0.21-0.90)
Test set	34	0.77 (0.51-0.97)
Repetition set	31	0.76 (0.60-0.89)
Intraobserver agreement	31	0.77 (0.61-0.95)

Figure 2. Distribution of κ scores for each set of the e-Learning: (A) training set, (B) test set, (C) repetition set, and (D) intraobserver agreement.

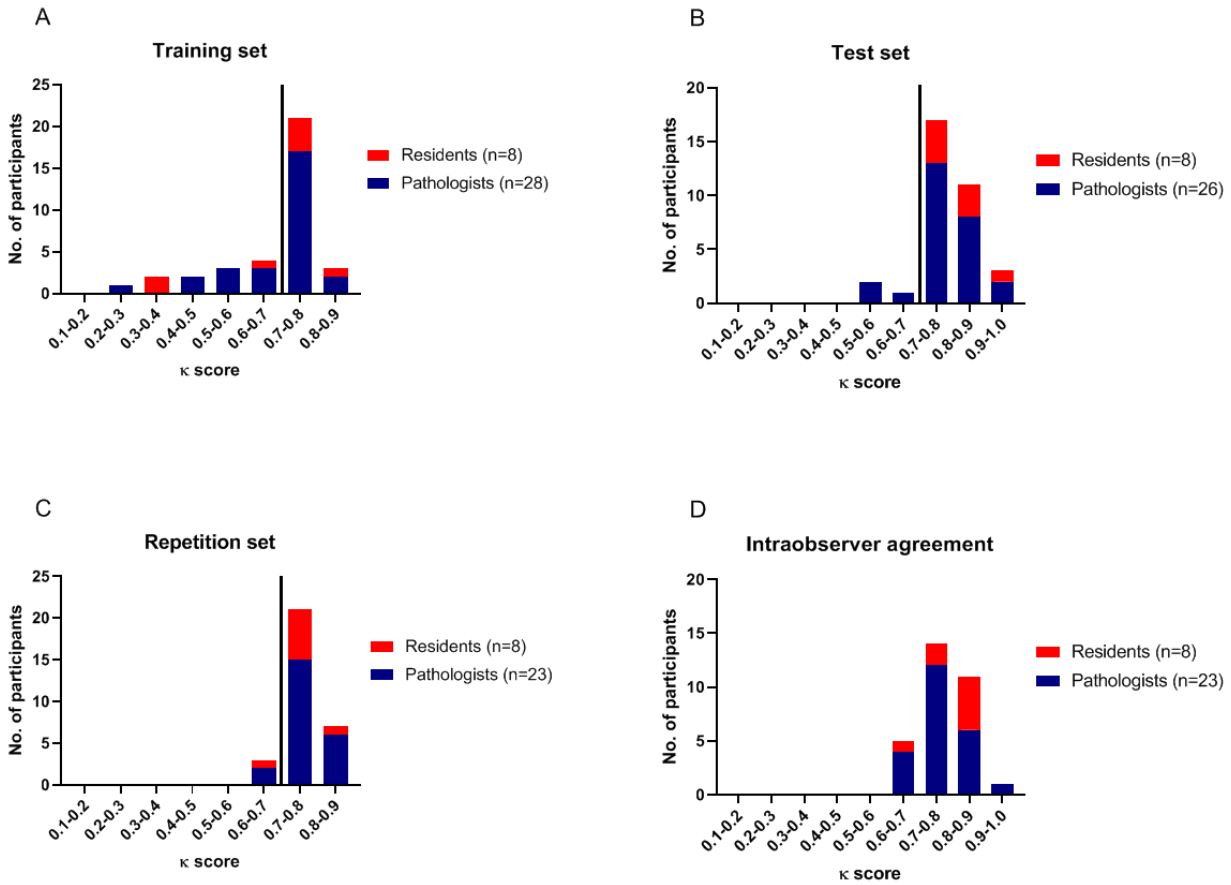
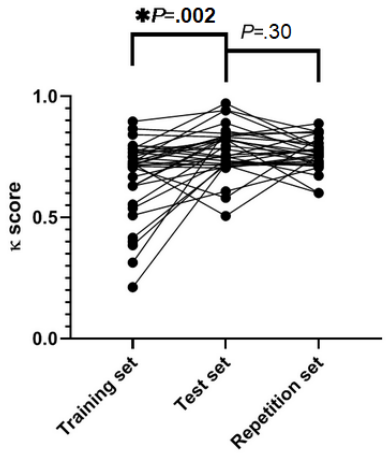


Figure 3. Progress of κ scores per participant during the e-Learning module.



Difficulty of the Questions

For each case, participants were asked whether scoring the TSR of the section was easy, normal, or difficult. If answered as difficult, a reason could be given. A case was classified as difficult when at least 40% of the participants agreed with this assessment. Eleven cases (9 different cases) were classified as difficult (4 in the training set, 2 in the test set, and 5 in the repetition set). The cases classified as difficult in the test set remained difficult in the repetition set. The three other cases of the repetition set were more difficult than average in the test set. As expected, most of the difficult cases were those close to the cut-off value of 50%, mucinous tumors, tumors with a lot of necrosis, and tumors in which the distinction between the stroma and the smooth muscle was difficult (see Multimedia Appendix 5 for examples of difficult cases). Overall, the 11 cases were more often answered wrong (29% of the answers) by participants who classified (one of) these cases as difficult compared to 19% of wrong answers for cases that were assessed as not difficult (see Multimedia Appendix 6 for the subdivision per case).

Drawing Annotations

A few cases were used in both the training and test sets. Analysis of these cases showed progress of scoring at the hotspot, as the annotations were more centered (Figure 4). Furthermore, in stroma-low cases, annotations were more widely spread over the entire tumor area, whereas one or two hotspots were more often identified in stroma-high cases (Figure 5).

Figure 4. Example of improvement of selecting the most optimal area for scoring the tumor-stroma ratio. Diffuse spread of annotations was seen in the training set (A) compared to less variation of the scoring area in the test set (B).

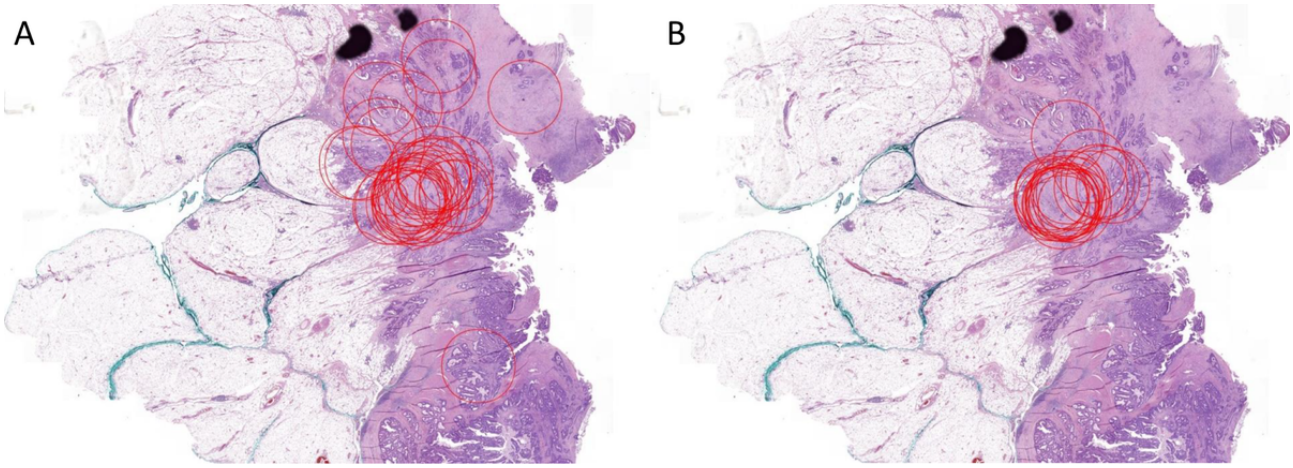
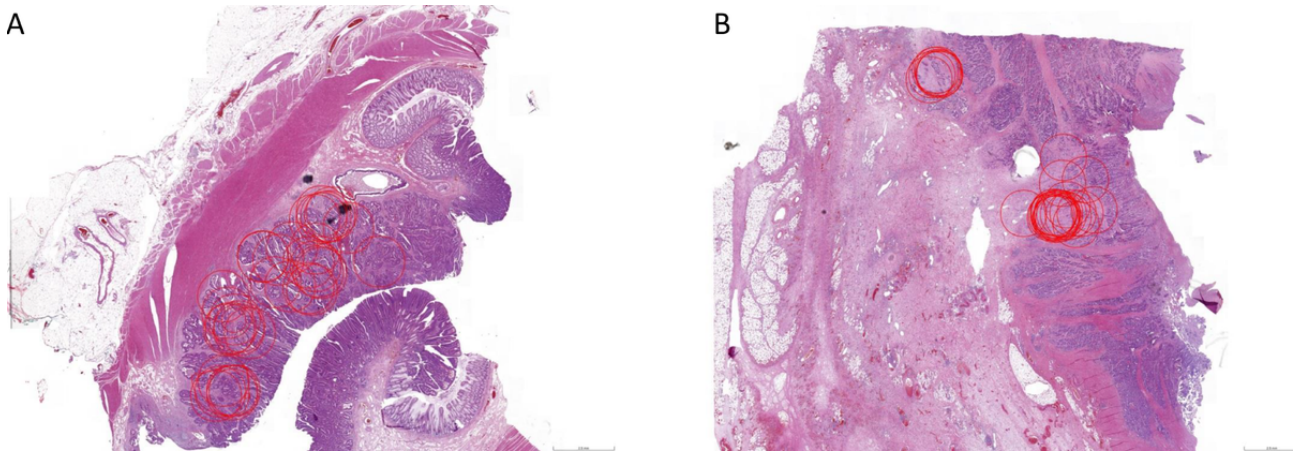


Figure 5. Examples of the distribution of annotations. In stroma-low cases (A), a wide distribution was more frequent, whereas in stroma-high cases, a hotspot (or sometimes more) formed (B).



Discussion

This study shows that e-Learning is an accurate method to instruct pathologists and pathology residents for scoring the TSR. A κ score >0.7 was used to define the reproducibility of the method. The median κ score improved and the minimum value increased after completing the consecutive sets for scoring the TSR. Significant progress (*P*=.002) was observed from the training to the test set. If a set was not passed the first time, feedback was given to the participant before they repeated the same set; however, participants did not get insight into their precise mistakes. The feedback was personalized, but not case-specific.

A decrease was observed between the number of created accounts, the number of participants who started the e-Learning, and the number of participants who completed the whole module. There are multiple reasons for this decrease. Most of the participants were registered by the principal investigator of an institute; however, not everyone responded to their registration. Another reason for dropout was withdrawal of the center from the UNITED study.

In daily diagnostics, H&E-stained sections are used for determining cancer stage. To ensure the high quality of the sections used for e-Learning, the original H&E-stained slides of the cases used for diagnostic purposes were scanned. Furthermore, all slides were scanned at the same magnification to avoid differences in the quality of the digital images.

When reviewing the e-Learning sets, in some cases it seemed as if the participants had scored the tumor percentage instead of the stroma percentage. This might be explained by the fact that pathologists are accustomed to scoring the neoplastic cell percentage for molecular analysis. Another explanation might be more related to semantics. The amount of tumor is in the numerator, which might be a source of confusion. In these doubtful cases, the participant was asked to reevaluate the case as well as some others. Thus, scoring the percentage of stroma remains a point of attention.

Overall, participants were well able to choose the right area for scoring the stroma percentage and to estimate whether a section was stroma-high or stroma-low. A common misinterpretation was scoring at the invasive front instead of looking for an area with as much stroma as possible within the section. This might be explained by the fact that pathologists are accustomed to scoring the tumor budding at the invasive front [26]. Furthermore, as the scoring protocol describes the use of the section from the deepest part of the tumor (usually the section used for determining the T-status) [3], the distinction between scoring the TSR in the whole tumor area and not necessarily at the invasive front might not have been made clear enough in the instructions. When comparing the results of pathologists and residents, pathologists showed significant improvement from the training to the test set, whereas residents did not. A possible explanation is the small group of residents (n=8) or the fact that pathologists are more experienced than residents.

Research performed on other pathology biomarkers used in daily practice such as lymphovascular invasion [27], tumor grading [28,29], classification and grading of colorectal polyps or adenomas [19,20,30-33], and the estimation of tumor cell percentage [34] has shown weak to moderate interobserver agreement. With three median κ scores above 0.7, the interobserver agreement for scoring the TSR in this study was found to be good. Although no comparison arm was included in this study, the median κ values obtained in this study are lower than those reported previously. This can be explained by the fact that the scores were low in the training set, which improved in the test set.

Digital pathology is increasingly entering pathology practice, although most pathology departments are not yet (fully) digitalized. In the future, a digital image analysis program will be useful for more accurate scoring and even better reproducibility. Digital pathology for teaching goals has some advantages and disadvantages. The advantages include easy distribution of samples and being able to score a set at the same time while reaching a worldwide group of pathologists, and disadvantages include possible software flaws or bugs when using a digitized workflow. In this study, the placed annotations were not always saved correctly, which sometimes made it difficult to analyze the results for a participant. In these particular cases, the participant was asked to reevaluate the case.

In conclusion, this study showed that e-Learning is a good and effective method to instruct pathologists and residents in scoring

the TSR. Furthermore, this study confirmed the reproducibility of the scoring method.

Acknowledgments

The authors would like to thank Professor Vincent Smit for scoring the cases in the e-Learning module to develop an answer model, Professor Hein Putter for advice on statistical analysis, and Ms. Cleo Keppens and Ms. Kelly Dufraing for their help with the e-Learning software. The UNITED study is funded by The Dutch Cancer Society (KWF Kankerbestrijding project 10174) and the Stichting Fonds Oncologie Holland.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of the participating countries and number of participating pathologists or residents per country.
[PDF File (Adobe PDF File), 112 KB - formative_v5i3e19408_app1.pdf]

Multimedia Appendix 2
Example of a multiple choice question of the e-Learning. Each participant was asked to select the annotation at the most optimal position following the criteria mentioned in the scoring protocol.
[PDF File (Adobe PDF File), 2755 KB - formative_v5i3e19408_app2.pdf]

Multimedia Appendix 3
Flowchart for the instruction of participating pathologists using the e-Learning module.
[PDF File (Adobe PDF File), 168 KB - formative_v5i3e19408_app3.pdf]

Multimedia Appendix 4
Progress of κ scores per participant separated for pathologists (A) and residents (B).
[PDF File (Adobe PDF File), 141 KB - formative_v5i3e19408_app4.pdf]

Multimedia Appendix 5
Examples of difficult cases for scoring the tumor-stroma ratio (TSR) as mentioned by the participants. (A) A case around the cut off-value of 50%. (B) Necrosis makes it difficult to select the most optimal site for the annotation. (C) A mucinous tumor can sometimes be difficult to score, because mucus has to be visually excluded from the scorings area. (D) Stromal tissue and smooth muscle tissue.
[PDF File (Adobe PDF File), 244 KB - formative_v5i3e19408_app5.pdf]

Multimedia Appendix 6
Subdivision of the 11 questions classified as difficult, showing the proportion of wrong answers per question.
[PDF File (Adobe PDF File), 108 KB - formative_v5i3e19408_app6.pdf]

References

1. Labianca R, Nordlinger B, Beretta G, Mosconi S, Mandalà M, Cervantes A, ESMO Guidelines Working Group. Early colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2013 Oct;24(Suppl 6):vi64-vi72 [FREE Full text] [doi: 10.1093/annonc/mdt354] [Medline: 24078664]
2. Mesker WE, Liefers G, Junggeburt JMC, van Pelt GW, Alberici P, Kuppen PJK, et al. Presence of a high amount of stroma and downregulation of SMAD4 predict for worse survival for stage I-II colon cancer patients. Cell Oncol 2009;31(3):169-178 [FREE Full text] [doi: 10.3233/CLO-2009-0478] [Medline: 19478385]
3. van Pelt GW, Kjær-Frifeldt S, van Krieken JHJM, Al Dieri R, Morreau H, Tollenaar RAEM, et al. Scoring the tumor-stroma ratio in colon cancer: procedure and recommendations. Virchows Arch 2018 Oct 20;473(4):405-412 [FREE Full text] [doi: 10.1007/s00428-018-2408-z] [Medline: 30030621]
4. Mesker WE, Junggeburt JMC, Szuhai K, de Heer P, Morreau H, Tanke HJ, et al. The carcinoma-stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumor stage. Cell Oncol 2007;29(5):387-398 [FREE Full text] [doi: 10.1155/2007/175276] [Medline: 17726261]
5. Wu J, Liang C, Chen M, Su W. Association between tumor-stroma ratio and prognosis in solid tumor patients: a systematic review and meta-analysis. Oncotarget 2016 Oct 18;7(42):68954-68965 [FREE Full text] [doi: 10.18632/oncotarget.12135] [Medline: 27661111]

6. Zhang R, Song W, Wang K, Zou S. Tumor-stroma ratio (TSR) as a potential novel predictor of prognosis in digestive system cancers: A meta-analysis. Clin Chim Acta 2017 Sep;472:64-68. [doi: 10.1016/j.cca.2017.07.003] [Medline: 28729135]
7. Kramer CJH, Vangangelt KMH, van Pelt GW, Dekker TJA, Tollenaar RAEM, Mesker WE. The prognostic value of tumour-stroma ratio in primary breast cancer with special attention to triple-negative tumours: a review. Breast Cancer Res Treat 2019 Jan 9;173(1):55-64 [FREE Full text] [doi: 10.1007/s10549-018-4987-4] [Medline: 30302588]
8. Wang K, Ma W, Wang J, Yu L, Zhang X, Wang Z, et al. Tumor-stroma ratio is an independent predictor for survival in esophageal squamous cell carcinoma. J Thorac Oncol 2012 Sep;7(9):1457-1461 [FREE Full text] [doi: 10.1097/JTO.0b013e318260dfe8] [Medline: 22843085]
9. Dekker TJA, van de Velde CJH, van Pelt GW, Kroep JR, Julien J, Smit VTHBM, et al. Prognostic significance of the tumor-stroma ratio: validation study in node-negative premenopausal breast cancer patients from the EORTC perioperative chemotherapy (POP) trial (10854). Breast Cancer Res Treat 2013 Jun 25;139(2):371-379. [doi: 10.1007/s10549-013-2571-5] [Medline: 23709090]
10. Eriksen AC, Sørensen FB, Lindebjerg J, Hager H, dePont Christensen R, Kjær-Frifeldt S, et al. The prognostic value of tumour stroma ratio and tumour budding in stage II colon cancer. A nationwide population-based study. Int J Colorectal Dis 2018 Aug 21;33(8):1115-1124 [FREE Full text] [doi: 10.1007/s00384-018-3076-9] [Medline: 29785462]
11. Peng C, Liu J, Yang G, Li Y. The tumor-stromal ratio as a strong prognosticator for advanced gastric cancer patients: proposal of a new TSNM staging system. J Gastroenterol 2018 May 16;53(5):606-617 [FREE Full text] [doi: 10.1007/s00535-017-1379-1] [Medline: 28815347]
12. Kemi N, Eskuri M, Herva A, Leppänen J, Huhta H, Helminen O, et al. Tumour-stroma ratio and prognosis in gastric adenocarcinoma. Br J Cancer 2018 Aug 30;119(4):435-439 [FREE Full text] [doi: 10.1038/s41416-018-0202-y] [Medline: 30057407]
13. Pongsuvareeyakul T, Khunamornpong S, Settakorn J, Sukpan K, Suprasert P, Intaraphet S, et al. Prognostic evaluation of tumor-stroma ratio in patients with early stage cervical adenocarcinoma treated by surgery. Asian Pac J Cancer Prev 2015;16(10):4363-4368 [FREE Full text] [doi: 10.7314/apjcp.2015.16.10.4363] [Medline: 26028100]
14. Smit M, van Pelt G, Roodvoets A, Meershoek-Klein Kranenbarg E, Putter H, Tollenaar R, et al. Uniform noting for international application of the tumor-stroma ratio as an easy diagnostic tool: protocol for a multicenter prospective cohort study. JMIR Res Protoc 2019 Jun 14;8(6):e13464 [FREE Full text] [doi: 10.2196/13464] [Medline: 31199317]
15. Parker EU, Reder NP, Glasser D, Henriksen J, Kilgore MR, Rendi MH. NDER: a novel web application for teaching histology to medical students. Acad Pathol 2017 Feb 10;4:2374289517691061 [FREE Full text] [doi: 10.1177/2374289517691061] [Medline: 28725787]
16. Helle L, Nivala M, Kronqvist P, Gegenfurtner A, Björk P, Säljö R. Traditional microscopy instruction versus process-oriented virtual microscopy instruction: a naturalistic experiment with control group. Diagn Pathol 2011 Mar 30;6(Suppl 1):S8 [FREE Full text] [doi: 10.1186/1746-1596-6-S1-S8] [Medline: 21489203]
17. Sahota M, Leung B, Dowdell S, Velan GM. Learning pathology using collaborative vs. individual annotation of whole slide images: a mixed methods trial. BMC Med Educ 2016 Dec 12;16(1):311 [FREE Full text] [doi: 10.1186/s12909-016-0831-x] [Medline: 27955651]
18. Samulski TD, Taylor LA, La T, Mehr CR, McGrath CM, Wu RI. The utility of adaptive eLearning in cervical cytopathology education. Cancer Cytopathol 2018 Feb 20;126(2):129-135. [doi: 10.1002/cncy.21942] [Medline: 29053223]
19. IJspeert JEG, Madani A, Overbeek LIH, Dekker E, Nagtegaal ID. Implementation of an e-learning module improves consistency in the histopathological diagnosis of sessile serrated lesions within a nationwide population screening programme. Histopathology 2017 May 08;70(6):929-937. [doi: 10.1111/his.13155] [Medline: 28000308]
20. Madani A, Kuijpers CCHJ, Sluijter CE, Von der Thüsen JH, Grünberg K, Lemmens VEPP, et al. Decrease of variation in the grading of dysplasia in colorectal adenomas with a national e-learning module. Histopathology 2019 May 10;74(6):925-932. [doi: 10.1111/his.13834] [Medline: 30725483]
21. Nakanishi H, Doyama H, Ishikawa H, Uedo N, Gotoda T, Kato M, et al. Evaluation of an e-learning system for diagnosis of gastric lesions using magnifying narrow-band imaging: a multicenter randomized controlled study. Endoscopy 2017 Oct 21;49(10):957-967. [doi: 10.1055/s-0043-111888] [Medline: 28637065]
22. Canty D, Barth J, Yang Y, Peters N, Palmer A, Royse A, et al. Comparison of learning outcomes for teaching focused cardiac ultrasound to physicians: A supervised human model course versus an eLearning guided self- directed simulator course. J Crit Care 2019 Feb;49:38-44. [doi: 10.1016/j.jcrc.2018.10.006] [Medline: 30359924]
23. Maertens H, Madani A, Landry T, Vermassen F, Van Herzeele I, Aggarwal R. Systematic review of e-learning for surgical training. Br J Surg 2016 Oct;103(11):1428-1437. [doi: 10.1002/bjs.10236] [Medline: 27537708]
24. Tumor-stroma ratio. UNITED. URL: http://watchstroma.com/ [accessed 2019-05-02] [WebCite Cache ID 784El3ZGH]
25. Huijbers A, Tollenaar R, v Pelt G, Zeestraten E, Dutton S, McConkey C, et al. The proportion of tumor-stroma as a strong prognosticator for stage II and III colon cancer patients: validation in the VICTOR trial. Ann Oncol 2013 Jan;24(1):179-185 [FREE Full text] [doi: 10.1093/annonc/mds246] [Medline: 22865778]
26. Lugli A, Kirsch R, Ajioka Y, Bosman F, Cathomas G, Dawson H, et al. Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. Mod Pathol 2017 Sep;30(9):1299-1311. [doi: 10.1038/modpathol.2017.46] [Medline: 28548122]

27. Harris E, Lewin D, Wang H, Lauwers G, Srivastava A, Shyr Y, et al. Lymphovascular invasion in colorectal cancer: an interobserver variability study. Am J Surg Pathol 2008 Dec;32(12):1816-1821 [FREE Full text] [doi: 10.1097/PAS.0b013e3181816083] [Medline: 18779725]
28. Chandler I, Houlston RS. Interobserver agreement in grading of colorectal cancers-findings from a nationwide web-based survey of histopathologists. Histopathology 2008 Mar;52(4):494-499. [doi: 10.1111/j.1365-2559.2008.02976.x] [Medline: 18315602]
29. Barresi V, Reggiani Bonetti L, Ieni A, Branca G, Tuccari G. Histologic prognostic markers in stage IIA colorectal cancer: a comparative study. Scand J Gastroenterol 2016 Mar;51(3):314-320. [doi: 10.3109/00365521.2015.1084646] [Medline: 26554618]
30. Jensen P, Krogsgaard MR, Christiansen J, Braendstrup O, Johansen A, Olsen J. Observer variability in the assessment of type and dysplasia of colorectal adenomas, analyzed using kappa statistics. Dis Colon Rectum 1995 Feb;38(2):195-198. [doi: 10.1007/BF02052450] [Medline: 7851176]
31. Terry MB, Neugut AI, Bostick RM, Potter JD, Haile RW, Fenoglio-Preiser CM. Reliability in the classification of advanced colorectal adenomas. Cancer Epidemiol Biomarkers Prev 2002 Jul;11(7):660-663 [FREE Full text] [Medline: 12101114]
32. Denis B, Peters C, Chapelain C, Kleinclaus I, Fricker A, Wild R, et al. Diagnostic accuracy of community pathologists in the interpretation of colorectal polyps. Eur J Gastroenterol Hepatol 2009 Oct;21(10):1153-1160. [doi: 10.1097/MEG.0b013e3283221124] [Medline: 19696682]
33. Costantini M, Sciallero S, Giannini A, Gatteschi B, Rinaldi P, Lanzanova G, SMAC Workgroup. Interobserver agreement in the histologic diagnosis of colorectal polyps. the experience of the multicenter adenoma colorectal study (SMAC). J Clin Epidemiol 2003 Mar;56(3):209-214. [doi: 10.1016/s0895-4356(02)00587-5] [Medline: 12725874]
34. Smits AJJ, Kummer JA, de Bruin PC, Bol M, van den Tweel JG, Seldenrijk KA, et al. The estimation of tumor cell percentage for molecular testing by pathologists is not accurate. Mod Pathol 2014 Feb;27(2):168-174. [doi: 10.1038/modpathol.2013.134] [Medline: 23887293]

Abbreviations

H&E: hematoxylin and eosin
LUMC: Leiden University Medical Centre
TNM: tumor node metastasis
TSR: tumor-stroma ratio
UNITED: Uniform Noting for International Application of the Tumor-Stroma Ratio as an Easy Diagnostic Tool

Edited by G Eysenbach; submitted 17.04.20; peer-reviewed by Y Chu, CS Wu; comments to author 25.09.20; revised version received 14.12.20; accepted 03.03.21; published 19.03.21.

Please cite as:
Smit MA, van Pelt GW, Dequeker EMC, Al Dieri R, Tollenaar RAEM, van Krieken JHJM, Mesker WE, UNITED Group
e-Learning for Instruction and to Improve Reproducibility of Scoring Tumor-Stroma Ratio in Colon Carcinoma: Performance and Reproducibility Assessment in the UNITED Study
JMIR Form Res 2021;5(3):e19408
URL: https://formative.jmir.org/2021/3/e19408
doi:10.2196/19408
PMID:33739293

Original Paper

A Mobile Patient-Facing App for Tracking Patient-Reported Outcomes in Head and Neck Cancer Survivors: Single-Arm Feasibility Study

Sewit Teckie[1,2], MD; Jeffrey Solomon[3], BFA; Karthik Kadapa[4], MS; Keisy Sanchez[3], BS; David Orner[1,2], MPH; Dennis Kraus[5,6], MD; Dev P Kamdar[5,6], MD; Lucio Pereira[5,6], MD; Douglas Frank[5,6], MD; Michael Diefenbach[3], PhD

[1]Academic Department of Radiation Medicine, Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, New York, NY, United States
[2]Department of Radiation Medicine, Northwell Health Cancer Institute, Lake Success, NY, United States
[3]Center for Health Innovations and Outcomes Research, Department of Medicine, Northwell Health, Manhasset, NY, United States
[4]Center for Research Informatics & Innovation, Feinstein Institutes for Medical Research, Northwell Health, Manhasset, NY, United States
[5]Department of Otolaryngology, Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, United States
[6]Department of Otolaryngology, Northwell Health Cancer Institute, Lake Success, NY, United States

Corresponding Author:
Sewit Teckie, MD
Academic Department of Radiation Medicine
Donald and Barbara Zucker School of Medicine at Hofstra/Northwell
130 East 77th Street
New York, NY, 10010
United States
Phone: 1 212 434 2919
Email: steckie@northwell.edu

Abstract

Background: Patients with head and neck cancer (HNC) frequently experience disease-related symptoms and treatment adverse effects that impact their overall quality of life. Cancer-specific mobile health apps for patient-related outcomes allow patients to communicate with their clinicians and proactively track their symptoms, which have been shown to improve clinical management and disease outcomes.

Objective: The purpose of this study was to evaluate the feasibility of LogPAL, a novel iPhone-based mobile health app designed to help HNC survivors track and manage their posttreatment symptoms.

Methods: Patients who completed curative treatment for HNC in the preceding 24 months were recruited from 2 clinical sites within a single institution. Upon enrollment, participants completed a brief sociodemographic survey, downloaded the app onto their iPhone devices, and were asked to complete a series of biweekly questionnaires (based on the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events) via the app for an 8-week study period. The primary feasibility endpoints included retention (retaining >80% of the enrolled participants for the duration of the study period), adherence (>50% of the participants completing 100% of the questionnaires over the study period), and usability (a mean system usability scale [SUS] score >68). Additional postintervention questions were collected to assess perceived usefulness, acceptance, and overall satisfaction.

Results: Between January and October 2019, 38 participants were enrolled in the study. Three participants dropped out, and 3 were classified as nonusers. The remaining 32 (87%) were eligible for analysis. Their mean age was 57.8 (SD 12.3) years (range 24-77 years, 81% [26/32] male). Overall, 375 of 512 (73.2%) questionnaires were completed, with 17 (53%) of the 32 participants adherent. Participant-reported usability was acceptable; the mean SUS score was 71.9 (95% CI 64.3-79.5) with high satisfaction of LogPAL usefulness and likelihood to recommend to other cancer survivors.

Conclusions: This single-arm prospective pilot study showed that LogPAL is a feasible, regularly used, accepted app for HNC survivors, justifying a full-scale pilot. Based on the findings from this study, future iterations will aim to improve usability and test intervention efficacy.

(JMIR Form Res 2021;5(3):e24667) doi:10.2196/24667

KEYWORDS

mHealth; ePROs; head and neck cancer; mobile phone

Introduction

Background

The demographics of head and neck cancer (HNC) are changing [1-4], with an increase in the incidence of HNC in younger patients without significant smoking or alcohol use history. HNC survival rates are increasing as a result of these changing demographics and more effective multidisciplinary treatment options [5-8]. Despite these advances in disease outcomes, HNC survivors often experience significant toxicities and unique functional impairments such as dysphagia, mucositis, xerostomia, and dysphonia, which are distinctive from those reported in other cancer survivors [9-11]. These can lead to debilitating and lifelong consequences due to adverse effects during and after treatment, which impair the quality of life. Nevertheless, patients commonly underreport symptoms or delay reporting of symptoms, which, in turn, lead to delay in the best clinical management [12,13]. Mobile health (mHealth) interventions such as smartphone apps have been advocated as promising strategies in patient self-management [14]. These tools have the potential to increase accessibility of patient's health information, provide real-time reporting of the concerning symptoms to providers, and improve overall patient satisfaction via proactive self-management care [15-19]. The proliferation of mHealth apps, with over 325,000 mHealth apps developed [20], is changing how patients interact with the health care system. Yet, there remains a dearth in knowledge regarding the feasibility of using mHealth electronic patient-reported outcomes (ePROs) to assess and address the vulnerable population of HNC survivors. To date, only few studies have evaluated a smartphone-based self-management system during HNC treatment [16,21]. To fill this knowledge gap, our study aimed to evaluate the feasibility of LogPAL, a novel patient-facing mHealth app specific for HNC survivors.

Objectives of This Study

The primary objectives of this study were to (1) assess the feasibility of LogPAL and (2) explore the perceived usefulness, acceptance, and overall satisfaction through validated questionnaires and participant feedback. We hypothesized that there would be an >80% retention rate (proportion of enrolled participants who completed at least one questionnaire during the study period), >50% adherence rate (proportion of participants who will complete 100% of scheduled questionnaires), and a mean system usability scale (SUS) score >68. Secondary exploratory objectives reviewed additional engagement metrics and preliminary associations between outcome measures and participant characteristics (covariates). As there is no consensus on how best to evaluate mHealth pilot studies, criteria selection and determination were based on insights from the one of the co-principal investigators (MAD) who is experienced in evaluating web-based and app-based health tools for patients with cancer.

Methods

Recruitment and Enrollment

All HNC survivors were screened for eligibility from 2 Northwell Health otolaryngology surgical oncology clinics. Potential participants were identified through a review of electronic medical records to determine if the following inclusion criteria were met prior to their upcoming in-clinic appointment: HNC survivors ≥18 years old who completed all curative treatments for primary HNC (lip/oral cavity, pharynx, larynx, salivary gland, paranasal sinus) within the preceding 24 months, who had no serious self-reported cognitive impairment, who were able to read and speak English fluently, and who had access to a smartphone that operates the iPhone operating system software (iPhone/iPad). Eligible participants were contacted by phone to discuss meetings about enrollment at the clinic during an upcoming appointment. Upon meeting, additional details about the study were reviewed and participants were informed that participation was voluntary. If informed consent was provided, participants underwent a training session with a study team member on how to use the app, with additional time for any questions prior to the first log-in.

Study Design

A nonrandomized, prospective, single-arm pilot study was carried out from January to October 2019. The study period consisted of an 8-week intervention during which participants were instructed to answer and complete ePROs twice weekly for 16 sessions. The frequency and duration of the ePROs were determined by a team of HNC specialists to reflect clinically relevant timepoints that encouraged self-monitoring, but did not compromise HNC patient safety during their recovery. This study was reviewed and approved by the Institutional Review Board of Northwell Health.

LogPAL App

LogPAL Development

The LogPAL app was developed using an iterative user-centered design approach [22,23], which systematically engaged end users to identify requirements and app core functionality to enhance the effectiveness, efficiency, and usability [24,25]. The goal was to develop a patient-facing mHealth app that could provide HNC survivors with a tool to better manage and track their symptoms by self-reporting any adverse effects and offer immediate informational resources to help develop and practice health self-management skills.

LogPAL Overview

For this pilot study, a version of LogPAL, currently available in the Apple App Store, that consisted of the following 4 core features was used: (1) Start Tracking, (2) View My Progress, (3) Self Care Tips, and (4) Resources (Figure 1). Users were not specifically required to access other features other than *Start Tracking* during the intervention period.

Figure 1. Snapshots and overview of the LogPAL app features.

Structure of the PROs

A preliminary list of PROs was selected from the existing validated National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events library [26]. Individual items were selected by the study team for their relevance to symptoms and complications experienced by patients with HNC. To allow for iterative refinement, 2-week usability testing sessions were conducted using a purposive cohort of patients with HNC. At the end of the 2-week period, participants engaged in focus group interviews, which used the think-aloud method to elicit feedback regarding the perceived understanding of PRO questions and overall experience using the app. PROs were rated on a 5-tier scale ranging from "none" to "very severe." Questions were categorized into 10 common disease-specific symptoms, which included difficulty swallowing and chewing, dry mouth, loss of appetite, changes in taste, impaired speech, sores/pain in the mouth, overall pain, cough, nausea, and fatigue (Figure 2). Through this careful selection process, 2 versions were created. A weekly questionnaire (consisting of a series of 26 symptom-based questions) and a monthly questionnaire (weekly questionnaire with an additional 16 questions) were asked at the end of every month.

Figure 2. Examples of patient-reported outcome questions and illustrations of the relevant adverse effects. (Figure created on BioRender).

Data Collection

Questionnaires and app-generated data analytics were used to record participant sociodemographic characteristics at baseline, to measure engagement throughout the study period, and to perform postintervention analysis. Questionnaire data were entered into the Research Electronic Data Capture (REDCap) software and encrypted and stored in a secure server, which was selected and approved by the Office of the Chief Information Officer. REDCap's web-based app uses secure 2-factor web authentication, data logging, and encryption that ensures the security and confidentiality of private information for obtaining informed consent [27].

Measures

Patient Demographics

Participants completed an 18-item survey at enrollment. The survey included demographic characteristics about the participant's sex, racial and ethnic background, highest level of education, and marital status. Health-related information about the treatment type received, year of diagnosis, number of years since last treatment, smoking history, self-reported physical health and physical activity relative to peers, comorbidities, and clinic site of enrollment were also obtained. Self-reported physical health and physical activity relative to peers were determined using a 5-point Likert scale (excellent, very good, good, fair, and poor) and a 3-point Likert scale (more, about the same, less), respectively.

Follow-up Survey

After 8 weeks of using the app, participants were invited to complete a postintervention survey via email. The survey comprised of (1) the validated SUS, a standardized questionnaire commonly used to assess participants' perceptions of usability of an electronic system or device [28,29]; (2) a 5-point Likert response scale (1=strongly disagree to 5=strongly agree) acceptability questions; and (3) open-ended questions to prompt ideas for app improvement.

Outcome Measures

For this feasibility study, a priori criteria were defined as ≥80% retention rate (defined as the number of participants who enrolled and completed at least one questionnaire during the study period), ≥50% adherence rate (defined as the percentage of participants who completed all scheduled sessions), and mean SUS>68. This criterion was based on prior studies that assessed feasibility in applied intervention research [30-32] and self-management apps among cancer survivors [33]. To assess implementation outcomes such as recruitment and retention, rates were tabulated based on data collected from the research team tracking logs and summarized using a CONSORT diagram. Secondary exploratory outcomes tracked additional engagement metrics such as number of log-ins and frequency of participant interactions with other features [34]. Insights on different types of engagement indicators could provide opportunities for designing more engaging and clinically effective mHealth interventions [35].

Statistical Analysis

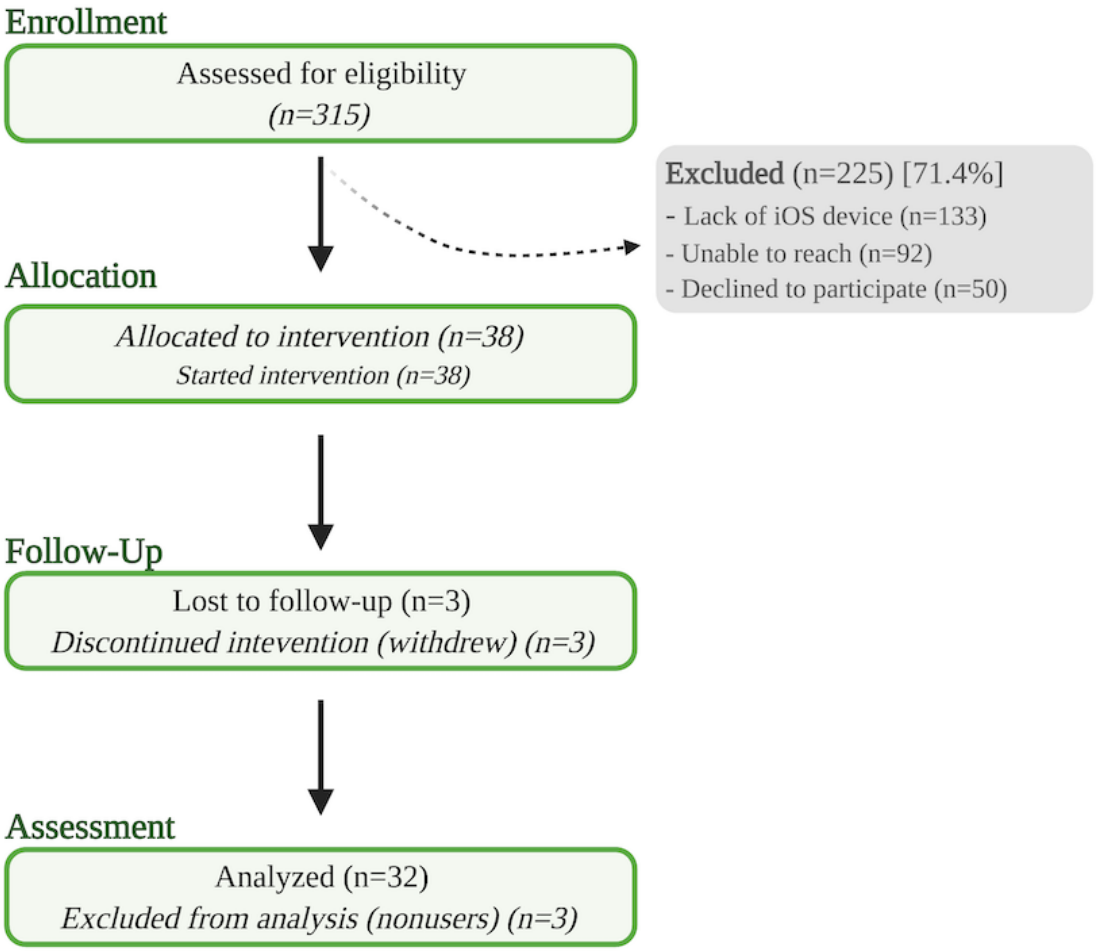
For this feasibility study, 50 participants were sought to review the app. This estimate was drawn on prior experiences accruing participants [36]. Taking into consideration a retention rate of 80% of recruited patients, it is estimated that 41 patients will remain in the study. All statistical analyses were conducted using Prism 8.4.0 (GraphPad Software) and SPSS version 25.0 (IBM Corp). Descriptive statistics was utilized to summarize participant characteristics and rates of engagement. Categorical variables were reported as frequencies (n) and percentages (%); continuous variables were reported as mean (SD) or median (IQR), that is, 25th-75th percentile, as needed. Nonparametric Mann-Whitney *U* test for 2 independent groups or the Kruskal-Wallis test for more than 2 groups were used for exploratory analyses of adherence scores (out of 16). Qualitative responses to open-ended questions were categorized using a thematic analysis, and relevant quotations were included to illustrate those themes. Results with *P* values less than .05 were considered to be statistically significant.

Results

Recruitment and Enrollment

A total of 315 patients were screened, of which 90 (28.6%) met the eligibility criteria. Of the 90 patients, 38 (42%) enrolled in the study. The primary reason for ineligibility was the lack of an iPhone device (133/315, 42.2%). The primary reason that eligible patients declined to participate was limited time and disinterest (50/90, 56%). Attrition rate was 16% (6/38) consisting of 3 dropouts and 3 classifieds as nonusers (participants who did not complete at least one PRO questionnaire during the study period) (Figure 3).

Figure 3. CONSORT flowchart for this study.



Participant Characteristics

Participants were predominantly males (26/32, 81%), Whites (25/32, 78%), and married (21/32, 66%), reflecting the regional demographics of HNC. The mean age of the participants was 57.8 (SD 12.3) years (range 24-77 years), and the median time from treatment completion to study enrollment was 10.8 months (range 1-23 months). The majority were college educated, including some college (7/32, 22%), bachelor's degree (9/32, 28%), or higher (9/32, 28%) (Table 1).

Table 1. Sociodemographic and clinical characteristics of the participants (N=32).

Characteristics[a,b]	Values, n (%)
Gender	
Female	6 (19)
Male	26 (81)
Ethnicity	
White	25 (78)
Black	3 (9)
Asian	2 (6)
Other	2 (6)
Marital status	
Single	5 (16)
Married	21 (66)
Divorced	5 (16)
Widow	1 (3)
Highest level of education	
High school/GED	7 (22)
Some college	7 (22)
Bachelors or equivalent	9 (28)
> Bachelor's	9 (28)
Type of treatment	
Radiation	6 (19)
Chemotherapy	14 (44)
Combination	12 (38)
Self-reported physical activity compared to others in their age group	
More	6 (19)
Less	6 (19)
The same	20 (63)
Self-reported physical health	
Poor/Fair	8 (25)
Good	14 (44)
Very Good/Excellent	10 (31)
Facility	
Site 1	26 (81)
Site 2	6 (19)

[a]Mean (SD) age in years = 57.8 (12.3) years.

[b]Mean (SD) time since last treatment = 10.8 (8.0) months.

App Use Tracking

Of the 32 participants eligible for analysis, 17 (53%) completed all of the scheduled sessions, 20 (63%) completed 75% or more of the sessions, and 25 (78%) completed at least 50% of the scheduled sessions by the end of the study period. Overall, 73.2% (375/512) of the questionnaires were completed (range 6.25%-100%) with participants opening the app 693 times over the course of 8 weeks.

Postintervention Survey

Usability

At the end of the study, 17 of the 32 participants (53%) reconsented to complete the SUS. The mean SUS score (95% CI) was 71.9 (64.3-79.5), which was an "acceptable" rating based on the standard SUS [28]. Further analysis of the subscales showed that the mean SUS (95% CI) learnability domain was 78.7 (71.2-86.1) and the mean (95% CI) usability domain was

70.2 (61.8-78.7). In the responses to the SUS questionnaire, 88% (15/17) found LogPAL "easy to use," 94% (16/17) felt that "most people could learn to use LogPAL very quickly," and 82% (14/17) felt "very confident using the system."

Acceptability

Among the participants, 76% (13/17) agreed that LogPAL was useful, with 59% (10/17) and 71% (12/17) agreeing with the frequency and length of the PROs, respectively. Additionally, 76% (13/17) of the participants agreed that they would recommend LogPAL to other cancer survivors. A total of 5 (29%) of the 17 participants gave additional feedback on their experience using LogPAL. In terms of what they liked, participants used the words "informative," "helpful," and "valuable." One patient stated, "I thought it was definitely helpful, it really let you understand what was going on with your body and it just wasn't you experiencing these symptoms." In terms of areas of improvements, participants mentioned the need for the app to be available to patients during or immediately after treatment, an alert tone as a reminder to update, and the ability to provide additional comments on changes. As one patient mentioned, the app "...does not allow for comments when an answer changes due to new variable in my situation, to explain why I changed my answer from previous weeks."

Exploratory Analyses

Additional App Features

Engagement with additional app features were reviewed among eligible participants. Findings showed that 32 (100%) used *View My* Progress, 27 (84%) used the *Self-Care Tips*, and 5 (16%) used the *Resource* feature. During the study period, the *View My Progress* feature was clicked 3445 times and the *Self-Care Tips* was clicked 79 times.

Relationship Among Covariates

Analysis of mean (SD) adherence scores and participant characteristics showed higher scores among those who self-reported conducting "more" physical activity (16.0 [SD 0.0]) than among those who self-reported "about the same" (12.2 [SD 4.9]) as others their age (P=.04). Similarly, the scores were higher among participants who self-reported their physical health as "very good"/"excellent" (14.5 [SD 4.4]) than among those who self-reported their physical health as "fair"/"poor" (7.8 [SD 2.1]) (P=.05). Lastly, the mean (SD) scores of the participants at clinic site 1 (13.2 [SD 4.5]) were higher than those of the participants at clinic site 2 (5.8 [SD 5.7]) (P=.003).

Discussion

Principal Results

To the best of our knowledge, this is the first study to evaluate the feasibility of a patient-facing mobile app that collects ePROs specifically for HNC survivors. Our results indicate that participants considered LogPAL as a feasible approach by meeting the a priori criteria. Our findings highlight the potential that mHealth apps have to improve symptom control and promote self-management of symptoms to improve health outcomes and quality of life. These findings are encouraging now more than ever, as the COVID-19 pandemic has placed greater importance on remote technology-enabled monitoring of high-risk patients through a digital platform. We believe that there is a greater need to develop an approach that allows HNC survivors to feel that their unique symptom needs are met and to have the ability to access straightforward information and valued material that address relevant issues.

Comparison With Prior Work

With the ubiquity of mobile phones in our society, there is a growing interest in and use of mobiles apps for patients with cancer to self-manage symptoms during treatment and those that persist into survivorship [37]. These apps have the potential to provide individually tailored self-management advice for different participants during their survivorship at home [38]. Furthermore, mHealth apps can increase patient engagement in their own recovery, provide better patient-provider communication, and flag patients at risk for readmission, thereby facilitating potential early interventions. Results from this study demonstrate that participant adherence (17/32, 53%) was congruent with that reported in previous studies that defined usage as over 50% of the participants that actually use an eHealth PRO intervention as intended. Other recent single-arm pilot trials have shown higher adherence rates; 1 mHealth ePRO study among patients with prostate cancer found that 86% (25/29) of the participants satisfactorily completed 60% of the weekly questions over a 3-month period [39]. Another pilot study of 10 patients with gynecological cancers who received palliative chemotherapy showed >70% adherence to daily smartphone surveys >4 days per week [40]. One explanatory factor for this inconsistency is the lack of clarity regarding evaluation methodologies, leading to substantial heterogeneity in the reported outcomes [41]. Recent attempts to meaningfully summarize indicators used in pilot studies have led to the creation of a lexicon of the most commonly used terms to identify effective app components [34,42]; however, a conceptually coherent framework is yet to be adapted. Another possibility to explain the lower engagement is the relationship between usage and descriptive variables such as age, marital status, years of education, and socioeconomic status [43]. Concomitantly, we did not find a significant relationship between ePRO usage among patients with HNC and the common descriptive mediating factors [44]. However, additional variables such as physical health, physical activity, and site of recruitment were not previously investigated. As such, these findings were unexpected and suggest that those who have self-perceived better physical health and those who are more physically active than those of their age may lead to higher engagement. It is important to note that there is strong evidence to show that usage of mobile phones and wearable devices increases physical activity and health, which significantly reduces cancer-related symptoms/side effects, leads to greater quality of life [45,46], as well as good retention rates and adherence. However, further investigation is needed to determine the relationships.

Additionally, clinic site 2 had notably lower retention and adherence rates compared to clinic site 1. Clinic site 2 is located in an ambulatory clinic within an urban hospital with 1 surgeon seeing those patients. In contrast, clinic site 1 is located in an outpatient ambulatory clinic in a suburban hospital campus, with 3 different surgeons seeing patients. The recruitment staff

at site 1 were different from those at site 2. Lastly, our results corroborated the need to develop cancer-supportive digital interventions that are interactive and tailored [47]. Integrating relevant information such as *Self Care Tips* as participants progress has been suggested as a way to maintain and enhance patients' experience of personal relevance [42]. In the future iterations, it might be beneficial to integrate an option for participants to include or supplant symptoms or concerns in PROs to make it more person-facing.

Limitations

This study is not without limitations. First, the app is currently only available on iPhone operating system (Apple iPhone and iPad) devices, which potentially caused selection bias (during enrollment) for age, ethnicity/race, education, and socioeconomic factors. Second, although the demographic characteristics of our cohort were slightly more heterogenous than those in most cancer pilot studies [48], this study was conducted in a single multi-site health system, thereby limiting its generalizability and not typifying those of the wider population. Third, patients who were enrolled in the study may be healthier than a general HNC follow-up population, thereby potentially biasing the engagement and usage figures to appear higher than they would be in a generalized population. Lastly, more rigorous recruitment protocols are needed to ascertain equitable retention rates. Additional updates and development of an Android version will help reduce these biases.

Conclusions

This study has provided preliminary evidence to suggest that the LogPAL app is a feasible and acceptable mHealth intervention that collects PROs to improve symptom management and proactively detect serious downstream complications among HNC survivors. The success of this feasibility study presents support for conducting a larger, multisite, randomized clinical trial to assess the efficacy of LogPAL with an active control and more heterogenous sample size.

Acknowledgments

The authors would like to thank all those at the Center for Health Innovations and Outcomes Research who helped with the development of the app as well as those who provided additional support with the recruitment of participants for this study.

Conflicts of Interest

None declared.

References

1. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin D, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer 2019 Apr 15;144(8):1941-1953 [FREE Full text] [doi: 10.1002/ijc.31937] [Medline: 30350310]
2. GBD 2016 Disease Injury Incidence Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1211-1259 [FREE Full text] [doi: 10.1016/S0140-6736(17)32154-2] [Medline: 28919117]
3. Murphy C, Galloway T, Handorf E, Egleston B, Wang L, Mehra R, et al. Survival Impact of Increasing Time to Treatment Initiation for Patients With Head and Neck Cancer in the United States. JCO 2016 Jan 10;34(2):169-178 [FREE Full text] [doi: 10.1200/jco.2015.61.5906]
4. Miller M, Shuman A, American HeadNeck Society's Committee on Survivorship. Survivorship in Head and Neck Cancer: A Primer. JAMA Otolaryngol Head Neck Surg 2016 Oct 01;142(10):1002-1008. [doi: 10.1001/jamaoto.2016.1615] [Medline: 27414519]
5. So W, Wong C, Choi K, Chan C, Chan J, Law B, et al. A Mixed-Methods Study of Unmet Supportive Care Needs Among Head and Neck Cancer Survivors. Cancer Nurs 2019;42(1):67-78 [FREE Full text] [doi: 10.1097/ncc.0000000000000542]
6. Lo Nigro C, Denaro N, Merlotti A, Merlano M. Head and neck cancer: improving outcomes with a multidisciplinary approach. CMAR 2017 Aug;Volume 9:363-371 [FREE Full text] [doi: 10.2147/cmar.s115761]
7. Cramer J, Burtness B, Le Q, Ferris R. The changing therapeutic landscape of head and neck cancer. Nat Rev Clin Oncol 2019 Nov;16(11):669-683. [doi: 10.1038/s41571-019-0227-z] [Medline: 31189965]
8. Jackson L, Ridner S, Deng J, Bartow C, Mannion K, Niermann K, et al. Internal Lymphedema Correlates with Subjective and Objective Measures of Dysphagia in Head and Neck Cancer Patients. J Palliat Med 2016 Sep;19(9):949-956 [FREE Full text] [doi: 10.1089/jpm.2016.0018] [Medline: 27227341]
9. Cramer J, Johnson J, Nilsen M. Pain in Head and Neck Cancer Survivors: Prevalence, Predictors, and Quality-of-Life Impact. Otolaryngol Head Neck Surg 2018 Nov;159(5):853-858. [doi: 10.1177/0194599818783964] [Medline: 29943677]
10. Rogers L, Courneya K, Robbins K, Malone J, Seiz A, Koch L, et al. Physical activity and quality of life in head and neck cancer survivors. Support Care Cancer 2006 Oct;14(10):1012-1019. [doi: 10.1007/s00520-006-0044-7] [Medline: 16538497]
11. Murphy B, Gilbert J, Ridner S. Systemic and global toxicities of head and neck treatment. Expert Rev Anticancer Ther 2007 Jul;7(7):1043-1053. [doi: 10.1586/14737140.7.7.1043] [Medline: 17627463]

12. Pakhomov SV, Jacobsen SJ, Chute CG, Roger VL. Agreement between patient-reported symptoms and their documentation in the medical record. Am J Manag Care 2008 Aug;14(8):530-539 [FREE Full text] [Medline: 18690769]
13. Hansen D, Larsen P, Holm L, Rottmann N, Bergholdt S, Søndergaard J. Association between unmet needs and quality of life of cancer patients: A population-based study. Acta Oncologica 2012 Dec 17;52(2):391-399 [FREE Full text] [doi: 10.3109/0284186x.2012.742204]
14. Faiola A, Papautsky E, Isola M. Empowering the Aging with Mobile Health: A mHealth Framework for Supporting Sustainable Healthy Lifestyle Behavior. Curr Probl Cardiol 2019 Aug;44(8):232-266 [FREE Full text] [doi: 10.1016/j.cpcardiol.2018.06.003] [Medline: 30185374]
15. Basch E, Deal A, Dueck A, Scher HI, Kris MG, Hudis C, et al. Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment. JAMA 2017 Jul 11;318(2):197-198 [FREE Full text] [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
16. Falchook A, Tracton G, Stravers L, Fleming ME, Snavely AC, Noe JF, et al. Use of mobile device technology to continuously collect patient-reported symptoms during radiation therapy for head and neck cancer: A prospective feasibility study. Adv Radiat Oncol 2016;1(2):115-121 [FREE Full text] [doi: 10.1016/j.adro.2016.02.001] [Medline: 28740878]
17. Benze G, Nauck F, Alt-Epping B, Gianni G, Bauknecht T, Ettl J, et al. PROutine: a feasibility study assessing surveillance of electronic patient reported outcomes and adherence via smartphone app in advanced cancer. Ann Palliat Med 2019 Apr;8(2):104-111 [FREE Full text] [doi: 10.21037/apm.2017.07.05] [Medline: 29156896]
18. Childes JM, Palmer AD, Fried-Oken MB. Communication Support Before, During, and After Treatment for Head and Neck Cancer. In: Doyle P, editor. Clinical Care and Rehabilitation in Head and Neck Cancer. Cham, Switzerland: Springer Nature; 2019:247-264.
19. Aapro M, Bossi P, Dasari A, Fallowfield L, Gascón P, Geller M, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. Support Care Cancer 2020 Oct;28(10):4589-4612 [FREE Full text] [doi: 10.1007/s00520-020-05539-1] [Medline: 32533435]
20. 325,000 mobile health apps available in 2017? Android now the leading mHealth platform. Research2Guidance. 2017. URL: https://research2guidance.com/325000-mobile-health-apps-available-in-2017/ [accessed 2020-09-29]
21. Zini E, Lanzola G, Quaglini S, Bossi P, Licitra L, Resteghini C. A pilot study of a smartphone-based monitoring intervention on head and neck cancer patients undergoing concurrent chemo-radiotherapy. Int J Med Inform 2019 Sep;129:404-412. [doi: 10.1016/j.ijmedinf.2019.06.004] [Medline: 31445284]
22. De Vito Dabbs A, Myers B, Mc Curry KR, Dunbar-Jacob J, Hawkins RP, Begey A, et al. User-centered design and interactive health technologies for patients. Comput Inform Nurs 2009;27(3):175-183 [FREE Full text] [doi: 10.1097/NCN.0b013e31819f7c7c] [Medline: 19411947]
23. 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]
24. Nacheva R. Standardization Issues of Mobile Usability. Int J Interact Mob Technol 2020 May 06;14(07):149 [FREE Full text] [doi: 10.3991/ijim.v14i07.12129]
25. Yen P, Bakken S. Review of health information technology usability study methodologies. J Am Med Inform Assoc 2012;19(3):413-422 [FREE Full text] [doi: 10.1136/amiajnl-2010-000020] [Medline: 21828224]
26. Dueck A, Mendoza T, Mitchell S, Reeve BB, Castro KM, Rogak LJ, National Cancer Institute PRO-CTCAE Study Group. Validity and Reliability of the US National Cancer Institute's Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). JAMA Oncol 2015 Nov;1(8):1051-1059 [FREE Full text] [doi: 10.1001/jamaoncol.2015.2639] [Medline: 26270597]
27. Harris P, Taylor R, Thielke R, Payne J, Gonzalez N, Conde J. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
28. Bangor A, Kortum P, Miller J. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 2008 Jul 30;24(6):574-594 [FREE Full text] [doi: 10.1080/10447310802205776]
29. Lewis JR, Sauro J. The Factor Structure of the System Usability Scale. In: Kurosu M, editor. Human Centered Design. San Diego, CA, USA: Springer; 2009:94-103.
30. Kazak A, Simms S, Alderfer M, Rourke MT, Crump T, McClure K, et al. Feasibility and preliminary outcomes from a pilot study of a brief psychological intervention for families of children newly diagnosed with cancer. J Pediatr Psychol 2005 Dec;30(8):644-655. [doi: 10.1093/jpepsy/jsi051] [Medline: 16260434]
31. Diefenbach M, Mohamed N, Butz B, Bar-Chama N, Stock R, Cesaretti J, et al. Acceptability and preliminary feasibility of an internet/CD-ROM-based education and decision program for early-stage prostate cancer patients: randomized pilot study. J Med Internet Res 2012 Jan 13;14(1):e6 [FREE Full text] [doi: 10.2196/jmir.1891] [Medline: 22246148]
32. Zhao J, Freeman B, Li M. Can Mobile Phone Apps Influence People's Health Behavior Change? An Evidence Review. J Med Internet Res 2016 Oct 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]

33. Melissant H, Verdonck-de Leeuw I, Lissenberg-Witte B, Konings I, Cuijpers P, Van Uden-Kraan C. 'Oncokompas', a web-based self-management application to support patient activation and optimal supportive care: a feasibility study among breast cancer survivors. Acta Oncologica 2018 Feb 16;57(7):924-934 [FREE Full text] [doi: 10.1080/0284186x.2018.1438654]
34. Pham Q, Graham G, Carrion C, Morita PP, Seto E, Stinson JN, et al. A Library of Analytic Indicators to Evaluate Effective Engagement with Consumer mHealth Apps for Chronic Conditions: Scoping Review. JMIR Mhealth Uhealth 2019 Jan 18;7(1):e11941 [FREE Full text] [doi: 10.2196/11941] [Medline: 30664463]
35. Zhang R, Nicholas J, Knapp A, Graham AK, Gray E, Kwasny MJ, et al. Clinically Meaningful Use of Mental Health Apps and its Effects on Depression: Mixed Methods Study. J Med Internet Res 2019 Dec 20;21(12):e15644 [FREE Full text] [doi: 10.2196/15644] [Medline: 31859682]
36. Leach C, Diefenbach M, Fleszar S, Alfano, Stephens RL, Riehman K, et al. A user centered design approach to development of an online self-management program for cancer survivors: Springboard Beyond Cancer. Psychooncology 2019 Oct;28(10):2060-2067. [doi: 10.1002/pon.5193] [Medline: 31379069]
37. Osborn J, Ajakaiye A, Cooksley T, Subbe C. Do mHealth applications improve clinical outcomes of patients with cancer? A critical appraisal of the peer-reviewed literature. Support Care Cancer 2020 Mar;28(3):1469-1479 [FREE Full text] [doi: 10.1007/s00520-019-04945-4] [Medline: 31273501]
38. Ruland C, Andersen T, Jeneson A, Moore S, Grimsbø GH, Børøsund E, et al. Effects of an internet support system to assist cancer patients in reducing symptom distress: a randomized controlled trial. Cancer Nurs 2013;36(1):6-17. [doi: 10.1097/NCC.0b013e31824d90d4] [Medline: 22495503]
39. Tran C, Dicker A, Leiby B, Gressen E, Williams N, Jim H. Utilizing Digital Health to Collect Electronic Patient-Reported Outcomes in Prostate Cancer: Single-Arm Pilot Trial. J Med Internet Res 2020 Mar 25;22(3):e12689 [FREE Full text] [doi: 10.2196/12689] [Medline: 32209536]
40. Wright A, Raman N, Staples P, Schonholz S, Cronin A, Carlson K, et al. The HOPE Pilot Study: Harnessing Patient-Reported Outcomes and Biometric Data to Enhance Cancer Care. JCO Clinical Cancer Informatics 2018 Dec(2):1-12 [FREE Full text] [doi: 10.1200/cci.17.00149]
41. Warrington L, Absolom K, Conner M, Kellar I, Clayton B, Ayres M, et al. Electronic Systems for Patients to Report and Manage Side Effects of Cancer Treatment: Systematic Review. J Med Internet Res 2019 Jan 24;21(1):e10875 [FREE Full text] [doi: 10.2196/10875] [Medline: 30679145]
42. Short C, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring Engagement in eHealth and mHealth Behavior Change Interventions: Viewpoint of Methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
43. Jiang Y, West BT, Barton DL, Harris MR. Acceptance and Use of eHealth/mHealth Applications for Self-Management Among Cancer Survivors. Stud Health Technol Inform 2017;245:131-135 [FREE Full text] [Medline: 29295067]
44. Head B, Keeney C, Studts J, Khayat M, Bumpous J, Pfeifer M. Feasibility and Acceptance of a Telehealth Intervention to Promote Symptom Management during Treatment for Head and Neck Cancer. J Support Oncol 2011 Jan 01;9(1):e1-e11 [FREE Full text] [doi: 10.1016/j.suponc.2010.12.006] [Medline: 21499540]
45. Granger C, Connolly B, Denehy L, Hart N, Antippa P, Lin KY, et al. Understanding factors influencing physical activity and exercise in lung cancer: a systematic review. Support Care Cancer 2017 Mar;25(3):983-999. [doi: 10.1007/s00520-016-3484-8] [Medline: 27900549]
46. Schaffer K, Panneerselvam N, Loh K, Herrmann R, Kleckner IR, Dunne RF, et al. Systematic Review of Randomized Controlled Trials of Exercise Interventions Using Digital Activity Trackers in Patients With Cancer. J Natl Compr Canc Netw 2019 Jan;17(1):57-63 [FREE Full text] [doi: 10.6004/jnccn.2018.7082] [Medline: 30659130]
47. Moradian S, Voelker N, Brown C, Liu G, Howell D. Effectiveness of Internet-based interventions in managing chemotherapy-related symptoms in patients with cancer: a systematic literature review. Support Care Cancer 2018 Feb;26(2):361-374. [doi: 10.1007/s00520-017-3900-8] [Medline: 28948360]
48. Unger J, Cook E, Tai E, Bleyer A. The Role of Clinical Trial Participation in Cancer Research: Barriers, Evidence, and Strategies. Am Soc Clin Oncol Educ Book 2016;35:185-198 [FREE Full text] [doi: 10.1200/EDBK_156686] [Medline: 27249699]

Abbreviations

ePRO: electronic patient-reported outcome
HNC: head and neck cancer
mHealth: mobile health
PRO: patient-reported outcome
REDCap: Research Electronic Data Capture
SUS: system usability scale

Edited by G Eysenbach; submitted 30.09.20; peer-reviewed by D Di Stasio, J Jabson; comments to author 18.11.20; revised version received 02.12.20; accepted 17.01.21; published 19.03.21.

Please cite as:
Teckie S, Solomon J, Kadapa K, Sanchez K, Orner D, Kraus D, Kamdar DP, Pereira L, Frank D, Diefenbach M
A Mobile Patient-Facing App for Tracking Patient-Reported Outcomes in Head and Neck Cancer Survivors: Single-Arm Feasibility Study
JMIR Form Res 2021;5(3):e24667
URL: https://formative.jmir.org/2021/3/e24667
doi:10.2196/24667
PMID:33739291

Original Paper

Gender Differences in Adolescent Sleep Disturbance and Treatment Response to Smartphone App–Delivered Cognitive Behavioral Therapy for Insomnia: Exploratory Study

Sophie H Li[1], MPsychol, PhD; Bronwyn M Graham[2], MPsychol, PhD; Aliza Werner-Seidler[1], MPsychol, PhD

[1]Black Dog Institute, University of New South Wales, Randwick, Australia

[2]School of Psychology, University of New South Wales, Kensington, Australia

Corresponding Author:
Sophie H Li, MPsychol, PhD
Black Dog Institute
University of New South Wales
Hospital Road
Randwick, 2031
Australia
Phone: 61 2 9382 4530
Email: s.h.li@blackdog.org.au

Abstract

Background: Insomnia and sleep disturbance are pervasive and debilitating conditions affecting up to 40% of adolescents. Women and girls are at greater risk of insomnia, yet differences in treatment responsiveness between genders have not been adequately investigated. Additionally, while women report greater symptom severity and burden of illness than men, this discrepancy requires further examination in adolescents.

Objective: The purpose of this study was to examine gender differences in sleep symptom profiles and treatment response in adolescents.

Methods: Digital cognitive behavioral therapy for insomnia (CBT-I) treatment responsiveness, as indexed by changes in Insomnia Severity Index (ISI) and Global Pittsburgh Sleep Quality Index (PSQI) scores, was compared in boys and girls (aged 12-16 years; N=49) who participated in a pilot evaluation of the Sleep Ninja smartphone app. Gender differences in self-reported baseline insomnia symptom severity (ISI), sleep quality (PSQI), and sleep characteristics derived from sleep diaries were also examined.

Results: Compared with boys, we found that girls reported greater symptom severity (P=.04) and nighttime wakefulness (P=.01 and P=.04) and reduced sleep duration (P=.02) and efficiency (P=.03), but not poorer sleep quality (P=.07), more nighttime awakenings (P=.16), or longer time to get to sleep (P=.21). However, gender differences in symptom severity and sleep duration were accounted for by boys being marginally younger in age. Treatment response to CBT-I was equivalent between boys and girls when comparing reductions in symptom severity (P=.32); there was a trend showing gender differences in improvements in sleep quality, but this was not statistically significant (P=.07).

Conclusions: These results demonstrate the presence of gender differences in insomnia symptoms and severity in adolescents and suggest further research is required to understand gender differences in insomnia symptom profiles to inform the development of gender-specific digital interventions delivered to adolescents.

(JMIR Form Res 2021;5(3):e22498) doi:10.2196/22498

KEYWORDS

insomnia; gender differences; adolescents; sleep disturbance; sleep quality; sleep; gender; digital interventions

Introduction

Insomnia, defined as difficulty initiating or maintaining sleep with an associated impairment in daytime functioning [1], is a pervasive problem affecting 4%-18.5% of adolescents, in addition to 40% experiencing subthreshold symptoms and insufficient sleep [2-5]. Insomnia and disturbed and insufficient sleep impact adolescents' academic, social, emotional, and behavioral development [6]. Difficulties with sleep also often

accompany mental health disorders, such as depression, and contribute to their maintenance and burden of illness [7]. Moreover, there is a well-established link between sleep disturbance and the onset of depression in adolescents, emphasizing the need for effective interventions to ensure adequate sleep duration and quality in this population [8-10].

The exact cause of insomnia has not yet been identified but is likely to involve complex interactions between genetic, neurological, behavioral, cognitive, and emotional factors (eg, [11,12], particularly during adolescence [13]. One factor identified to be a likely contributing candidate is gender. Girls and women are 50% more likely to develop insomnia [14,15] compared with adolescent boys and men. One recent large-scale study examining insomnia symptoms in boys and girls aged 6-17 years found that pubertal maturation in girls was associated with increased prevalence and severity of insomnia symptoms, indicating that gender discrepancies in insomnia emerge at puberty [16]. Women also report greater symptom severity and perceived burden of illness [17] and more difficulty falling asleep and nighttime wakefulness than men [18]. Only one study has investigated gender discrepancies in sleep characteristics in adolescents. This study recruited a large cohort of more than 10,000 Norwegian adolescents with a mean age of 17 years and found that girls reported taking longer to get to sleep (sleep onset latency [SOL]: 51 minutes vs 43 minutes), experienced greater durations of wakefulness after sleep onset (wake after sleep onset [WASO]: 17 minutes vs 12 minutes), and reported a greater perceived need for sleep (8 hours 43 minutes vs 8 hours 26 minutes) [4].

Cognitive behavioral therapy for insomnia (CBT-I) is the gold-standard psychological treatment for insomnia and is effective for adults [19-22] and adolescents [23]. CBT-I consists of two main components. The first is cognitive therapy, which provides the individual with strategies to reevaluate catastrophic beliefs about sleep and manage sleep-inhibiting cognitive processes, such as worry and rumination. The second is the behavioral component, which encourages the application of behaviors associated with improved sleep, such as a consistent sleep routine and the elimination of daytime naps (for review, see [1]). CBT-I can be delivered in its traditional face-to-face format [20,24] or digitally as a web-based program [23] or smartphone app [25]. Face-to-face and digitally delivered CBT-I have been found to be equally beneficial [23]. However, digitally delivered CBT-I has the benefit of addressing several barriers to treatment identified in adolescents, such as accessibility and affordability, and allowing a degree of privacy to reduce concerns related to stigma, making digital delivery a particularly good treatment option for this age group. Regardless of delivery format, there is significant variability between individuals in treatment response. One potential source of treatment response variability may be the aforementioned gender differences in sleep symptom profiles. However, despite well-documented gender differences in insomnia prevalence, gender differences in insomnia treatment *responsiveness* have not been adequately investigated in either adults or adolescents. Specifically, only 2 studies have intentionally investigated the contribution of gender to treatment outcomes. One study examined the extent to which several demographic factors, including income, race, gender, age, and education, moderated treatment response in 358 adults with insomnia [26]. Neither gender nor any of the remaining demographic factors were significant moderators of CBT-I treatment response as measured by the Insomnia Severity Index (ISI). In a smaller study, Lami et al [27] investigated differential CBT-I treatment response in adult men (n=13) and women (n=15) with fibromyalgia (a condition characterized by persistent pain that disproportionately affects women). Consistent with Cheng et al [26], there were no gender differences in treatment response. We are unaware of any studies that have explicitly explored gender differences in treatment response in adolescents. Understanding whether gender differences in symptom profiles and treatment response are present during adolescence, when symptoms often first emerge [28], could assist in improving CBT-I outcomes by tailoring them to the needs of the individual's gender, rather than applying a "one size fits all" approach. This knowledge is particularly important when designing digitally delivered interventions as a face-to-face therapist is not present to adjust the intervention according to the individual's needs.

The purpose of this study was primarily to investigate possible gender differences in self-reported insomnia symptom severity, sleep quality, and sleep characteristics in adolescents; the secondary purpose was to examine gender differences in CBT-I treatment response in terms of improvements in self-reported insomnia symptoms and sleep quality. Baseline and posttreatment scores on the ISI and Pittsburgh Sleep Quality Index (PSQI) and baseline-only sleep diary variables were compared between girls and boys aged 12-16 years that participated in a published single-group, pre-post pilot trial evaluating the efficacy of a digital CBT-I intervention for adolescent insomnia [29]. Sleep diary variables were included in addition to the PSQI to provide a report of current sleep characteristics (eg, nighttime wakefulness, difficulty falling asleep). Based on the adult literature, we predicted girls would report worse symptom severity, sleep quality, and sleep characteristics consistent with sleep disturbance compared to boys. Based again on the adult literature, we did not anticipate finding significant gender differences in CBT-I treatment response in an adolescent sample.

Methods

Participants

This study used data from 49 young people who participated in a pilot trial evaluating Sleep Ninja, a CBT-I smartphone app for young people [29]. Three hundred individuals expressed interest in participating in the Sleep Ninja pilot trial. Of those, 60 reported meeting inclusion criteria and provided written informed consent, along with parental or carer consent, before being formally screened for eligibility. Of these, 10 were not enrolled in the trial, including 4 who did not meet inclusion criteria and 6 who withdrew prior to the trial. Withdrawal reasons included change of mind (n=1), lack of time (n=2), and finding participation a chore (n=1), and 1 person did not provide a reason. Of the remaining 50 participants, 1 was excluded from this study because they did not disclose gender, and 1 boy and 1 girl completed fewer than 6 sleep diaries so were not included

in baseline sleep variable comparisons. Gender was assessed via an item on the demographic survey, which asked "What is your gender?", with possible responses being male, female, or other. Participants were not excluded based on reported gender. Participants were 33 girls and 16 boys aged between 12 and 16 years who met the study inclusion criteria of at least mild insomnia symptoms (operationalized as endorsement of at least one of the first three items on the ISI: difficulty falling asleep, difficulty staying asleep, or waking up too early) and access to a smartphone, internet, and a valid email address.

Thirty-three adolescents completed postintervention assessment. Two were not invited to complete the postintervention assessment because they did not download the Sleep Ninja app. The attrition rates from preintervention to postintervention assessment were similar among boys (6/16, 38% attrition) and girls (10/33, 30% attrition).

Measures

Insomnia Severity Index (ISI)

The ISI is a 7-item self-report measure of insomnia symptoms over the previous 2 weeks that is psychometrically sound [30]. Responses are reported on a Likert scale from 0 to 4, producing total scores of 0-28. Cutoff scores are as follows: 0-7 reflects no clinically significant insomnia, 8-14 indicates subthreshold insomnia, 15-21 suggests moderate severity insomnia, and 22-28 indicates severe insomnia [30]. The ISI was designed for use in adults but has been widely administered to, and validated in, adolescent samples [31-33]. In one adolescent validation study, reliability was strong (Cronbach α=.83), and test-retest reliability was acceptable (r=0.79) [33].

Pittsburgh Sleep Quality Index (PSQI)

The PSQI is a 19-item self-report scale that is widely used to assess usual sleep habits and experiences over the preceding month. It has been validated in adolescent samples, with strong internal consistency (α=.72) and test-retest reliability over a 6-week period (r=0.81) [34]. There are 7 subscales, which are sleep quality, sleep latency, sleep duration, habitual sleep efficiency (SE), sleep disturbances, use of sleeping medications, and daytime dysfunction [35]. Each component is scored from 0 (no difficulty) to 3 (severe difficulty), and the components are summed to obtain a Global PSQI score ranging from 0 to 21 [34]. The Global PSQI score was used in this study because it is a valid representation of self-reported sleep quality [21,36,37].

Sleep Diary

The 10-item sleep diary was developed by the research team by incorporating the questions from the Consensus Sleep Diary [38], with the addition of 2 questions regarding daytime naps and use of sleep medication. Participants answered 10 questions, which included bedtime, time taken to fall asleep (SOL), number and duration of nighttime awakenings (number of awakenings; NWAK), duration of wakefulness after sleep onset (WASO), time of final awakening, time participants got out of bed for the day, subjective sleep quality, how refreshed participants felt on awakening, duration of any daytime naps, and use of sleep medication. Sleep diary variables obtained via self-report have been found to be consistent with more objective (eg, polysomnography) measures of sleep characteristics [39-41] and were included in this study to provide as objective a report of sleep characteristics as possible in the absence of objective measures. Further details regarding the sleep diary can be found in Werner-Seidler et al [42].

Intervention: Sleep Ninja

The Sleep Ninja app and the process employed in its design are described in detail by Werner-Seidler et al [42,43]. It is derived from CBT-I and functions as a fully automated smartphone app consisting of six sequential lessons (each taking 5-10 minutes to complete) delivering core CBT-I strategies: psychoeducation, stimulus control, sleep hygiene, and sleep-focused cognitive therapy. The intervention also includes a sleep tracking function, recommended bedtimes based on sleep guidelines, reminders to start a wind-down routine each night, a series of sleep tips, and general information about sleep. The lessons are delivered via a chatbot feature where the sleep ninja acts as a sleep coach. The chatbot feature contains forced choice responses allowing the "chat" to be responsive to the input of the user by personalizing information and recommendations based on the selections and sleep profile recorded by the user. Upon completion of a lesson and 3 nights of sleep tracking (out of a 7-night period), users level up and reach their next "belt."

Procedure

This study was conducted with written consent from each participant and their parent or carer, and all procedures were carried out in accordance with the Declaration of Helsinki and approved by the University of New South Wales Human Research Ethics Committee (approval number: HC16702). Participants were recruited through media and social media channels, including the Black Dog Institute's website and paid Facebook advertisements that targeted potential participants and their parents between April and June 2017. After providing their consent and parental consent via the study website, participants were invited to complete screening questionnaires to verify study eligibility. They then completed demographic and baseline questionnaires (delivered online and described in detail by Werner-Seidler et al [29]): ISI [30] and PSQI [34] to measure symptom severity and sleep quality, respectively, and a 7-day sleep diary to provide a current report of sleep characteristics. We included measures of symptom severity (reflecting symptoms of disorder, including functional impact) and sleep quality (a more general construct related to sleep quality and duration) because, despite typically co-occurring, a subclinical population may experience poor sleep quality in the absence of insomnia symptoms. The Patient Health Questionnaire–Adolescent Version and Generalized Anxiety Disorder–7 were also administered as part of the larger study with outcomes reported by Werner-Seidler et al [29]. Following completion of the sleep diary, participants gained access to Sleep Ninja for 6 weeks before completing the same battery of questionnaires (posttreatment), resulting in a 6-week interval between baseline and posttreatment assessments. Lesson completion was automatically recorded by the app.

Statistical Analysis

To determine gender differences in symptom severity and sleep quality, two-tailed independent-samples *t* tests were conducted on baseline ISI and Global PSQI scores, respectively. To determine possible differences between girls and boys on sleep diary variables, summary scores for each variable at baseline, including total sleep time (TST), total wake time (TWT), SOL, WASO, NWAK, and SE, were derived by averaging the 7 baseline sleep diary entries and compared using two-tailed independent-samples *t* tests (participants completing fewer than 6 entries were excluded to ensure reliability [44]). Gender differences in demographics between boys and girls were examined using Fisher exact tests.

Since there was a trend in differences in age between boys and girls, though these differences were not significant, post hoc linear regressions were conducted for each sleep outcome, entering gender (coded as 0=male, 1=female) and age as predictors to determine whether gender predicted insomnia symptom severity, sleep quality, and sleep characteristics when controlling for age.

To examine whether gender predicted declines in ISI and PSQI from pretreatment to posttreatment, we conducted hierarchical linear mixed models using SPSS (version 25; IBM Corp) with restricted maximum likelihood estimation. This modelling approach handles missing data by incorporating all available data from each participant into the analysis. Data from the pretreatment and posttreatment time points were grouped by participant. Participant was specified as a random factor, and time, gender, and age were specified as fixed factors. The outcome variables were ISI and PSQI at both time points, with gender as the predictor. Time was coded as 0 for pre and 1 for post, and gender was coded as 0 for male and 1 for female participants. We specified two sets of multilevel models for each outcome variable. In the first model we entered time, gender, and age as predictors, and in the second model we added the two-way interactions between time and gender, time and age, and age and gender. For all models, we specified a random intercept, which allows participants' mean levels of each outcome to vary. Intraclass correlations (ICCs) were calculated for ISI and PSQI based on intercept-only models. ICC values indicate the amount of variance accounted for by within-person variability. Convention suggests that variables with ICC values greater than 0.10 show sufficient dependency to be analyzed with multilevel modelling [45]. ICCs were 0.42 for ISI and 0.81 for PSQI.

Results

Gender Differences in Baseline Symptoms

Baseline score comparisons are presented in Table 1. Boys and girls did not significantly differ in age or city residence; however, more girls than boys were currently using medication and receiving mental health treatment. Girls had significantly higher ISI scores at baseline (t_{47}=−2.13, *P*=.04; Cohen *d*=0.60), suggesting significantly greater symptom severity compared to boys. Differences in PSQI scores did not reach significance (t_{47}=−1.90, *P*=.07; *d*=0.57), suggesting negligible differences between girls and boys in sleep quality. Girls reported significantly greater WASO and TWT and significantly reduced TST and sleep efficiency compared to boys (t_{44}=−2.72, *P*=.01, *d*=0.54; t_{45}=−2.12, *P*=.04, *d*=0.24; t_{45}=2.34, *P*=.02, *d*=0.68; t_{45}=2.28, *P*=.03, *d*=0.30, respectively) but no difference in SOL or NWAK (*P*=.21 and *P*=.16, respectively).

Post hoc linear regressions found that when controlling for age, gender predicted WASO (t_{44}=2.07, *P*=.04; f^2=0.09), TWT (t_{45}=2.39, *P*=.02; f^2=0.13), and SE (t_{45}=−2.18, *P*=.03; f^2=0.11). In contrast, when controlling for age, gender did not predict baseline ISI (t_{47}=1.42, *P*=.16; f^2=0.04), PSQI (t_{47}=1.31, *P*=.20' f^2=0.04), SOL (t_{45}=−0.57, *P*=.23; f^2=0.04), or NWAK (t_{45}=0.88, *P*=.39; f^2=0.02) or TST (t_{45}=1.45, *P*=.15; f^2=0.04). Notably, two variables identified as showing gender differences in the *t* tests, ISI and TST, were not significant predictors in the linear regressions, suggesting baseline gender differences in these variables are accounted for by boys being marginally younger in age.

Of relevance to the current study, there were no differences between boys and girls in the number of Sleep Ninja lessons completed (*P*=.42). Information regarding uptake and adherence to the intervention can be found in Werner-Seidler et al [42].

Table 1. Participant characteristics and baseline sleep symptoms.

Characteristic	Male	Female	*P* value
Age (years), mean (SD)	13.70 (1.04)	14.47 (1.51)	.07
Reside in city, n (%)	13 (81)	30 (91)	.38
MH[a] treatment,[b] n (%)	0 (0)	13 (39)	.004
Medication use,[c] n (%)	1 (6)	13 (39)	.02
ISI,[d] mean (SD)	7.90 (4.63)	9.87 (5.00)	.04
Global PSQI,[e] mean (SD)	7.40 (3.78)	7.91 (3.90)	.07
SOL,[f,g] mean (SD)	0.82 (0.75)	0.73 (0.54)	.21
NWAK,[f,h] mean (SD)	1.50 (2.31)	0.65 (0.69)	.16
WASO,[f,i] mean (SD)	0.12 (0.18)	0.24 (0.26)	.01
TWT,[f,j] mean (SD)	1.13 (0.88)	1.40 (1.32)	.04
TST,[f,k] mean (SD)	8.74 (0.65)	8.14 (1.07)	.02
SE[l] (%), mean (SD)	88.70 (8.41)	85.80 (10.74)	.03

[a]MH: mental health.

[b]Predominantly psychology.

[c]Predominantly antidepressants, hormonal contraceptives, and melatonin.

[d]ISI: Insomnia Severity Index.

[e]PSQI: Pittsburgh Sleep Quality Index.

[f]In hours, derived by averaging the 7 baseline sleep diary entries.

[g]SOL: sleep onset latency.

[h]NWAK: number of awakenings.

[i]WASO: wake after sleep onset.

[j]TWT: total wake time.

[k]TST: total sleep time.

[l]SE: sleep efficiency.

Gender Differences in Treatment Response

Change in ISI and PSQI scores from pretreatment to posttreatment for each participant, grouped by gender, is represented in Figure 1. Results from the mixed-model analysis are presented in Table 2. For insomnia symptom severity, time was a significant predictor of ISI in both models. The coefficient was negative, indicating that scores on the ISI decreased from pretreatment to posttreatment. Gender did not reach significance in either model. The gender by time interaction in model 2 was not significant. Age was not a moderator of any effect.

For sleep quality, time was a significant predictor of PSQI in model 1. The coefficient was negative, indicating that scores on the PSQI decreased from pretreatment to posttreatment. However, time did not remain a significant predictor of PSQI in model 2. Gender did not reach significance in either model. There was a trend in the gender by time interaction in model 2, but it was not significant. Age was not a moderator of any effect.

Figure 1. Pretreatment and posttreatment scores for the ISI (A) and PSQI (B) for each participant, grouped by gender. ISI: Insomnia Severity Index. PSQI: Pittsburgh Sleep Quality Index.

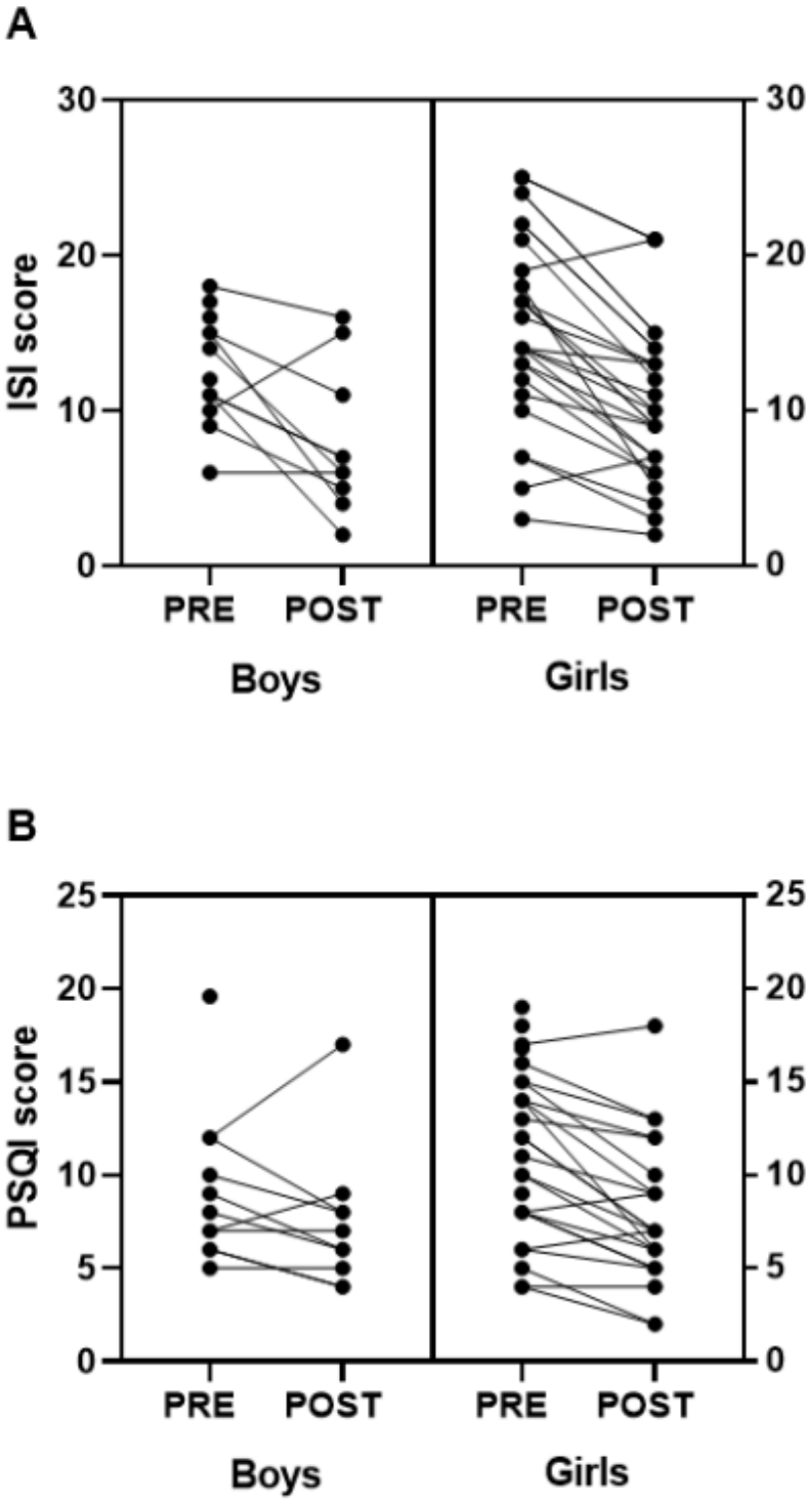


Table 2. Results from the hierarchical linear mixed models.

Characteristic	Parameter estimate	SE	95% CI	*P* value
ISI[a]				
Model 1				
Intercept	1.94	6.21	−10.55 to 14.43	.76
Time	−4.83	0.63	−6.1 to −3.56	<.001
Gender	1.6	1.32	−1.06 to 4.26	.23
Age	0.8	0.46	−0.13 to 1.73	.09
Model 2				
Intercept	0.72	14.61	−28.66 to 30.12	.96
Time	−15.44	6.43	−28.5 to −2.38	.02
Gender	8.13	16.15	−24.39 to 40.66	.62
Age	0.87	1.1	−1.34 to 3.09	.43
Time×gender	−1.38	1.4	−4.13 to 1.37	.32
Time×age	0.84	0.47	−0.13 to 1.8	.09
Gender×age	−0.45	1.2	−2.87 to 1.97	.71
PSQI[b]				
Model 1				
Intercept	1.25	5.82	−10.49 to 12.99	.83
Time	−2.13	0.45	−3.01 to −1.21	<.001
Gender	1.02	1.24	−1.43 to 3.57	.39
Age	0.61	0.43	−2.66 to 1.48	.17
Model 2				
Intercept	8.87	13.63	−18.61 to 36.34	.52
Time	−2.06	4.58	−11.42 to 7.29	.66
Gender	−7.77	15.15	−38.32 to 22.78	.61
Age	0.004	1.03	−2.06 to 2.07	>.99
Time×gender	−1.82	0.96	−3.79 to 0.16	.07
Time×age	0.08	0.34	−0.6 to 0.77	.80
Gender×age	0.71	1.13	−1.57 to 2.98	.54

[a]ISI: Insomnia Severity Index.

[b]PSQI: Pittsburgh Sleep Quality Index.

Discussion

This study compared self-reported insomnia symptom severity, sleep quality, sleep characteristics, and CBT-I treatment response between adolescent girls and boys with at least mild symptoms of insomnia. The results supported some of our predictions by showing that girls more often than boys reported sleep characteristics consistent with disturbed sleep, including increased nighttime wakefulness and reduced sleep efficiency. However, contrary to predictions, we found no differences in self-reported sleep quality on the PSQI or in difficulty falling asleep, and gender differences in symptom severity and total sleep duration were accounted for by boys being marginally younger in age. It is not immediately clear why girls reported sleep characteristics consistent with disturbed sleep but not poorer sleep quality on the PSQI or insomnia symptoms on the ISI (when controlling for age) compared to boys. One possible explanation is the nonclinical sample examined in the current study. Individuals with milder symptoms may experience greater symptom fluctuation consistent with individuals in the early stage of disorder development (prodromal phase; eg, [46]), which may be more effectively captured in state-based measures (eg, daily sleep diary) compared to trait-based measures reliant on self-report (eg, PSQI and ISI measure sleep parameters over the past 4 and 2 weeks, respectively).

Increased wakefulness after sleep onset in girls in the current study is consistent with previous reports showing women and adolescent girls more often report sleep characteristics consistent with increased nighttime wakefulness [4,18]. However, unlike these previous studies, we did not specifically find that girls

reported more difficulty falling asleep. The previous studies were conducted in adults and in a cohort of adolescents with an average age of 17 years (compared to 14.5 years in the current study), which might account for these differences. Regardless, our results indicate that gender differences in sleep characteristics in older adolescents and adults are at least partially present in younger adolescents (in addition to reduced sleep efficiency in girls). One possible explanation for increased sleep disturbance in girls is the propensity for adolescent and adult females to engage in unhelpful cognitive processes, such as rumination and worry, relative to males [47,48]. Rumination and worry play a central role in sleep disturbance in a range of contexts, including in individuals with sleep disturbances [49], individuals with other mental disorders [50], and healthy individuals [51]. For example, Harvey's [52] cognitive model of insomnia postulates that excessive worry about sleep and the consequences of poor-quality sleep triggers autonomic arousal, emotional distress, and excessive monitoring of threats to sleep quality (eg, indicators of insufficient sleep or poor functioning during the day), which ultimately results in real deficits in sleep. Thus, established gender differences in worry make these factors that are critical to the development and maintenance of sleep disturbance good candidates to explain gender differences shown in this study. Future studies could employ measures of rumination and worry alongside sleep measures to determine the extent to which they moderate gender differences in symptoms.

Adolescence has been identified as a period of substantial change in terms of biopsychosocial development, which has been attributed to changing sleep patterns in both genders in this age group [13]. However, pubertal maturation seems to be particularly significant for girls in regard to sleep. Several studies have found that in girls, but not boys, the onset of pubertal maturation is a significant risk factor in the development of insomnia and is associated with increased symptom severity [16,17]. Therefore, another possible mechanism underlying gender differences in pretreatment symptoms is hormonal fluctuations associated with the onset of puberty in girls. The menstrual cycle is a major source of hormonal fluctuations in estradiol and progesterone over a 4-week period. Estradiol and progesterone levels are initially low (follicular phase), with a midcycle peak in estradiol indicating ovulation, followed by a rise in both estradiol and progesterone during the second half of the cycle before declining to the initial low levels (luteal phase). In support of this notion, adult women self-report poorer sleep quality during periods of hormonal flux, including the luteal phase of the menstrual cycle, perinatal period and menopause [17,18,53]; however, the association between sex hormones and more objectively measured sleep variables (eg, via polysomnography) is more contentious [54]. The nature of the association between pubertal maturation and insomnia prevalence and disturbed sleep characteristics, including the potential role of fluctuating sex hormones, requires further examination.

Despite pretreatment differences in symptom severity, boys and girls responded comparably to CBT-I, as evidenced by comparable reductions in symptom severity. On the one hand, it is encouraging that despite greater pretreatment symptom severity, girls gain the same benefits from treatment as boys. On the other hand, equivalent treatment-elicited symptom reduction would suggest gender differences in symptom severity at baseline are maintained posttreatment, which we speculate may render girls more vulnerable to relapse than boys. More work in this area is required to determine if differences remain between boys' and girls' posttreatment symptom levels. We were unable to adequately analyze this here due to participant attrition and therefore an incomparable sample at baseline and posttreatment. Future studies could also examine if there are gender differences in responsiveness to certain treatment elements. For example, as girls engage in more rumination and worry [47,48], which are perpetuating features of insomnia (as discussed above), they may benefit from extending the cognitive aspect of CBT-I to reduce the frequency of these cognitive maintaining factors and thus limit ongoing sleep disturbance. These ideas require further investigation to refine gender-specific delivery of CBT-I.

Our study had several limitations, namely the lack of a nontreatment control group to exclude the possibility that participants responded to some nontherapeutic aspect of CBT-I. Our study also relied on subjectively reported sleep diary variables. The inclusion of polysomnography or actigraphy would help establish gender differences in objectively measured sleep parameters. However, the presence of gender differences on some sleep characteristics but not others in the current study suggests that our results are not merely driven by a bias for girls to overreport (or boys to underreport) symptoms in general. A disproportionate number of girls participated in this study. This is consistent with male to female participant ratios reported in similar studies evaluating CBT-I in adolescents that use a range of recruitment strategies (eg, community recruitment via a study website [23] and outpatient mental health clinics [55]). This suggests gender disparity in participation is representative of the increased prevalence of sleep problems in girls compared to boys in the community. The small sample size, particularly the small sample of boys, should also be taken into consideration when interpreting the current results. It is possible that a lack of power to detect significant effects concealed true differences between boys and girls in treatment response (particularly in reductions in sleep quality, in which there was a trend, though it was not statistically significant) and baseline sleep characteristics. Future studies could ensure adequate power to detect possible gender differences by recruiting larger samples with equal numbers of boys and girls.

A further limitation is that the current study did not assess pubertal maturation, which has previously been shown to be associated with gender differences in insomnia [16]. This information would have provided greater insight into the potential role of hormonal fluctuation in girls' experience of sleep disturbance. Despite these limitations, our study has demonstrated the importance of continued investigation into understanding the nature of gender differences in insomnia and sleep disturbance in adolescents. This knowledge could be applied in future research aimed at refining insomnia interventions according to the unique needs of each gender, especially in digital interventions as this could inform

onboarding processes to facilitate the personalization of the intervention to the benefit of both genders.

Acknowledgments

This study was supported by a grant awarded to AW-S and the Black Dog Institute by the Corella Foundation. The funding source had no involvement in any aspect of the research study.

Authors' Contributions

AW-S developed the study concept and design and collected data. SHL and BMG performed the data analysis. SHL drafted the paper, and BMG and AW-S provided critical revisions. All authors approved the final version of the paper for submission.

Conflicts of Interest

None declared.

References

1. Dewald-Kaufmann J, de Bruin E, Michael G. Cognitive Behavioral Therapy for Insomnia (CBT-i) in School-Aged Children and Adolescents. Sleep Med Clin 2019 Jun;14(2):155-165. [doi: 10.1016/j.jsmc.2019.02.002] [Medline: 31029183]
2. Carskadon MA. Patterns of sleep and sleepiness in adolescents. Pediatrician 1990;17(1):5-12. [Medline: 2315238]
3. Chung K, Kan KK, Yeung W. Insomnia in adolescents: prevalence, help-seeking behaviors, and types of interventions. Child Adolesc Ment Health 2014 Feb;19(1):57-63. [doi: 10.1111/camh.12009] [Medline: 32878360]
4. Hysing M, Pallesen S, Stormark KM, Lundervold AJ, Sivertsen B. Sleep patterns and insomnia among adolescents: a population-based study. J Sleep Res 2013 Oct;22(5):549-556 [FREE Full text] [doi: 10.1111/jsr.12055] [Medline: 23611716]
5. Ohayon MM, Roberts RE, Zulley J, Smirne S, Priest RG. Prevalence and patterns of problematic sleep among older adolescents. J Am Acad Child Adolesc Psychiatry 2000 Dec;39(12):1549-1556. [doi: 10.1097/00004583-200012000-00019] [Medline: 11128333]
6. Roberts RE, Roberts CR, Duong HT. Sleepless in adolescence: prospective data on sleep deprivation, health and functioning. J Adolesc 2009 Oct;32(5):1045-1057 [FREE Full text] [doi: 10.1016/j.adolescence.2009.03.007] [Medline: 19361854]
7. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). Washington, DC: American Psychiatric Association Publishing; 2013.
8. Liu X, Buysse DJ, Gentzler AL, Kiss E, Mayer L, Kapornai K, et al. Insomnia and hypersomnia associated with depressive phenomenology and comorbidity in childhood depression. Sleep 2007 Jan;30(1):83-90. [doi: 10.1093/sleep/30.1.83] [Medline: 17310868]
9. Lovato N, Gradisar M. A meta-analysis and model of the relationship between sleep and depression in adolescents: recommendations for future research and clinical practice. Sleep Med Rev 2014 Dec;18(6):521-529. [doi: 10.1016/j.smrv.2014.03.006] [Medline: 24857255]
10. Marver JE, McGlinchey EA. Sex differences in insomnia and risk for psychopathology in adolescence. Curr Opin Psychol 2020 Aug;34:63-67. [doi: 10.1016/j.copsyc.2019.09.004] [Medline: 31655365]
11. Sivertsen B, Pallesen S, Stormark KM, Bøe T, Lundervold AJ, Hysing M. Delayed sleep phase syndrome in adolescents: prevalence and correlates in a large population based study. BMC Public Health 2013 Dec 11;13(1):1163 [FREE Full text] [doi: 10.1186/1471-2458-13-1163] [Medline: 24330358]
12. Morin C, Drake CL, Harvey AG, Krystal AD, Manber R, Riemann D, et al. Insomnia disorder. Nat Rev Dis Primers 2015 Sep 03;1(1):15026-15018. [doi: 10.1038/nrdp.2015.26] [Medline: 27189779]
13. Carskadon MA. Sleep in adolescents: the perfect storm. Pediatr Clin North Am 2011 Jun;58(3):637-647 [FREE Full text] [doi: 10.1016/j.pcl.2011.03.003] [Medline: 21600346]
14. Johnson EO, Roth T, Schultz L, Breslau N. Epidemiology of DSM-IV insomnia in adolescence: lifetime prevalence, chronicity, and an emergent gender difference. Pediatrics 2006 Feb 01;117(2):e247-e256. [doi: 10.1542/peds.2004-2629] [Medline: 16452333]
15. Zhang B, Wing YK. Sex differences in insomnia: a meta-analysis. Sleep 2006 Jan;29(1):85-93. [doi: 10.1093/sleep/29.1.85] [Medline: 16453985]
16. Zhang J, Chan NY, Lam SP, Li SX, Liu Y, Chan JWY, et al. Emergence of Sex Differences in Insomnia Symptoms in Adolescents: A Large-Scale School-Based Study. Sleep 2016 Aug 01;39(8):1563-1570 [FREE Full text] [doi: 10.5665/sleep.6022] [Medline: 27091537]
17. Suh S, Cho N, Zhang J. Sex Differences in Insomnia: from Epidemiology and Etiology to Intervention. Curr Psychiatry Rep 2018 Aug 09;20(9):69. [doi: 10.1007/s11920-018-0940-9] [Medline: 30094679]
18. Mong JA, Cusmano DM. Sex differences in sleep: impact of biological sex and sex steroids. Philos Trans R Soc Lond B Biol Sci 2016 Feb 19;371(1688):20150110 [FREE Full text] [doi: 10.1098/rstb.2015.0110] [Medline: 26833831]

19. Luik AI, van der Zweerde T, van Straten A, Lancee J. Digital Delivery of Cognitive Behavioral Therapy for Insomnia. Curr Psychiatry Rep 2019 Jun 04;21(7):50 [FREE Full text] [doi: 10.1007/s11920-019-1041-0] [Medline: 31161406]
20. Trauer JM, Qian MY, Doyle JS, Rajaratnam SM, Cunnington D. Cognitive Behavioral Therapy for Chronic Insomnia: A Systematic Review and Meta-analysis. Ann Intern Med 2015 Aug 04;163(3):191-204. [doi: 10.7326/M14-2841] [Medline: 26054060]
21. van der Zweerde T, Bisdounis L, Kyle SD, Lancee J, van Straten A. Cognitive behavioral therapy for insomnia: A meta-analysis of long-term effects in controlled studies. Sleep Med Rev 2019 Dec;48:101208. [doi: 10.1016/j.smrv.2019.08.002] [Medline: 31491656]
22. Zachariae R, Lyby MS, Ritterband LM, O'Toole MS. Efficacy of internet-delivered cognitive-behavioral therapy for insomnia - A systematic review and meta-analysis of randomized controlled trials. Sleep Med Rev 2016 Dec;30:1-10. [doi: 10.1016/j.smrv.2015.10.004] [Medline: 26615572]
23. de Bruin EJ, Bögels SM, Oort FJ, Meijer AM. Efficacy of Cognitive Behavioral Therapy for Insomnia in Adolescents: A Randomized Controlled Trial with Internet Therapy, Group Therapy and A Waiting List Condition. Sleep 2015 Dec 01;38(12):1913-1926 [FREE Full text] [doi: 10.5665/sleep.5240] [Medline: 26158889]
24. Okajima I, Komada Y, Inoue Y. A meta-analysis on the treatment effectiveness of cognitive behavioral therapy for primary insomnia. Sleep and Biological Rhythms 2011;9(1):24-34. [doi: 10.1111/j.1479-8425.2010.00481.x]
25. Espie CA, Emsley R, Kyle SD, Gordon C, Drake CL, Siriwardena AN, et al. Effect of Digital Cognitive Behavioral Therapy for Insomnia on Health, Psychological Well-being, and Sleep-Related Quality of Life: A Randomized Clinical Trial. JAMA Psychiatry 2019 Jan 01;76(1):21-30 [FREE Full text] [doi: 10.1001/jamapsychiatry.2018.2745] [Medline: 30264137]
26. Cheng, Luik AI, Fellman-Couture C, Peterson E, Joseph CLM, Tallent G, et al. Efficacy of digital CBT for insomnia to reduce depression across demographic groups: a randomized trial. Psychol Med 2019 Feb;49(3):491-500 [FREE Full text] [doi: 10.1017/S0033291718001113] [Medline: 29792241]
27. Lami MJ, Martínez MP, Sánchez AI, Miró E, Diener FN, Prados G, et al. Gender Differences in Patients with Fibromyalgia Undergoing Cognitive-Behavioral Therapy for Insomnia: Preliminary Data. Pain Pract 2016 Feb;16(2):E23-E34. [doi: 10.1111/papr.12411] [Medline: 26841198]
28. Colrain IM, Baker FC. Changes in sleep as a function of adolescent development. Neuropsychol Rev 2011 Mar 12;21(1):5-21 [FREE Full text] [doi: 10.1007/s11065-010-9155-5] [Medline: 21225346]
29. Werner-Seidler A, Wong Q, Johnston L, O'Dea B, Torok M, Christensen H. Pilot evaluation of the Sleep Ninja: a smartphone application for adolescent insomnia symptoms. BMJ Open 2019 May 27;9(5):e026502 [FREE Full text] [doi: 10.1136/bmjopen-2018-026502] [Medline: 31133584]
30. Bastien C, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001 Jul;2(4):297-307. [doi: 10.1016/s1389-9457(00)00065-4] [Medline: 11438246]
31. Alvaro PK, Roberts RM, Harris JK. The independent relationships between insomnia, depression, subtypes of anxiety, and chronotype during adolescence. Sleep Med 2014 Aug;15(8):934-941. [doi: 10.1016/j.sleep.2014.03.019] [Medline: 24958244]
32. Chahoud M, Chahine R, Salameh P, Sauleau EA. Reliability, factor analysis and internal consistency calculation of the Insomnia Severity Index (ISI) in French and in English among Lebanese adolescents. eNeurologicalSci 2017 Jun;7:9-14 [FREE Full text] [doi: 10.1016/j.ensci.2017.03.003] [Medline: 29260018]
33. Chung K, Kan KK, Yeung W. Assessing insomnia in adolescents: comparison of Insomnia Severity Index, Athens Insomnia Scale and Sleep Quality Index. Sleep Med 2011 May;12(5):463-470. [doi: 10.1016/j.sleep.2010.09.019] [Medline: 21493134]
34. de la Vega R, Tomé-Pires C, Solé E, Racine M, Castarlenas E, Jensen MP, et al. The Pittsburgh Sleep Quality Index: Validity and factor structure in young people. Psychol Assess 2015 Dec;27(4):e22-e27. [doi: 10.1037/pas0000128] [Medline: 26653055]
35. Cole J, Motivala SJ, Buysse DJ, Oxman MN, Levin MJ, Irwin MR. Validation of a 3-factor scoring model for the Pittsburgh sleep quality index in older adults. Sleep 2006 Jan;29(1):112-116. [doi: 10.1093/sleep/29.1.112] [Medline: 16453989]
36. Zachariae R, Amidi A, Damholdt MF, Clausen CDR, Dahlgaard J, Lord H, et al. Internet-Delivered Cognitive-Behavioral Therapy for Insomnia in Breast Cancer Survivors: A Randomized Controlled Trial. J Natl Cancer Inst 2018 Aug 01;110(8):880-887 [FREE Full text] [doi: 10.1093/jnci/djx293] [Medline: 29471478]
37. Blake M, Schwartz O, Waloszek JM, Raniti M, Simmons JG, Murray G, et al. The SENSE Study: Treatment Mechanisms of a Cognitive Behavioral and Mindfulness-Based Group Sleep Improvement Intervention for At-Risk Adolescents. Sleep 2017 Jun 01;40(6). [doi: 10.1093/sleep/zsx061] [Medline: 28431122]
38. Carney C, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep 2012 Feb 01;35(2):287-302 [FREE Full text] [doi: 10.5665/sleep.1642] [Medline: 22294820]
39. Armitage R, Trivedi M, Hoffmann R, Rush A. Relationship between objective and subjective sleep measures in depressed patients and healthy controls. Depress Anxiety 1997;5(2):97-102. [doi: 10.1002/(sici)1520-6394(1997)5:2<97::aid-da6>3.0.co;2-2] [Medline: 9262940]
40. Riedel BW, Lichstein KL. Objective sleep measures and subjective sleep satisfaction: how do older adults with insomnia define a good night's sleep? Psychol Aging 1998 Mar;13(1):159-163. [doi: 10.1037//0882-7974.13.1.159] [Medline: 9533198]

41. Sharpley A, Clements A, Hawton K, Sharpe M. Do patients with "pure" chronic fatigue syndrome (neurasthenia) have abnormal sleep? Psychosom Med 1997;59(6):592-596. [doi: 10.1097/00006842-199711000-00006] [Medline: 9407577]
42. Werner-Seidler A, Johnston L, Christensen H. Digitally-delivered cognitive-behavioural therapy for youth insomnia: A systematic review. Internet Interv 2018 Mar;11:71-78 [FREE Full text] [doi: 10.1016/j.invent.2018.01.007] [Medline: 30135762]
43. Werner-Seidler A, O'Dea B, Shand F, Johnston L, Frayne A, Fogarty AS, et al. A Smartphone App for Adolescents With Sleep Disturbance: Development of the Sleep Ninja. JMIR Ment Health 2017 Jul 28;4(3):e28 [FREE Full text] [doi: 10.2196/mental.7614] [Medline: 28754651]
44. Short M, Arora T, Gradisar M, Taheri S, Carskadon MA. How Many Sleep Diary Entries Are Needed to Reliably Estimate Adolescent Sleep? Sleep 2017 Mar 01;40(3) [FREE Full text] [doi: 10.1093/sleep/zsx006] [Medline: 28199718]
45. Kenny DA, Kashy DA, Bolger N. Data analysis in social psychology. In: Gilbert DT, Fiske ST, Lindzey G, editors. The handbook of social psychology. US: McGraw-Hill; 1998:233-265.
46. Fava GA, Tossani E. Prodromal stage of major depression. Early Interv Psychiatry 2007 Feb;1(1):9-18. [doi: 10.1111/j.1751-7893.2007.00005.x] [Medline: 21352104]
47. Jose PE, Brown I. When does the Gender Difference in Rumination Begin? Gender and Age Differences in the Use of Rumination by Adolescents. J Youth Adolescence 2007 Jan 23;37(2):180-192. [doi: 10.1007/s10964-006-9166-y]
48. Nolen-Hoeksema S, Aldao A. Gender and age differences in emotion regulation strategies and their relationship to depressive symptoms. Personality and Individual Differences 2011 Oct;51(6):704-708. [doi: 10.1016/j.paid.2011.06.012]
49. Harvey A. Unwanted intrusive thoughts in insomnia. In: Clark DA, editor. Intrusive thoughts in clinical disorders: Theory, research, and treatment. UK: Guilford Press; 2005:86-118.
50. Pillai V, Drake CL. Chapter 10 - Sleep and Repetitive Thought: The Role of Rumination and Worry in Sleep Disturbance. In: Babson KA, Feldner MT, editors. Sleep and Affect. San Diego: Academic Press; 2015:201-225.
51. Takano K, Iijima Y, Tanno Y. Repetitive Thought and Self-Reported Sleep Disturbance. Behavior Therapy 2012 Dec;43(4):779-789. [doi: 10.1016/j.beth.2012.04.002] [Medline: 23046780]
52. Harvey A. A cognitive model of insomnia. Behav Res Ther 2002 Aug;40(8):869-893. [doi: 10.1016/s0005-7967(01)00061-4] [Medline: 12186352]
53. Gervais NJ, Mong JA, Lacreuse A. Ovarian hormones, sleep and cognition across the adult female lifespan: An integrated perspective. Front Neuroendocrinol 2017 Oct;47:134-153 [FREE Full text] [doi: 10.1016/j.yfrne.2017.08.002] [Medline: 28803147]
54. Baker FC, Driver HS. Circadian rhythms, sleep, and the menstrual cycle. Sleep Med 2007 Sep;8(6):613-622. [doi: 10.1016/j.sleep.2006.09.011] [Medline: 17383933]
55. Cliffe B, Croker A, Denne M, Smith J, Stallard P. Digital Cognitive Behavioral Therapy for Insomnia for Adolescents With Mental Health Problems: Feasibility Open Trial. JMIR Ment Health 2020 Mar 03;7(3):e14842 [FREE Full text] [doi: 10.2196/14842] [Medline: 32134720]

Abbreviations

CBT-I: cognitive behavioral therapy for insomnia
ICC: intraclass correlation
ISI: Insomnia Severity Index
NWAK: number of awakenings
PSQI: Pittsburgh Sleep Quality Index
SE: sleep efficiency
SOL: sleep onset latency
TST: total sleep time
TWT: total wake time
WASO: wake after sleep onset

Edited by G Eysenbach; submitted 14.07.20; peer-reviewed by J Lunsford-Avery, M Kwon; comments to author 10.10.20; revised version received 10.11.20; accepted 19.12.20; published 23.03.21.

Please cite as:
Li SH, Graham BM, Werner-Seidler A
Gender Differences in Adolescent Sleep Disturbance and Treatment Response to Smartphone App–Delivered Cognitive Behavioral Therapy for Insomnia: Exploratory Study
JMIR Form Res 2021;5(3):e22498
URL: https://formative.jmir.org/2021/3/e22498
doi:10.2196/22498
PMID:33755029

Original Paper

Designing an mHealth Intervention for People With Visible Differences Based on Acceptance and Commitment Therapy: Participatory Study Gaining Stakeholders' Input

Fabio Zucchelli[1], BA, MSc; Olivia Donnelly[2], BSc, DClinPsy; Emma Rush[3], MA; Harriet Smith[1], BSc, MSc; Heidi Williamson[1], BSc, MSc, ProfDoc; The VTCT Foundation Research Team[4]

[1]The Centre for Appearance Research, University of the West of England, Bristol, United Kingdom

[2]The Outlook Service, Southmead Hospital, North Bristol NHS Trust, Bristol, United Kingdom

[3]Vitiligo Support UK, London, United Kingdom

[4]see Acknowledgments

Corresponding Author:
Fabio Zucchelli, BA, MSc
The Centre for Appearance Research
University of the West of England
Frenchay Campus
Coldharbour Lane
Bristol, BS16 1QY
United Kingdom
Phone: 44 11732 83882
Email: fabio.zucchelli@uwe.ac.uk

Abstract

Background: Given their growing popularity, mobile health (mHealth) apps may offer a viable method of delivering psychological interventions for people with an atypical appearance (ie, visible difference) who struggle with appearance-related distress. Acceptance and Commitment Therapy (ACT), a third-wave cognitive behavioral approach, has been used effectively in mHealth and is being increasingly applied clinically to common psychosocial difficulties associated with visible differences. We planned to design an ACT-based mHealth intervention (ACT It Out) for this population.

Objective: The aim of this study is to gain key stakeholder input from user representatives and psychological clinicians to optimize the intervention's design for future development and uptake. To do so, we explored considerations relating to mHealth as a delivery platform for adults with visible differences and elicited stakeholders' design preferences and ideas based on initial author-created content.

Methods: Within a participatory design framework, we used a mix of qualitative methods, including usability sessions and a focus group in a face-to-face workshop, and interviews and textual feedback collected remotely, all analyzed using template analysis. A total of 6 user representatives and 8 clinicians were recruited for this study.

Results: Our findings suggest that there are likely to be strengths and challenges of mHealth as an intervention platform for the study population, with key concerns being user safeguarding and program adherence. Participants expressed design preferences toward relatable human content, interactive and actionable features, flexibility of use, accessibility, and engaging content.

Conclusions: The findings offer valuable design directions for ACT It Out and related interventions, emphasizing the need to carefully guide users through the intervention while acknowledging the limited time and space that mHealth affords.

(JMIR Form Res 2021;5(3):e26355) doi:10.2196/26355

KEYWORDS

mobile health; acceptance and commitment therapy; appearance; qualitative; participatory design; mobile phone

Introduction

Background

There are multiple reasons why someone may have an unusual physical appearance, or *visible difference*. Some live with a visible difference from birth, such as people with congenital craniofacial conditions, whereas others acquire a difference as a result of skin disease, injury, and/or medical treatment. In the United Kingdom alone, for example, around 1 in 60 people are estimated to have a visible difference [1]. Many affected individuals thrive; however, in appearance-focused Western cultures in which intrusive scrutiny of those who look different is commonplace, many others experience difficulties, including social anxiety and withdrawal, depression, body dissatisfaction, and low quality of life [2]. Specialist cognitive behavioral interventions show promise in addressing these concerns, typically incorporating social skills training to manage difficult interpersonal interactions [3,4].

In the context of limited specialist face-to-face psychological services for adults with visible differences internationally [5], there is an established need for self-help interventions catering to the specific experiences of this population [6]. Some prefer remote support, for example, because it is less stigmatizing [7]; others have limited or no access to specialist face-to-face services [8]. The review by Muftin and Thompson [8] provides preliminary support for web-based self-help in addressing appearance-based anxiety in adults with visible differences. However, the self-management landscape has since shifted, smartphones having overtaken laptops as many people's primary electronic device [9], and individuals increasingly seek mobile health (mHealth) and mental health apps over web-based formats [10,11]. The aim of our overall project is to design, develop, and evaluate a standalone mHealth intervention, *ACT It Out*, for adults with visible differences experiencing appearance-related distress; to the authors' knowledge, this is the first of its kind. This study describes formal stakeholder involvement at the design stage.

Acceptance and Commitment Therapy

Acceptance and Commitment Therapy (ACT), an established third-wave cognitive behavioral therapy and behavior change model [12], underpins ACT It Out. ACT has been applied to mHealth interventions, with evidence for ACT-based mHealth in enhancing well-being and valued action [13], reducing social anxiety in a clinical population (alongside internet-delivered treatment) [14], and enabling smoking cessation [15]. Psychologists across Europe report using ACT with patients who have visible differences and note its suitability for the population [16,17]. A detailed exposition of how ACT fits the population's needs is given by Zucchelli et al [18].

The process of change targeted in ACT is psychological flexibility, the capacity to direct one's behaviors in accordance with personally held values, thus paying mindful attention both to facilitate valued action and to fully experience its fruition [19,20]. A total of 6 subprocesses mutually develop psychological flexibility: acceptance (willingness to experience private events including unwanted ones); cognitive defusion (loosening thoughts' literality); present-focused attention, self as context (deidentifying from private events); understanding and clarifying one's values (desired qualities of behavior); and committed value-oriented action [19]. Self-compassion is increasingly recognized as an inherent component of psychological flexibility and is specifically nurtured alongside the 6 subprocesses [21]. The converse of 2 of these subprocesses, cognitive fusion (converse of defusion) and experiential avoidance (converse of acceptance), have been shown to partially mediate the relationship between how people with visible differences evaluate their appearance and unhelpful appearance-focused behaviors, including avoidance of appearance-related situations and appearance-fixing behaviors such as covering areas of difference [22].

Psychological flexibility is promoted via mindfulness practices, experiential exercises, and metaphors, such as *passengers on a bus* [23]. This established metaphor describes how private events (thoughts, feelings, etc)—*passengers*—often seem to drive our lives. The metaphor invites us to imagine an alternative where we, the bus driver, take command of our direction of travel by establishing where we want to drive (via values) and by adopting an open, present, and detached relationship with the passengers and their protestations to accommodate their presence along the ride. The behavioral goal of ACT It Out is, therefore, to help users commit to more valued actions and engage in fewer avoidance-oriented behaviors (eg, avoiding social situations that evoke appearance anxiety).

Stakeholder Involvement

ACT It Out is a complex mHealth intervention, namely, one with multiple interacting components targeting new behaviors from its users [24]. In the United Kingdom, where this project is based, the UK Medical Research Council recommends involvement from key stakeholders, including end users, at the design stage of complex interventions [25]. Participatory design methods, which facilitate representative end users' perspectives, preferences, and ideas, are also vital in making any digital intervention appealing and accessible to its target user group [26]. Participatory methods and input from clinicians are also specifically recommended in mental health apps internationally to confer trustworthiness [27]. Accordingly, the authors collaborated in a participatory approach with experts by lived experience (*user representatives*) and clinical experience (*clinicians*) in designing ACT It Out. The aim of this study is to gain user representatives' and clinicians' perspectives using a participatory action procedure [28] to (1) explore the considerations relating to mHealth as a platform for delivering psychological intervention to the target group and (2) understand stakeholders' design preferences and elicit design ideas based on viewing an initial version of ACT It Out created by the authors to help shape the design of ACT It Out. The technical and financial aspects of development are beyond the scope of this study.

Methods

ACT It Out Content

History of the Initial Design

In February 2018, we formed a project team of appearance psychology researchers, a lead clinician with extensive knowledge of ACT and visible difference, a lead user representative also with experience of running a vitiligo support charity (all of whom are coauthors), a human-computer interaction expert, and an app developer.

The researchers and lead clinician first sketched out a preliminary overview of ACT It Out, drawing from knowledge of ACT (including clinical experience in the case of the lead clinician) and self-help development as well as literature on ACT-based mHealth [29] and existing web-based programs for adults with visible differences (eg, Face IT) [30]. With input from the app developer, the project team created a mock-up of a small portion of the intervention. This was presented to members of organizations who represent a range of appearance-affecting conditions at a meeting in February 2018, to gauge interest, elicit ideas, and gain early feedback on design ideas. This study was undertaken before the formal process of stakeholder involvement, which forms the subject of this study.

The Design Presented to Stakeholders

In April 2018, we created a wireframe (screen-by-screen interface illustration) of the first 2 (of 4) sessions of ACT It Out using the software tool Balsamiq, based on the aforementioned background knowledge of authors and relevant visible differences, mHealth and ACT research, as well as feedback of organization members from the meeting in February 2018 (a wireframe screen is shown in Figure 1).

Figure 1. Example wireframe screens produced using Balsamiq.

The wireframe designs subsequently presented to both stakeholder groups comprised 4 sequential training sessions of approximately 40 minutes, each subdivided into 3 subsections (Table 1). We envisaged that users would spend 1 week per session, during which they would engage in brief activities designed to cultivate psychological flexibility, such as guided

mindfulness practices and carrying out self-set valued actions, aided by reminder notifications.

ACT It Out is facilitated by a preprogrammed human *guide* (an expert clinician) who appears in introductory videos in each session and is shown photographically offering textual tips and guidance throughout. The ACT model is introduced with an animated video used in previous ACT self-help interventions [29] showing *passengers on a bus*, which continues as the central metaphor throughout the sessions (eg, with users recognizing their common appearance-focused thoughts as *Your Appearance Passengers* in session 2). The way in which the ACT model relates to the common challenges experienced by people with visible differences is incorporated into the guidance and through real visible difference case examples. Bespoke guided mindfulness practices target acceptance, present attention, cognitive defusion, and self as context (*Attention on your 5 senses* in session 1, *Attention training: Breath and body* in session 2, *Attention training: Managing distress* in session 3, and *Attention training: In daily life* in session 4), alongside specific experiential cognitive defusion and self-compassion exercises. A modified values-sorting exercise [23] helps users clarify their values (*Your values in action* in session 3), and users progressively set short-, medium-, and long-term value-based goals. Social skills training, an evidence-based approach for adults with visible differences [3,4], is also presented as a contextually significant facilitator of behavior change toward valued action (*Building on your social skills* in session 3).

Table 1. High-level layout of the ACT It Out version presented to stakeholders.

Session and subsection	Main ACT[a] processes targeted
Introduction: Getting started	
Why ACT It Out?	N/A[b]
How your data will be used	N/A
Getting to know you	N/A
1. Building awareness	
Passengers on a bus	Open up[c]
Your appearance passengers	Open up
Attention on your 5 senses	Be present[d]
2. Planning your route	
Attention training: Breath and body	Self-compassion
Recap and review of between-session activities	Valued action[e]
Your values in action	Open up, be present
3. Getting social	
Recap and review of between-session activities	Self-compassion
Attention training: Managing distress	Open up, be present
Building on your social skills	Valued action
4. Putting your training into action	
Recap and review of between-session activities	Self-compassion
Attention training: In daily life	Open up, be present
Optional section: Intimacy and visible difference	Valued action
Making a long-term action plan	Valued action

[a]ACT: Acceptance and Commitment Therapy.

[b]N/A: not applicable, as the introductory session does not target ACT processes.

[c]Open up: acceptance and cognitive defusion.

[d]Be present: present attention and self as context.

[e]Valued action: values clarification and committed action [23].

Participants

The participants were 14 stakeholders: 6 user representatives and 8 clinicians. The authors recruited user representatives through purposively selected charities that serve a cross-section of appearance-affecting conditions, including cleft lip and/or palate, alopecia, burns, neurofibromatosis, vitiligo, and Apert syndrome. We purposively recruited for a gender mix (2 males) in user representatives and a wide age range (25-68 years), in an effort not to disadvantage older potential users. Eligibility included age over 18 years, self-identifying as having a visible difference, and some experience of psychosocial challenges

relating to appearance. In addition to the 6 user participants, 1 individual the first author approached was unavailable, and 2 people agreed to take part but were later unable to attend because of personal reasons.

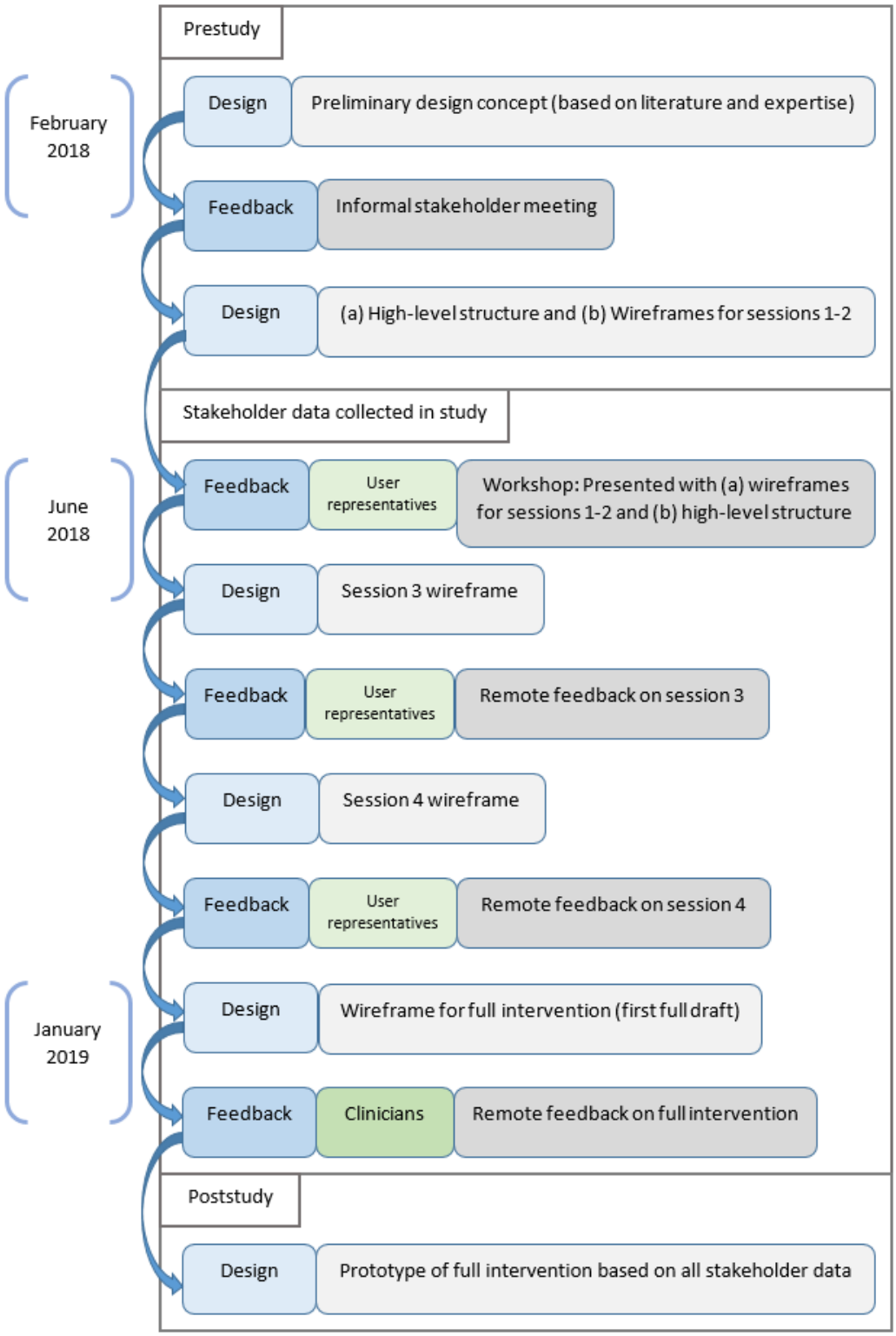
Through the authors' professional networks, we purposively recruited specialist psychological clinicians who have experience of supporting individuals with visible differences and/or applying ACT in a health setting. Most were clinical psychologists (n=6) working for the UK National Health Service or support charities. In total, 4 clinicians were male and 4 were female. Another clinician the authors approached was unavailable to participate. All participants were White and were based in the United Kingdom.

Design

Overview

We employed participatory design methods to create a design via a series of iterations based on stakeholder contributions, as recommended for complex interventions and mHealth [25,27]. The authors front-loaded user representative input to ensure that the user perspective was incorporated into the first full draft of the ACT It Out design, before seeking clinicians' input. The full iterative design process within which this study is conducted is shown in Figure 2. Ethical approval for this research was obtained from the University Research Ethics Committee.

Figure 2. Diagram of the iterative design process used in ACT It Out. ACT: Acceptance and Commitment Therapy.



User Representatives

We arranged a user representative workshop in June 2018. The authors chose a group workshop so that participants could meet each other, conducive to a sense of commonality, and to mitigate any potential power imbalance between participants and researchers. A total of 7 people accepted invitation and 4 attended (all 3 nonattendees notified the researchers).

The first author began the workshop by welcoming user representatives and giving an overview of the project, and then, the second author facilitated an *icebreaker* exercise. This was followed by one-to-one usability sessions, facilitated by trained researchers, of the ACT It Out introduction and sessions 1 and 2, in which user representatives viewed paper wireframes (following feedback from the February 2018 meeting that a nonclickable smartphone mock-up can cause frustration). Participants were invited to complete study-specific usability

feedback forms and describe aloud their experience of using it as they progressed (*think-aloud* protocol) [31]. After a break, the participants took part in a semistructured focus group facilitated by the first and fifth authors. Topics included (1) advantages and disadvantages of mHealth as a self-help platform for individuals struggling with visible differences, (2) preferences for using mHealth (eg, duration and setting), (3) feedback on the high-level ACT It Out structure, (4) preferences regarding screen interface, and (5) feedback on specific design elements of ACT It Out. The focus group lasted for 1 hour and 39 minutes.

Following the workshop, one nonattendee viewed the introduction and sessions 1 and 2 wireframes remotely and completed the usability feedback form electronically. From the combined feedback, the first author produced a wireframe of session 3, then emailed this as a PDF file to user representatives to view independently and complete a feedback form. The first author then repeated this process in designing and gaining feedback for session 4, which 4 participants completed. In both stages, 4 participants took part. One additional workshop nonattendee viewed the entire wireframe remotely and provided written feedback. In total, 6 user representatives contributed to the design phase. The authors then discussed key design changes, and the first author incorporated user representatives' feedback into the full intervention wireframes.

Clinicians

We then obtained feedback from expert clinicians from January 2019 onward. Due to busy work commitments of clinicians, the first author sent clinicians the wireframes as PDF files to view in their own time and asked them to complete a feedback form. The first author also arranged a telephone interview within a week of viewing the wireframes. Interviews were semistructured, following a largely equivalent schedule to the users' focus group, along with more clinically oriented topics, including (1) defining the user group for whom ACT It Out is suitable (and unsuitable) and (2) design elements for behavior change. Interviews lasted for 43 minutes to 67 minutes (mean 51.63, SD 7.32).

Data Analysis

The first and fourth authors pseudonymized and transcribed verbatim data from the user representative focus group, usability sessions, and clinician interviews. Together with written feedback, all data were transferred into NVivo version 12 for analysis. Data were analyzed using template analysis, an iterative form of thematic analysis in which a coding template is developed and modified throughout the analysis [32]. The authors chose template analysis for its compatibility with the project's iterative, action-research paradigm and its accommodation of combined deductive and inductive analytical approaches. We deductively applied a priori themes from relevant mHealth literature (Textbox 1) to the first template, which were either removed or modified in light of new data. The authors also applied an inductive approach to capture any novel insights. The first author applied the first template to further data and incrementally refined the template where new data did not fit existing themes, culminating after 8 versions in a final template that encompassed all relevant data, on which the below results are based. The fifth author reviewed the analysis by checking transcripts and templates, resulting in the subsequent addition of the final (eighth) theme presented below.

Textbox 1. A priori themes from a literature review.

A priori themes kept in the final coding template (although the wording was usually modified to better reflect participants' accounts).

Considerations of mobile health as a platform

- **Themes retained**
  - Mobile health can augment therapy [34]
  - Ease of access and portability [28]
  - High dropout in mobile health [47]
  - Fear of iatrogenic affects [10]
- **Themes not retained**
  - Users' data privacy [10,34,35,47]
  - Trustworthiness or credibility [11,27,35,47]

Design preferences

- **Themes retained**
  - Real-time engagement [33]
  - Reminders and notifications to engage [28,33,34]
  - Links to crisis support [27,33,47]
  - Immediate feedback [28,34]
  - Adaptable functionalities (eg, font and layout) [27]
- **Themes not retained**
  - Skepticism toward gamification [28]
  - Concise content [28]
  - Clear instructions for use [27]

Results

Overview

We produced 7 first-order themes, each split into lower-level themes, and 1 integrative theme (full details of the themes, their relation to the research questions, and which participant groups contributed toward each theme are given in Multimedia Appendix 1). The central findings from the combined participant groups, with illustrative quotes, are provided in the following sections. To preserve anonymity, participants' comments were labeled by the stakeholder group only.

1. mHealth Has Advantages for Users

Participants expressed the advantages of mHealth interventions over traditional talking therapies and other self-management platforms. Participants valued the privacy afforded by mHealth in terms of its discretion and suitability for those preferring not to seek face-to-face support for appearance concerns:

> *[An app] is perfect for someone who doesn't want to go and get counselling, this is a self-help thing that we all do and it's very modern so it's something that is needed...plus reading it on an app, it's not like you're reading a self-help book and everybody else can see what you're reading; it's personal.* [User representative, focus group]

Participants' accounts also highlighted the autonomy offered by mHealth through its portability and accessibility:

> *I like apps as an idea because you can use them wherever you are. Suppose you've got your meditation tabs [on the app]...Going to the pub and having a difficult moment, it's good to be able to access it there and then.* [Clinician, interview]

2. mHealth Should Add to—Not Replace—Existing Face-to-Face Resources

Participants felt that mHealth could complement existing face-to-face psychological support before, during, or after face-to-face support:

> *[mHealth] might be a step toward thinking about what help [users] might need. So it could be a platform for other things...I got excited [looking at the wireframe]; I could use this alongside working with somebody. You could alternate them doing something and then having a conversation about it.* [Clinician, interview]

Clinicians, in particular, voiced their view that mHealth cannot and should not replace face-to-face support:

> *As an adjunct, or something that's available where nothing else is available I think [mHealth] has potentially a huge benefit—but I'd not want to see it as a replacement for all individual therapy.* [Clinician, interview]

Relatedly, participants suggested that mHealth may be less suitable for some, including individuals with high levels of need or recently acquired visible differences:

> *[An app is] possibly [not suitable for] people who have only acquired a visible difference recently, because they may not know they have trauma responses, or they may need more time before they jump into something like this.* [Clinician, interview]

3. Safeguard Users' Well-Being

Participants emphasized the need to embed content and features that safeguard users' well-being:

> *...have a section [in the app] with useful links so you could fast-track to all of the different links and organisations where you can go [for support].* [User representative, focus group]

> *Is there a fail-safe mechanism, if someone has checked too hard all the way through [in response to questions asking if users would like to work on specific social skills]—have a screen that pops up and checks in with them.* [User representative, written feedback]

> *Possible discomfort is part of the process [of ACT], and [users] need to know that before [they] commit to it. And you might decide that if you've got particular things going on in your life, "You know what, not right now," or "Now is exactly the time I want to do this." But doing it knowing this isn't going to be a series of aromatherapy massages where at the end of each one you'll expect to feel much nicer than when you started.* [Clinician, interview]

4. There Has to be That Human Link

Participants described the importance of ACT It Out establishing a human connection with users. Many preferred the idea of a single preprogrammed human *guide* with relevant experience over multiple guides:

> *I think there's a lot to be said about there being a teacher. Certainly within the world of mindfulness, there's a healthy attachment to a teacher figure...You want a face to go with the voice. There has to be that human link to get people going.* [Clinician, interview]

User representatives, in particular, felt that the in-built human guide needs to be responsive to user input, to validate the common experiences of people with visible differences and offer encouragement:

> *Just have options with a dropdown menu, [eg,] If [a task] made me feel [unconfident], and then you give a bit of advice like "It's ok to feel like this, it's ok that you feel a little bit less confident right now."* [User representative, focus group]

> *...if there was some feedback loop in there, [users] might be more inclined to put something in.* [User representative, focus group]

Participants from both groups highlighted the need for ACT It Out to offer a way of normalizing users' experiences and saw real-life case examples as the best way of doing so, while also helping users to buy into the intervention's processes:

> *...the [real] examples are very good, [to see] that's how other people cope and see real experiences not just what experts think. It has to come from an actual person.* [User representative, usability session]

5. Engender Action

Participants were clear that a core design priority of the app should be to elicit value-based behavior change. To do so, participants suggested that content should focus on actionable tasks:

> *I felt like one really strong bit is the social skills bit. Because it was really practical, gives some clear guidance.* [Clinician, written feedback]

User-set notification reminders for activities were seen as a crucial tool unique to mHealth to aid engagement with behavioral tasks:

User representative (usability session): Would you be able to choose when you get your notifications?

Researcher: Yeah, does that feel important to you?

User representative: Yeah

> *I love the capacity to have reminders and the user to have control around the timings etc.* [Clinician, written feedback]

Clinicians described the importance of ACT It Out facilitating both immediate and sustainable behavior change, for example, by providing sufficient time between sessions for practice and by breaking behavioral goals into discrete, manageable segments:

> *Selecting your values and then implementing them as short, medium and long-term goals. I could see the benefit of that.* [Clinician, interview]

6. Design for a Range of Users

Participants' divergent app use preferences highlighted the need to build in flexibility about how the program is used:

> *It's helpful to have a journal and plus at the end of it you can always go back to see how far you've come. I don't want to sit there and write down how I'm feeling.* [User representative, focus group]

> *My train journey from where I am to go anywhere is 45 minutes and [I] sit on a different app for the full 45 minutes, so that doesn't seem very long to me [to use an app].* [User representative, focus group]

> *I just think it's kind of human nature, we don't want to sit down and spend a lot of time on [an app].* [User representative, focus group]

With ACT It Out designed for people with a range of causes of visible difference, participants noted the physical usability issues that may exist for some and the need to design the screen interface accordingly:

> *As well as facial differences, a lot of people with burns or other conditions might not have [fully functional] fingers so it's being able to press something...I found it yesterday logging onto [a public transport app], a tiny tick-box for terms and conditions, I couldn't press it because the button was far too small for my finger...it would have to be a nice big button.* [User representative, usability session]

7. Design for Learning

Participants' feedback also suggested that the content and structure of ACT It Out should actively facilitate user learning. Participants sought a clear rationale for each element of training:

> *Can you say something more here [in the first mindfulness training section] about mindfulness? Why is mindfulness helpful? Why are we asking people to do it?* [Clinician, written feedback]

Participants expressed a desire for training content to be organized into short, sequential chunks that consolidate and build on preceding sessions:

> *I like that it's broken down into different sections...and also within each session I think is helpful. Because psychological information can be a bit jargon-heavy, a bit much.* [Clinician, interview]

> *I can imagine that people going through it maybe have a tendency to go "I know what that's all about" and move on, but it looks like you've built it in so you'd have to actively click Next. I think the way it's laid out is quite easy to follow and gradually builds up.* [User representative, focus group]

Clinicians felt that learning would be aided by linking all training material to a small number of simple, actionable models:

> *I think having the Choice Point [a visual ACT model] in each session would be a really good way of tying the content together. Something quite visual...if [a user] is new to [ACT], just having one or two things that we just keep coming back to again and again.* [Clinician, interview]

8. Mitigate Dropout

An integrative theme that permeated across themes 3 to 7 conveyed participants' concern over users discontinuing the use of ACT It Out:

> *I suppose the fear would be someone starts and then doesn't carry on...before they get to the good bit.* [User representative, focus group]

> *...if somebody comes into [ACT It Out] with high expectations but significant problems, are you actually going to add to the problem because they're going to fail through the app? And how to manage that process?* [Clinician, interview]

Discussion

Principal Findings

In this study, we aim to identify the considerations of mHealth as a therapeutic platform for adults with visible differences and the design preferences and ideas of key stakeholders to optimize the design of ACT It Out. The 8 themes subsequently informed the full redesign of ACT It Out to be piloted in the prototype form.

Participants' accounts highlighted strengths, challenges, and limitations of mHealth as a mode of delivering psychological interventions for adults with visible differences. The strengths expressed by participants, namely, user discretion, accessibility, and portability, have been reported in the mHealth literature [33,34]. These features are likely to be of particular use to those experiencing social anxiety, a common challenge for people with visible differences [2,36-38]. Using a mobile platform and aided by social skills training, value clarification, and behavioral action plans, ACT It Out may offer a suitable and accessible medium for users to potentially transition into greater social activity.

In terms of challenges, participants clearly expressed the need to safeguard users' well-being and mitigate potentially deleterious effects. The potential for iatrogenic effects has previously been highlighted as a concern for mental health apps [10]. Within the training content, clinicians, in particular, highlighted the need to provide clear, regular, and timely information on the ACT processes underpinning ACT It Out to manage users' expectations and offer informed consent on its ongoing use. Although the effectiveness of trigger warnings before potentially distressing material and exercises remains unsubstantiated [39], providing users with relevant information about all exercises in advance and treating informed consent as an ongoing process is in keeping with ethical guidelines for psychological intervention more broadly [40].

Participants' feedback suggests that the absence of live interaction with a real therapist may pose an advantage for those who prefer remote support and a limitation for others. Establishing a *human* element to ACT It Out was, though, favored unanimously, corresponding with a growing interest in facilitating a therapeutic alliance in mHealth [41] and user preferences for embodied, empathic chatbot hosts rather than anonymous avatars [42]. Using a human *guide* via videos, photographs, and text, who responds to user input, was deemed crucial to providing meaningful interaction and validation of users' experiences. Nevertheless, clinicians' feedback highlighted that mHealth cannot fully replicate face-to-face support, and hence, ACT It Out would be unable to offer adequate support for those with greater clinical needs. Therefore, we need to provide clear guidance on who it is designed for and for whom professional support may be more suitable, both in promoting ACT It Out and in its content.

Various design preferences expressed by participants have since been incorporated into ACT It Out and may be informative for those developing any related interventions. Real case examples were resoundingly popular, offering a way of normalizing the

typical difficulties experienced by people with visible differences, counteracting a felt sense of difference common to this group [2]. Such case examples may also confer credibility to the ACT approach. In keeping with the participatory action paradigm, some of the user representative participants offered their own stories for future iterations of ACT It Out.

Participants' accounts pointed to the need to design content and features to encourage concrete behaviors. User-set notifications were valued by all participants as reminders for activities, echoing previous research on mental health apps [43]. Designing precise and time-limited value-based goals, with reminder notifications that appear on the same device through which many of the training activities are undertaken, offers a powerful tool for enacting implementation intentions (a strategy of specifying when and how an individual engages in goal-directed behavior) of the type targeted in ACT It Out [44]. As raised by clinicians, sufficient time is needed to establish and sustain behavioral changes, such as regular mindfulness and social skills training practice. To offer greater opportunity for behavior change and to reflect the mean duration spent by users on cognitive behavioral apps (5.4 weeks) [10], the authors have since increased the ACT It Out training content from 4 to 6 weekly sessions, without increasing the overall content.

Participants' preference for building the training content in a systematic, step-by-step fashion with manageable amounts of information corresponds to the user experience principle of progressive disclosure, in which only information essential to any given step of a process is provided when users need it [45]. This is especially important in mHealth design, where there is less space and time to engage users' attention [46]. Therefore, we simplified and divided the training content in accordance with this principle.

Participants' feedback also highlighted that not all user design facets are universal, with users likely to vary in their interaction preferences. Therefore, we included an optional journal or notes feature throughout the training for those who wish to use it. The need to accommodate users' varying physical needs is also paramount, especially for a user group in which many people have conditions that also affect physical function (eg, impaired digit functioning, hearing loss) and appearance. Features such as large font and buttons and optional subtitles and scripts for recordings offer the sort of adaptive functionality recommended in mHealth design [27].

The relatively high mHealth user dropout rate established in the literature [35,47] was echoed by participants' concerns. Many of their design preferences sought to foster engagement and counter potential causes of attrition, such as overly challenging materials.

Limitations

A limitation of the study was the small sample size, especially of the user representative group. The project was conducted primarily as stakeholder collaboration rather than in a traditional researcher-as-expert paradigm, as befits the development of a complex intervention [25]. This meant that the authors prioritized the quality of relationships with user representatives over their quantity. Collaborating with a small number of engaged user representatives over a year meant that the group was well informed about the project's objectives and their role within it. A second limitation was that the first author and lead designer of ACT It Out was heavily involved in data collection, creating the possibility of acquiescence bias in participants' responses [48]. The first authors' unique knowledge of ACT It Out meant that he was nevertheless best placed to gain feedback from participants and could follow up on participants' responses during data collection.

Conclusions

By collaborating with key stakeholders, namely, user representatives with visible differences and clinicians, this study established several actionable directions for the mHealth intervention (ACT It Out) under development. Gaining both user and clinician perspectives gave us a comprehensive picture of what an mHealth intervention based on ACT should look like for the target population. This paper also offers an example for other researchers involved in developing mHealth and other complex behavioral interventions and allows the authors' methods to be critiqued [25].

Acknowledgments

The authors are grateful to all the participants and charitable organizations who helped to facilitate this study. They are also grateful for the support of the wider Vocational Training Charitable Trust (VTCT) Foundation Research Team at the Centre for Appearance Research, who are Dr Amy Slater, Dr Nicola Marie Stock, Dr Claire Hamlet, Mr Nicholas D Sharratt, Ms Ella Guest, Ms Bruna Costa, Ms Jade Parnell, Ms Maia Thornton, Professor Nichola Rumsey, and Professor Diana Harcourt. This research was supported through a donation made to the Centre for Appearance Research, University of the West of England, by the Vocational Training Charitable Trust Foundation (a registered charity in England and Wales, no.1155360).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Themes mapped to research questions and their respective participant groups.
[DOCX File , 23 KB - formative_v5i3e26355_app1.docx]

References

1. Disfigurement in the UK. Changing Faces. URL: https://www.changingfaces.org.uk/campaign/dituk [accessed 2020-11-27]
2. Rumsey N, Harcourt D. Body image and disfigurement: issues and interventions. Body Image 2004 Jan;1(1):83-97. [doi: 10.1016/S1740-1445(03)00005-6] [Medline: 18089143]
3. Bessell A, Moss TP. Evaluating the effectiveness of psychosocial interventions for individuals with visible differences: a systematic review of the empirical literature. Body Image 2007 Sep;4(3):227-238. [doi: 10.1016/j.bodyim.2007.04.005] [Medline: 18089269]
4. Norman A, Moss TP. Psychosocial interventions for adults with visible differences: a systematic review. PeerJ 2015;3 [FREE Full text] [doi: 10.7717/peerj.870] [Medline: 25861556]
5. Williamson H, Antonelli P, Bringsén Å, Davies G, Dèttore D, Harcourt D, et al. Supporting patients with body dissatisfaction: a survey of the experiences and training needs of european multi-disciplinary healthcare professionals. NursingPlus Open 2018;4:8-13. [doi: 10.1016/j.npls.2018.05.001]
6. Bessell A, Clarke A, Harcourt D, Moss TP, Rumsey N. Incorporating user perspectives in the design of an online intervention tool for people with visible differences: face IT. Behav Cogn Psychother 2010 Oct;38(5):577-596. [doi: 10.1017/S1352465810000305] [Medline: 20609269]
7. Berger M, Wagner TH, Baker LC. Internet use and stigmatized illness. Soc Sci Med 2005 Oct;61(8):1821-1827. [doi: 10.1016/j.socscimed.2005.03.025] [Medline: 16029778]
8. Muftin Z, Thompson AR. A systematic review of self-help for disfigurement: effectiveness, usability, and acceptability. Body Image 2013 Sep;10(4):442-450. [doi: 10.1016/j.bodyim.2013.07.005] [Medline: 23962642]
9. Communications market reports. Ofcom. URL: https://www.ofcom.org.uk/research-and-data/multi-sector-research/cmr [accessed 2020-11-27]
10. Torous J, Levin ME, Ahern DK, Oser ML. Cognitive behavioral mobile applications: clinical studies, marketplace overview, and research agenda. Cognitive Beh Pract 2017 May;24(2):215-225. [doi: 10.1016/j.cbpra.2016.05.007]
11. Sucala M, Cuijpers P, Muench F, Cardo R, Soflau R, Dobrean A, et al. Anxiety: there is an app for that. A systematic review of anxiety apps. Depress Anxiety 2017 Jun;34(6):518-525. [doi: 10.1002/da.22654] [Medline: 28504859]
12. Hayes SC, Pistorello J, Levin ME. Acceptance and commitment therapy as a unified model of behavior change. Counsel Psychol 2012 Sep 26;40(7):976-1002. [doi: 10.1177/0011000012460836]
13. Krafft J, Potts S, Schoendorff B, Levin ME. A randomized controlled trial of multiple versions of an acceptance and commitment therapy matrix app for well-being. Behav Modif 2019 Mar;43(2):246-272. [doi: 10.1177/0145445517748561] [Medline: 29262693]
14. Ivanova E, Lindner P, Ly KH, Dahlin M, Vernmark K, Andersson G, et al. Guided and unguided acceptance and commitment therapy for social anxiety disorder and/or panic disorder provided via the internet and a smartphone application: a randomized controlled trial. J Anxiety Disord 2016 Dec;44:27-35. [doi: 10.1016/j.janxdis.2016.09.012] [Medline: 27721123]
15. Bricker JB, Mull KE, Kientz JA, Vilardaga R, Mercer LD, Akioka KJ, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. Drug Alcohol Depend 2014 Oct 01;143:87-94 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.07.006] [Medline: 25085225]
16. Harcourt D, Hamlet C, Feragen KB, Garcia-Lopez L, Masnari O, Mendes J, et al. The provision of specialist psychosocial support for people with visible differences: a European survey. Body Image 2018 Jun;25:35-39 [FREE Full text] [doi: 10.1016/j.bodyim.2018.02.001] [Medline: 29454284]
17. Stock NM, Zucchelli F, Hudson N, Kiff JD, Hammond V. Promoting psychosocial adjustment in individuals born with cleft lip and/or palate and their families: current clinical practice in the United Kingdom. Cleft Palate Craniofac J 2020 Feb;57(2):186-197. [doi: 10.1177/1055665619868331] [Medline: 31431061]
18. Zucchelli F, Donnelly O, Williamson H, Hooper N. Acceptance and commitment therapy for people experiencing appearance-related distress associated with a visible difference: a rationale and review of relevant research. J Cogn Psychother 2018 Aug;32(3):171-183. [doi: 10.1891/0889-8391.32.3.171] [Medline: 32746433]
19. Hayes SC, Strosahl KD, Wilson KG. Acceptance and commitment therapy: an experiential approach to behavior change. New York, United States: Guilford Press; 1999. URL: https://contextualscience.org/publications/hayes_strosahl_wilson_1999 [accessed 2021-03-01]
20. Stockton D, Kellett S, Berrios R, Sirois F, Wilkinson N, Miles G. Identifying the underlying mechanisms of change during Acceptance and Commitment Therapy (ACT): a systematic review of contemporary mediation studies. Behav Cogn Psychother 2019 May;47(3):332-362. [doi: 10.1017/S1352465818000553] [Medline: 30284528]
21. Luoma JB, Platt MG. Shame, self-criticism, self-stigma, and compassion in Acceptance and Commitment Therapy. Current Opinion Psychol 2015 Apr;2:97-101 [FREE Full text] [doi: 10.1016/j.copsyc.2014.12.016]
22. Zucchelli F, White P, Williamson H, VTCT Foundation Research Team at the Centre for Appearance Research. Experiential avoidance and cognitive fusion mediate the relationship between body evaluation and unhelpful body image coping strategies in individuals with visible differences. Body Image 2020 Mar;32:121-127. [doi: 10.1016/j.bodyim.2019.12.002] [Medline: 31869761]
23. Harris R. Act made simple: an easy-to-read primer on acceptance and commitment. Oakland, California, United States: New Harbinger; 2009.

24. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. Int J Nurs Stud 2013 May;50(5):587-592. [doi: 10.1016/j.ijnurstu.2012.09.010] [Medline: 23159157]
25. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open 2019 Aug 15;9(8) [FREE Full text] [doi: 10.1136/bmjopen-2019-029954] [Medline: 31420394]
26. Stratton E, Choi I, Peters D, Calvo RA, Harvey SB, Glozier N. Co-designing a web-based decision aid tool for employees disclosure of mental health conditions: a participatory study design using employee and organizational preferences. JMIR Form Res 2020 Nov 06;4(11) [FREE Full text] [doi: 10.2196/23337] [Medline: 33155982]
27. Alqahtani F, Orji R. Insights from user reviews to improve mental health apps. Health Informatics J 2020 Sep 10;26(3):2042-2066 [FREE Full text] [doi: 10.1177/1460458219896492] [Medline: 31920160]
28. Bradbury H, Reason P. Action Research: An Opportunity for Revitalizing Research Purpose and Practices. Qualitative Social Work 2003;2(2):155-175 [FREE Full text] [doi: 10.1177/1473325003002002003]
29. Barker E. The utility of ACT based apps in healthcare. University of Edinburgh. URL: https://era.ed.ac.uk/bitstream/handle/1842/23422/Barker2016.pdf?isAllowed=y&sequence=2 [accessed 2021-02-22]
30. Bessell A, Brough V, Clarke A, Harcourt D, Moss TP, Rumsey N. Evaluation of the effectiveness of Face IT, a computer-based psychosocial intervention for disfigurement-related distress. Psychol Health Med 2012;17(5):565-577. [doi: 10.1080/13548506.2011.647701] [Medline: 22313169]
31. Fonteyn ME, Kuipers B, Grobe SJ. A description of think aloud method and protocol analysis. Qual Health Res 2016 Jul;3(4):430-441. [doi: 10.1177/104973239300300403]
32. Brooks J, McCluskey S, Turley E, King N. The utility of template analysis in qualitative psychology research. Qual Res Psychol 2015 Apr 03;12(2):202-222 [FREE Full text] [doi: 10.1080/14780887.2014.955224] [Medline: 27499705]
33. Stawarz K, Preist C, Tallon D, Wiles N, Coyle D. User experience of cognitive behavioral therapy apps for depression: an analysis of app functionality and user reviewsuser experience of cognitive behavioral therapy apps for depression: an analysis of app functionality and user reviews. J Med Internet Res 2018 Jun 06;20(6) [FREE Full text] [doi: 10.2196/10120] [Medline: 29875087]
34. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 01;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
35. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
36. Alyami M, Giri B, Alyami H, Sundram F. Social anxiety apps: a systematic review and assessment of app descriptors across mobile store platforms. Evid Based Ment Health 2017 Aug;20(3):65-70. [doi: 10.1136/eb-2017-102664] [Medline: 28666986]
37. Attridge MD, Morfitt RC, Roseborough DJ, Jones ER. Internet-based cognitive-behavioral therapy for college students with anxiety, depression, social anxiety, or insomnia: four single-group longitudinal studies of archival commercial data and replication of employee user study. JMIR Form Res 2020 Jul 23;4(7) [FREE Full text] [doi: 10.2196/17712] [Medline: 32706662]
38. Stott R, Wild J, Grey N, Liness S, Warnock-Parkes E, Commins S, et al. Internet-delivered cognitive therapy for social anxiety disorder: a development pilot series. Behav Cogn Psychother 2013 Jul;41(4):383-397 [FREE Full text] [doi: 10.1017/S1352465813000404] [Medline: 23676553]
39. Sanson M, Strange D, Garry M. Trigger warnings are trivially helpful at reducing negative affect, intrusive thoughts, and avoidance. Clin Psychol Sci 2019 Mar 04;7(4):778-793. [doi: 10.1177/2167702619827018]
40. Code of ethics and conduct. The British Psychological Society. URL: https://www.bps.org.uk/news-and-policy/bps-code-ethics-and-conduct [accessed 2021-11-27]
41. 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]
42. Lucas GM, Gratch J, King A, Morency L. It's only a computer: virtual humans increase willingness to disclose. Comp in Hum Beh 2014 Aug;37:94-100. [doi: 10.1016/j.chb.2014.04.043]
43. Goodwin J, Cummins J, Behan L, O'Brien SM. Development of a mental health smartphone app: perspectives of mental health service users. J Ment Health 2016 Oct;25(5):434-440. [doi: 10.3109/09638237.2015.1124392] [Medline: 26732242]
44. Pirolli P, Mohan S, Venkatakrishnan A, Nelson L, Silva M, Springer A. Implementation intention and reminder effects on behavior change in a mobile health system: a predictive cognitive model. J Med Internet Res 2017 Nov 30;19(11):e397 [FREE Full text] [doi: 10.2196/jmir.8217] [Medline: 29191800]
45. Lowdermilk T. User–centered design: a developer's guide to building user-friendly. Sebatopol, CA: O'Reilly Media, Inc; 2013.
46. Roth WR, Vilardaga R, Wolfe N, Bricker JB, McDonell MG. Practical considerations in the design and development of smartphone apps for behavior change. J Contextual Behav Sci 2014 Oct 01;3(4):269-272 [FREE Full text] [doi: 10.1016/j.jcbs.2014.08.006] [Medline: 25485231]

47. Brown M, Hooper N, James P, Scott D, Bodger O, John A. A web-delivered acceptance and commitment therapy intervention with email reminders to enhance subjective well-being and encourage engagement with lifestyle behavior change in health care staff: randomized cluster feasibility study. JMIR Form Res 2020 Aug 07;4(8) [FREE Full text] [doi: 10.2196/18586] [Medline: 32763887]
48. Ray JJ. Reviving the problem of acquiescent response bias. J Soc Psychol 2010 Jul;121(1):81-96. [doi: 10.1080/00224545.1983.9924470]

Abbreviations

ACT: Acceptance and Commitment Therapy
mHealth: mobile health

Edited by G Eysenbach; submitted 08.12.20; peer-reviewed by A Norman, A van der Horst; comments to author 19.01.21; revised version received 02.02.21; accepted 04.02.21; published 24.03.21.

Please cite as:
Zucchelli F, Donnelly O, Rush E, Smith H, Williamson H, The VTCT Foundation Research Team
Designing an mHealth Intervention for People With Visible Differences Based on Acceptance and Commitment Therapy: Participatory Study Gaining Stakeholders' Input
JMIR Form Res 2021;5(3):e26355
URL: https://formative.jmir.org/2021/3/e26355
doi:10.2196/26355
PMID:33759791

Viewpoint

Development of Coaching Support for LiveWell: A Smartphone-Based Self-Management Intervention for Bipolar Disorder

Cynthia A Dopke[1*], PhD; Alyssa McBride[1*], BA; Pamela Babington[1*], BA; Geneva K Jonathan[1*], MS; Tania Michaels[2*], MD; Chloe Ryan[3*], LCSWA; Jennifer Duffecy[4*], PhD; David C Mohr[1*], PhD; Evan H Goulding[1*], MD, PhD

[1]Department of Psychiatry and Behavioral Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[2]General Pediatrics, Loma Linda Children's Hospital, Loma Linda, CA, United States

[3]Department of Social Work, UPMC Western Psychiatric Hospital, Pittsburgh, PA, United States

[4]Department of Psychiatry, University of Illinois at Chicago, Chicago, IL, United States

[*]all authors contributed equally

Corresponding Author:
Evan H Goulding, MD, PhD
Department of Psychiatry and Behavioral Sciences
Feinberg School of Medicine
Northwestern University
303 E Chicago Ave.
Suite 7-102
Chicago, IL, 60611
United States
Phone: 1 3125031189
Email: e-goulding@fsm.northwestern.edu

Abstract

Despite effective pharmacological treatment, bipolar disorder is a leading cause of disability due to recurrence of episodes, long episode durations, and persistence of interepisode symptoms. While adding psychotherapy to pharmacotherapy improves outcomes, the availability of adjunctive psychotherapy is limited. To extend the accessibility and functionality of psychotherapy for bipolar disorder, we developed LiveWell, a smartphone-based self-management intervention. Unfortunately, many mental health technology interventions suffer from high attrition rates, with users rapidly failing to maintain engagement with the intervention technology. Human support reduces this commonly observed engagement problem but does not consistently improve clinical and recovery outcomes. To facilitate ongoing efforts to develop human support for digital mental health technologies, this paper describes the design decisions, theoretical framework, content, mode, timing of delivery, and the training and supervision for coaching support of the LiveWell technology. This support includes clearly defined and structured roles that aim to encourage the use of the technology, self-management strategies, and communication with care providers. A clear division of labor is established between the coaching support roles and the intervention technology to allow lay personnel to serve as coaches and thereby maximize accessibility to the LiveWell intervention.

(JMIR Form Res 2021;5(3):e25810) doi:10.2196/25810

KEYWORDS

human support; adherence; self-management; behavior change; mHealth; bipolar disorder

Background

Bipolar disorder is a serious mental illness characterized by recurrent episodes of mania, hypomania, depression, and mixed states [1,2]. Even with pharmacological treatment, recurrence of acute episodes, long episode durations, and persistence of interepisode symptoms leads to significant disability, with three-quarters of those affected never achieving full recovery of psychosocial function [3-8]. The addition of empirically supported psychotherapy to pharmacotherapy improves clinical and recovery outcomes [4,9], but only half of individuals with bipolar disorder receive any therapy [10,11].

Mental health technologies (MHTs), including web and smartphone-based applications, provide a means to increase the

availability of empirically supported psychotherapeutic strategies for managing mental health problems. Additionally, relative to face-to-face (F2F) treatment, MHTs can increase the functionality of interventions by providing real-time assessments, feedback, and provider alerts. MHTs have been developed and found to be effective for depression and anxiety, and recent efforts have begun to focus on serious mental illnesses such as bipolar disorder and psychosis [12-17].

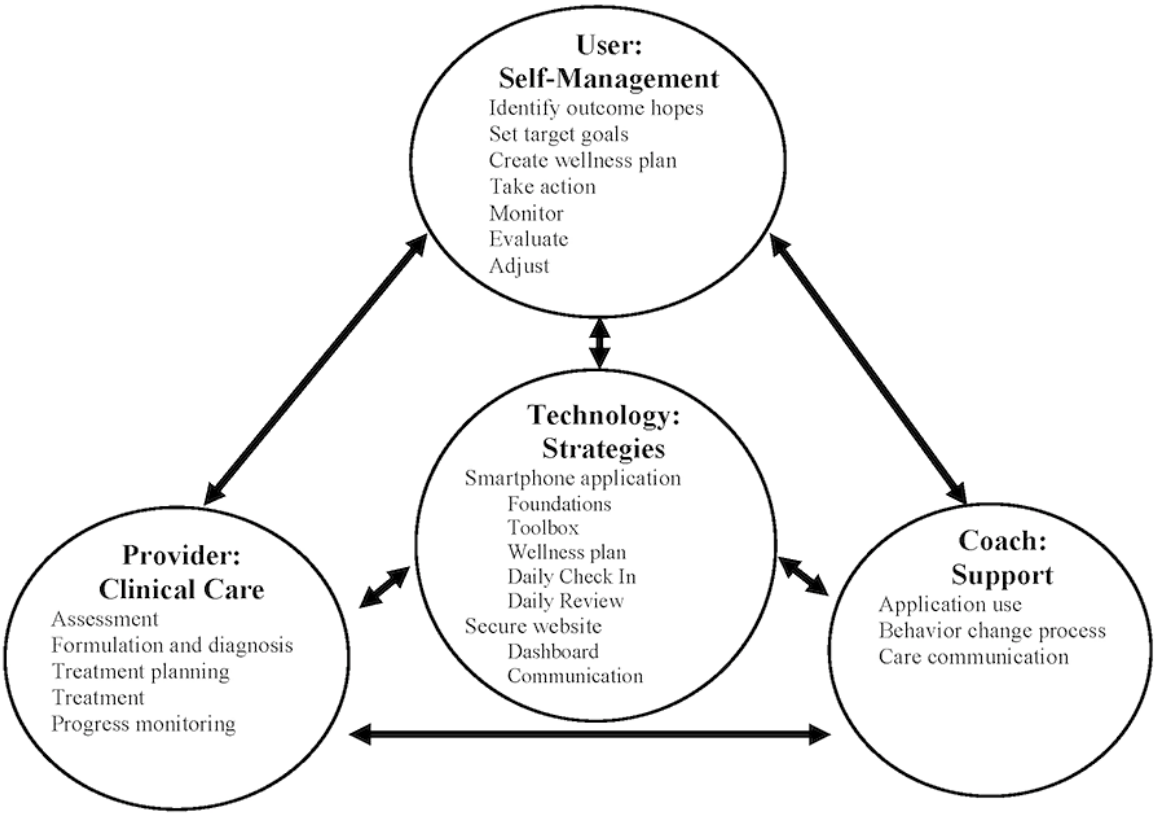
To extend the accessibility and functionality of psychotherapeutic strategies for bipolar disorder, we developed LiveWell, a smartphone-based self-management intervention. Unfortunately, many MHTs suffer from high attrition rates where users rapidly fail to maintain their use of the technology [18,19]. To address this challenge, human support has been introduced to help individuals utilize MHTs. This support often includes technical support to ensure that the technology is working as intended and use support to ensure that the technology is being used. It may also include clinical support to ensure that users identify the content and tools relevant to their needs, use them correctly, and translate this use into their daily lives [20]. Providing technical and use support has been shown to reduce attrition and improve adherence with MHTs, but this increased engagement is not always accompanied by improved outcomes [18,20-24]. Including clinical support may result in improved outcomes [20]. However, limited description of what this clinical support involves has made it difficult to replicate, improve, and implement this type of human support for MHTs [20,25-27].

To facilitate ongoing efforts to develop human support for MHTs, this paper describes the design decisions, theoretical framework, content, mode, and timing of delivery, as well as the training and supervision for coaching support, of the LiveWell technology. The design and development of LiveWell was guided by an intervention mapping and person-centered approach utilizing an iterative strategy (G. Jonathan et al, unpublished manuscript, 2021) [28-33]. This included an initial field trial of a simple self-monitoring application followed by the development of the complete self-management intervention (NCT02405117) via design interviews, usability testing, and a pilot trial. The final intervention design is being tested in a randomized control trial (NCT03088462). The supportive accountability model, which focuses on application use support, served as the starting point for the coaching development process (G. Jonathan et al, unpublished manuscript, 2021). However, F2F therapies for bipolar disorder involve high levels of personalization to meet individuals' varying clinical needs, symptom states, and commitment to self-management. As such, additional support roles were created to better meet the needs of people with bipolar disorder and potentially improve both adherence and outcomes. Specifically, 3 structured coach roles were developed to support (1) application use, (2) self-management, and (3) communication with clinical care providers.

Division of Labor

To reduce costs and increase access, human support for MHTs often involves personnel without professional training in mental health care [34-36]. To facilitate the use of lay personnel, the LiveWell coach's responsibilities are clearly defined in relation to the user, provider, and technology (Figure 1). The coach does not serve as a therapist but instead facilitates the use of the technology. Empirically supported psychotherapeutic strategies, such as providing information and assessments, goal setting, planning, monitoring, and suggesting skills to practice and adjustments to make, are embedded in the LiveWell technology. To ensure the coach operates within the scope of nonclinical practice, there is a clear division of labor between the technology and the coach. The technology operates as the psychotherapeutic strategy expert and provides status summaries and alerts to the coach who uses flowsheets and structured scripts to serve as a technology use concierge.

Figure 1. LiveWell division of labor.



User and Provider

LiveWell is an adjunctive intervention and requires users to be actively engaged with a psychiatrist. The psychiatrist is expected to work with the user to come to a mutual understanding of clinical problems and treatment plans and engage in active and sustained collaborative treatment and progress monitoring. Consistent with this chronic disease self-management model of care [37-48], the LiveWell intervention aims to support the user in learning and utilizing appropriate self-management strategies, including effective communication with their provider. The intervention also aims to assist the provider by delivering clinical information and alerts based on real-time user assessments. Overall, LiveWell seeks to support the functioning of the user-provider dyad to reduce problems caused by bipolar disorder and support the achievement of life goals.

Technology

The LiveWell technology supports the user in learning about and engaging in empirically supported self-management strategies through the delivery of information and assessments, provision of goal setting, planning, and monitoring tools, as well as dynamically suggesting materials to review, skills to practice, and adjustments to make (Figure 1). It presents foundational information on bipolar disorder and self-management (Foundations), a toolbox with a variety of self-assessment surveys and skills instructions (Toolbox), and enables the development of a personalized plan for living well (Wellness Plan). The core of the intervention is a Daily Check-In, where users are asked to monitor medication adherence, sleep duration, routine, and wellness ratings. Based on data from the Daily Check-In, an expert system (Daily Review) provides interactive, personalized feedback and directs individuals to relevant aspects of the application. Users also complete a Weekly Check-In of depressive and manic symptoms: the 8-item Patient Health Questionnaire (PHQ-8) and Altman Self-Rating Mania Scale (ASRM) [49,50]. Users, providers, and coaches can access a secure website that contains a summary of the self-assessment data. The expert system also provides email alerts to providers and coaches when clinical support may be needed.

Coach

The LiveWell coach supports application use adherence, self-management, and clinical care communication. The technology provides data summaries and alerts that direct the coach toward areas to be addressed. All coach-user conversations utilize structured scripts and protocols, which were refined based on user and coach feedback. Coaching starts with a F2F meeting that addresses how LiveWell can help the user utilize self-management strategies to live well with bipolar disorder (Multimedia Appendix 1, Application Training Script). Coaches then complete scheduled calls with users to encourage application use and support self-management (Multimedia Appendix 1, Scheduled Call Scripts). In addition, coaches may be prompted by the technology to contact users regarding application use adherence problems or to facilitate communication with providers when users' self-assessments indicate issues with treatment adherence, the presence of early warning signs, or worsening or severe symptoms (Multimedia Appendix 1, Ad Hoc Calls Script).

Application Use

Rationale

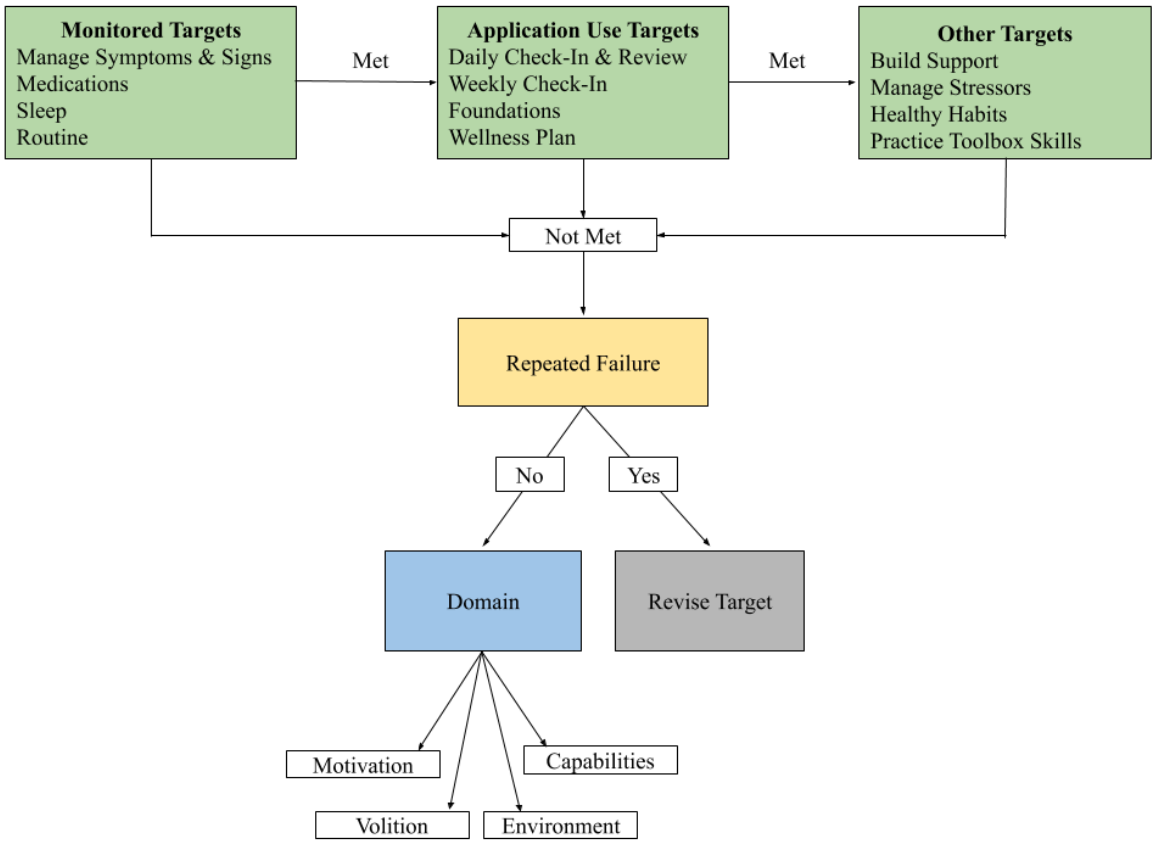
Systematic reviews of MHTs demonstrate that technical and application use support improves adherence and decreases attrition [51-53]. To develop this support role for LiveWell, we used the supportive accountability model, which is effective and proposes that accountability to a coach perceived as trustworthy, benevolent, and knowledgeable enhances adherence to MHTs (G. Jonathan et al, unpublished manuscript, 2021) [54]. To build accountability, the coach collaborates with the user to form clear, process-oriented expectations regarding how the use of the technology might help them achieve intrinsically motivated hopes related to the desired intervention outcomes, including decreasing relapse risk and interepisode symptoms as well as improving quality of life.

Implementation

The initial F2F coach meeting (60-75 minutes long) aims to establish a relationship with the user in which the coach is viewed as a trustworthy, benevolent, and knowledgeable support. The meeting starts with introductions followed by elicitation of the user's hopes ("What would you like to be different at the end of this program?"). The idea that the LiveWell technology can aid in using self-management strategies to stay well is introduced and discussed. The meeting's focus then shifts to support the user in engaging with the technology to learn about and utilize self-management strategies. The F2F meeting ends with explaining the coach's roles, including a clear summary of expectations regarding application use and when the coach might contact the user about use adherence or clinical care concerns. The coach ends the meeting by encouraging commitment to application use and acceptance of adherence monitoring to reinforce accountability.

During scheduled follow-up calls, the coach summarizes application use adherence since the previous call (Daily and Weekly Check-Ins completed; Foundation lessons, Daily Reviews and Toolbox content viewed). The coach then supports adherence and explores barriers to use when nonadherence occurs (Figure 2). The coach may also contact users (via calls, text messages, or emails) when the technology delivers nonadherence alerts to the coach. These alerts are triggered by the user missing Daily Check-Ins (≥ 3 times in a week) or Weekly Check-Ins (≥ 2 weeks in a row). During these ad hoc contacts, coaches elicit barriers to use and direct the user to sections of the application that may help them overcome these barriers. Coaches also provide ad hoc technical support to update application content and manage general technical issues.

Figure 2. Scheduled coach calls.



Self-Management

Rationale

Recent randomized trials suggest that MHT human support focused on technical and application use support does not consistently improve clinical outcomes [18,20-24]. Therefore, coach roles may need to be expanded to include clinical support. For example, the efficiency model of support suggests additional assistance is necessary for psychoeducation and implementation of skills, as well as alignment of these skills with users' unique needs [20]. Users with bipolar disorder may also specifically benefit from guidance in navigating self-management challenges [55]. To develop self-management support for LiveWell, we integrated information from empirically supported psychotherapies for bipolar disorder, the health psychology behavior change literature, and chronic disease self-management models [25,26,30-32,37,38,40,41,46-48,56-64].

LiveWell guides users to address specific behavioral targets, including medication adherence, obtaining adequate sleep duration, and maintaining regular routines. LiveWell emphasizes the importance of identifying early warning signs and symptoms of relapse, as well as developing and implementing a plan for managing these signs and symptoms. In addition, strengthening social support, managing stressors, and healthy habits regarding diet, exercise, and substance use are also addressed. These targets were selected because they have been proposed to underlie the improved outcomes (ie, reduced episode recurrence and interepisode symptoms, and improved quality of life) produced by empirically supported therapies for bipolar disorder [8,57,65].

To facilitate changes in these behavioral targets, the technology and coach deliver behavior change techniques (BCTs) that constitute the smallest intervention components impacting behavioral regulation [25,26,30-32,58-62]. BCTs can be grouped into nonoverlapping clusters hypothesized to alter specific behavioral determinants involved in enacting target behaviors [61,62,66-69]. Determinants and their corresponding techniques can be grouped into 4 domains: motivational determinants involved in developing an intention to engage in a behavior, volitional determinants involved in enacting the behavior, and environmental determinants and capabilities that impact both motivational and volitional processes. When difficulties arise for users in achieving their target goals, attention to these domains allows coaches to guide users to application content that addresses barriers and suggests solutions to reach their target goals.

The coach provides self-management support through behavior change counseling, a simplified adaptation of motivational interviewing that is effective with brief consultations administered by individuals without professional training in mental health care [70,71]. The goal is to encourage users to express commitment to and confidence in developing, enacting, and adjusting their desired behavior change plans. The coach provides support for setting appropriate target goals, personalizing a wellness rating scale and plan, and monitoring progress. When target goals are not met, the coach offers application content guidance to explore barriers and find solutions. When target goals are met, the coach reinforces target success and links this success to outcome hopes.

Implementation

At the F2F meeting, the coach works with the user to review their experiences with normal ups and downs, early warning signs and symptoms, episodes, and crisis situations. This review leads to the development of a personalized 9-point wellness rating scale to facilitate self-monitoring. The coach walks the user through the application and has the user complete a Daily Check-In and Daily Review. As part of this practice, the user sets specific behavioral targets for medication adherence, sleep duration, routine bedtime and rise time, and wellness rating range. The coach encourages the user to set parameters known to facilitate health, such as aiming to take their medications 100% of the time, sleep the recommended amount each day (7 to 9 hours, or 6 to 10 hours acceptable for some), go to bed and start their day within a 1.5-hour window, and keep their wellness ratings within a "balanced" range (with expected ups and downs due to routine events) [8,57,65,72].

Six scheduled coaching calls occur during weeks 1-4, 6, and 16. Each call lasts about 15 minutes, except for a longer week-4 call (~30 minutes), during which the Wellness Plan is personalized. Before each call, the coach reviews a dashboard that summarizes the user's application use and the percent of days their personalized target goals were met. The coach uses this summary and a flowsheet (Figure 2) in conjunction with call scripts to guide discussion and balance user autonomy needs with target priorities.

Each call starts with agenda-setting and invites collaboration. Next, the coach and user review progress toward the user's personalized target goals. If users fail to meet their personalized goals, the flowsheet and scripts provide tips to help the coach focus on behavior change domains relevant to their current situation. The coach encourages gentle exploration, aware that motivational, volitional, environmental, or capability determinants may have hindered progress. A tip sheet directs coaches to application sections relevant to users' specific behavioral target goals and behavior change determinants (Multimedia Appendix 1). If the report indicates that the user successfully met their target goals, the coach provides reinforcement and asks how this success might help achieve their outcome hopes. If application use adherence was not maintained, this is addressed. Otherwise, the coach works with the user to identify other personalized goals for using the application, such as building supports, managing stressors, creating or maintaining healthy habits, and practicing toolbox skills (Figure 2).

During the week-4 call, the coach helps the user personalize their wellness plan by guiding them to application content relevant to developing task and coping plans for each behavioral target. The user is also given an opportunity to make changes to their wellness rating scale and adjust their target goals. Going forward, the coach encourages the user to use the Wellness Plan to reduce risk when feeling well and take action when experiencing symptoms. At the end of each call, based on target progress during the prior week and the user's priorities, the user and the coach collaboratively set a target goal for the week

ahead. Users are urged to link their target behavior goals with their initial outcome hopes and are also asked to anticipate any potential obstacles and ways to overcome them. To assess users' experience of the psychoeducational materials, the coach also asks users if they had any thoughts, feelings, or questions about the materials. Each call ends with a recap of the personalized target goals for the week ahead. In summary, the coach acts as a concierge; directed by a simplified data summary, a flowsheet, and structured scripts, the coach suggests appropriate application content and tools to match users' needs in addressing their target goals.

To provide a detailed view of the BCTs that may be delivered by the coach, the F2F application training and scheduled call scripts were coded using published protocols for coding BCTs [62,73,74]. This coding allows the percent of pages (PP) with content addressing specific outcomes, targets, determinants, domains, and BCTs to be quantified (Multimedia Appendix 2). In addition, the BCTs were ranked (Rank) based on their coding frequency; the top 15 ranked BCTs are displayed in Table 1. In terms of outcomes, symptoms and episodes were most commonly addressed in the coaching script content (69%). In terms of targets, over half of the coaching script content addressed managing symptoms and signs and building supports (63%). In terms of determinants, support, planning, self-efficacy, intention, attitudes, and perceptions constituted the bulk of the content (70%), with agenda mapping, emphasize autonomy, coping planning, open-ended and desire-ability-reason-need questions, and review behavior goals being the top 5 ranked BCTs (42%).

Table 1. LiveWell coaching script content.

Domain (PP[a]), determinant (PP), and technique (PP)	Rank[b]
Motivation (39.4)	
Self-efficacy (12.3)	
Emphasize autonomy (8.5)	2
Affirmation (3.8)	8
Intention (8.9)	
Agenda mapping (8.9)	1
Attitudes & perceptions (9.2)	
Desire-ability-reason-need questions (4.8)	5
Elicit-provide-elicit (2.4)	12
Information about health consequences (2.0)	14
Insight (4.5)	
Guided discovery (4.5)	6
Knowledge (4.5)	
Information about a health condition (4.5)	6
Volition (25.9)	
Planning (13.3)	
Coping planning (7.5)	3
Task planning (3.8)	8
Consider change options (2.0)	14
Adjustment (4.8)	
Review behavior goals (4.8)	5
Goal setting (2.9)	
Process goal (2.9)	10
Evaluation (2.6)	
Feedback on behavior (2.6)	11
Monitoring (2.3)	
Self-monitoring of behavior (2.3)	13
Environment (24.9)	
Support (21.1)	
Open-ended questions (7.2)	4
Permission to provide information and advice (4.4)	7
Social support (practical; 3.8)	8
Social support (unspecified; 3.4)	9
Support change/persistence (2.3)	13
Prompts (2.0)	
Introduce cues (2.0)	14
Reinforcement (1.8)	
Social reward (1.8)	15

[a]PP: percent of total coaching script pages.

[b]Rank: behavioral change techniques rank-ordered based on their PP, BCTs with same PP given same rank-order.

Care Linkage

Rationale

Timeliness of care, especially during the onset of early warning signs, can significantly impact the course of a manic or depressive episode [75]. Additionally, increased communication and collaboration with a care provider are associated with reduced suicide risk, better medication adherence, and greater satisfaction with care [76,77]. For individuals with bipolar disorder, research shows that the use of a chronic disease self-management model reduces the number of weeks spent in manic episodes and improves overall functioning [78,79]. The LiveWell intervention strives to increase early intervention of illness-related problems by enhancing client-provider communication, a common target of empirically supported psychotherapy for bipolar disorder [9,80,81]. We structured the coach care linkage role using the chronic disease self-management model, wherein the coach utilizes active and sustained follow-up to monitor users' status, identify problems, and reinforce progress in implementing the care plan [81,82].

Implementation

LiveWell aims to enhance user-provider communication by prompting users to seek help when necessary and activating providers by offering access to a secure, web-based clinical status summary portal and automated email alerts. Within this system, the coach provides support to both users and providers. When a user experiences poor medication adherence, significant sleep duration changes, or the onset of early warning signs or worsening symptoms, the coach receives an automated email alert. Enrolled providers also receive an automated email or a phone call from the coach based on the provider's preference. When calling providers, the coach summarizes recent check-in data, reminds the provider that a self-assessment data summary is available on the secure portal, and welcomes the provider to reach out with questions.

Additionally, if a user enters a crisis rating on the Daily Check-In (+4 or -4) or has a Weekly Check-In score consistent with the onset of a mood episode (ie, a change from a below- to an above-threshold score: a PHQ-8 score of ≥ 9 or an ASRM score of ≥ 6), the coach receives an email alert [49,50]. In this case, the coach performs a suicidality assessment and functional impairment evaluation (SAFE). The coach utilizes a structured protocol (Multimedia Appendix 1) and flowsheets (Figures 3-4) to guide this risk evaluation and follow-up. After the alert is received, the coach attempts to contact the user up to 3 times within 24 hours to complete the SAFE protocol. If the coach cannot reach the user, the coach reaches out to the user's psychiatrist with available clinical information provided by the user's self-assessments.

Figure 3. Suicidality assessment protocol flowsheet. CS: clinical supervisor.

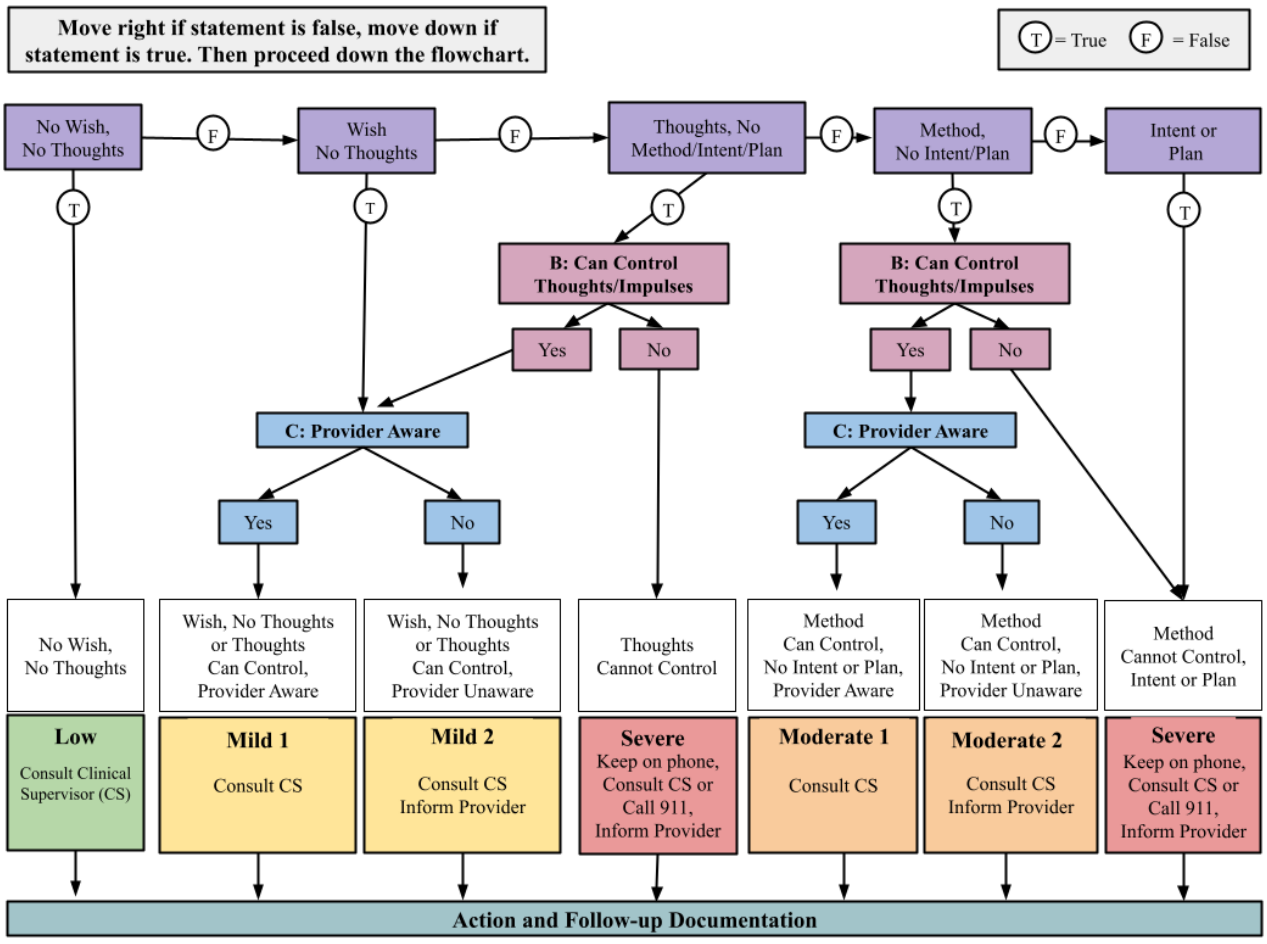
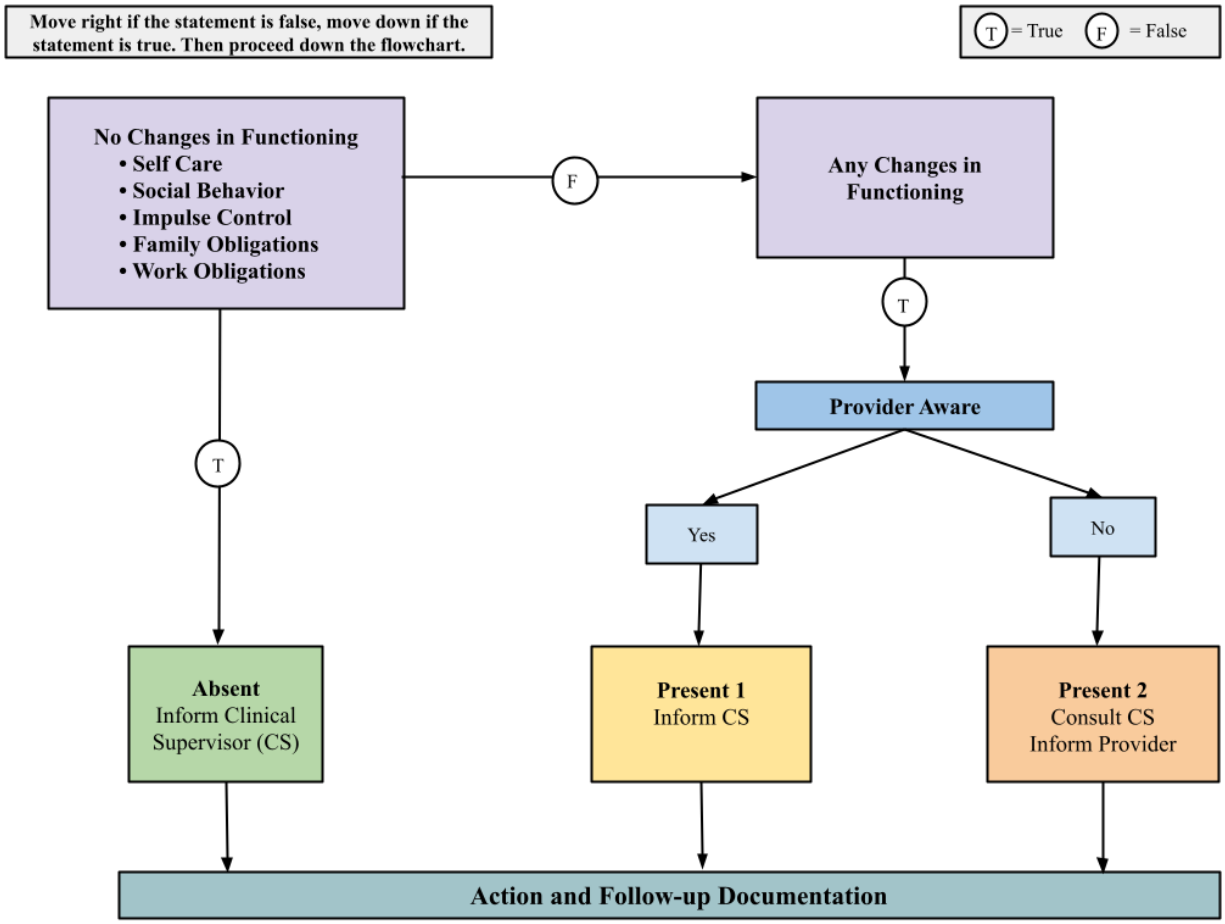


Figure 4. Functional evaluation protocol flowsheet. CS: clinical supervisor.



The suicidality assessment (SA) protocol contains 3 sections: assessing suicidal ideation, the ability to control these thoughts, and the providers' awareness of the information provided. The assessment of suicidal ideation utilizes the Columbia Suicide Severity Rating Scale (C-SSRS) [83]. Once the coach has reached the user by phone, the coach explains that they are calling due to the alert. The coach then explains that they are following up based on the safety monitoring discussed at the F2F meeting. The coach starts by asking 2 questions from the C-SSRS [83]: "Have you wished you were dead or wished you could go to sleep and not wake up?" (Wish); "Have you actually had any thoughts about killing yourself?" (Thoughts). Based on the user's response to these 2 questions, the coach then follows the logic of the SA protocol and flowsheet (Figure 3) to determine the user's risk level (low to severe risk), the provider's awareness, and the actions that the coach needs to take. This protocol allows coaches to proactively implement a standardized suicidal ideation evaluation, efficiently communicate the information they obtain, and promptly receive clinician support.

The functional impairment evaluation (FE), which draws upon components of the World Health Organization Quality of Life Questionnaire, Lehman's Quality of Life Interview, and the Global Assessment of Functioning [84-86], assesses users' life circumstances to evaluate impairment in daily activities. The user is asked if there is any new onset of changes or problems in their self-care, social activities, impulse control, or meeting of family or work obligations. If so, the coach asks if their provider is aware of this information and follows the logic of the FE protocol and flowsheet (Figure 4) to determine the actions to take. This procedure ensures that the coach is not responsible for determining the clinical significance of the reported changes in functioning.

Training and Supervision

Coaches receive initial training and weekly supervision. Prior to a full-day workshop, they are given a training manual (Multimedia Appendix 1), coaching scripts (Multimedia Appendix 1), and several reference papers. The training manual includes information about bipolar disorder and treatment, the LiveWell program, supportive accountability, behavior change processes, and coaching skills and responsibilities. The core training occurs in one full-day workshop. The morning involves a review and discussion of the material presented in the training manual. Case examples are provided to elucidate the material. The afternoon involves learning about the core coaching skills. All coaches engage in role-playing of scheduled calls with a supervising clinical psychologist. They train to mastery, defined as 3 consecutive contacts evaluated at or above 85% satisfactory ratings on an adapted version of the behavior change counseling index (Multimedia Appendix 1) [87]. Once some trained coaches

are available, new coaches also sit in on experienced coaches' initial visits and coaching calls. Experienced coaches then observe the new coaches' interactions to provide direct feedback. Beyond the initial training, weekly group supervision is provided. Supervision includes addressing any questions that arise during the week and feedback on audiotapes of coach F2F training or calls randomly selected for review.

Coaches also receive additional crisis support training. Supervisors and experienced coaches meet with new coaches and review the SAFE protocol and situations in which the protocol is activated. Coaches listen to 6 selected audio-recorded calls covering different risk levels during which the protocol was activated and successfully administered. They also shadow 3 live crisis calls and complete 3 calls under the supervision of an experienced coach. The training concludes with a protocol quiz regarding risk assessment in sample scenarios (Multimedia Appendix 1). New coaches are required to receive a score of 100%.

Themes and subthemes from the supervision of coaches (*N*=7) trained during our recent studies identify the importance of teaching lay personnel how to conduct structured interactions with users. As outlined in Table 2, posttraining supervision needs fell into 3 categories: (1) general counseling, (2) behavior change processes, and (3) LiveWell content. In the area of general counseling, coaches need additional instruction on when and how to point users back to parts of the application, how to respond to negative life events in an empathic way without leading to a deepening exploration of the issue, and how to delicately manage times when the perceptions of the user and coach are at odds. In regard to behavior change processes, coaches also receive further guidance around differentiating overall hopes for LiveWell and specific behavioral targets, how to set target goals as well as task and coping plans, and how to demonstrate understanding and reinforce change. Early on in training, coaches often try to directly engage users in problem-solving before assessing the obstacles involved; supervision engages users in thinking through difficulties in a more systematic and guided way. Supervision also helps coaches respond when users are unable or unmotivated to work on a specific behavioral target. Further, regarding the LiveWell content, coaches and the supervisor agree that an initial training followed by staged training is most effective. Staged training supervision sessions focus on Daily Check-Ins, the Wellness Plans, Foundation lessons, and the Toolbox. It appears that within 3 months, routine supervision feedback for new coaches is expended, and they can navigate encounters with mastery.

Table 2. Supervision themes and subthemes.

Themes and subthemes
General counseling
Maintaining integrity of coaching role
Establishing and maintaining boundaries
Responding to reports of negative life events
Managing discrepant coach-participant perceptions
Hows and whys of behavior change
Differentiating between hopes and targets
Setting goals, task, and coping plans
Understanding and reinforcing change
Assessing challenges before problem-solving
Reacting to inability or amotivation to work on specific targets
LiveWell content
Staged training
Suggestions guide
Daily Check-in
Wellness Plan
Foundations
Toolbox

Discussion

This paper describes the design decisions, theoretical framework, content, mode, timing of delivery, and the training and supervision for coaching support of the LiveWell technology. The development of this support initially focused on using the supportive accountability model to facilitate technology use adherence and reduce attrition. However, during the expansion of the application from a simple self-monitoring tool to a complete self-management application, feedback from users and coaches suggested that the intervention might benefit from expanding the coaching roles. Thus, to support the clinical needs of individuals with bipolar disorder and potentially improve

outcomes, coaching was expanded to include supporting self-management and communication with care providers.

To develop the LiveWell coaching support, we combined information from empirically supported psychotherapies for bipolar disorder, the health psychology behavior change literature, and chronic disease self-management models. In addition to guiding the development of the clinical support roles, this integration provides the ability to label, measure, and relate changes in (1) target behaviors proposed to improve outcomes, (2) behavioral determinants proposed to govern enactment of target behaviors, and (3) exposure to behavior change technique content and tool use proposed to alter behavioral determinants. This integrated approach to developing the coaching roles was initially used to develop the smartphone application content and derived from using an intervention mapping approach. Hopefully, using the same process to develop and investigate the mechanisms of action of the technology and its coaching support will provide a generally applicable means of further understanding how supported MHTs work and clarify how best to balance intervention delivery via technology and human support.

In developing the coaching support for LiveWell, we aimed to define a clear division of labor between the coaching roles and the technology to facilitate the use of personnel without professional training in mental health care. The use of a chronic disease self-management model was useful in defining the coaching support tasks in relation to the technology, the user, and the provider. To assist the coach, reports summarizing progress on clearly defined and monitored behavioral target goals and clinical alerts are provided by the technology. This information, combined with structured scripts, protocols, and flowsheets, allows the coach to deliver clinical support without extensive mental health training. Consequently, this should improve accessibility, as lay personnel can reduce costs and increase access to MHTs.

Obtaining feedback from users and coaches played an important role in developing the division of labor between the technology and coaching support. To facilitate coaches' use of the reports and alerts, users and coaches required a clear understanding of the behavioral target goals and the rationale for the clinical alerts. As a result, user and coach feedback led to multiple iterative revisions of the technology content and the coaching reports and alerts to ensure that the coaches felt confident in acting on the information provided by the technology. Developing straightforward, structured scripts and protocols for delivering self-management support and communicating with care providers was also essential in allowing the coaches to carry out these roles. In addition, the use of flowsheets appeared to assist lay personnel with learning the protocols.

The development of clear flowsheets and protocols was especially significant in allowing lay personnel to deliver the suicidality assessment and functional evaluation without undue distress. It is important to note that the use of lay personnel in this assessment role requires the availability of a trained clinical professional to provide real-time support to the coaches when severe suicidal ideation is detected. This need for real-time professional clinical support for the coaches and a mental health professional to deliver the coach training requires careful consideration when planning for implementing and disseminating supported MHT interventions. This consideration is especially relevant when these interventions are directed toward high-risk populations such as individuals with bipolar disorder.

MHT coaching is a nascent field. As research develops, we anticipate that the coaches' effectiveness at improving adherence and clinical outcomes produced by MHT interventions will increase. During the design and development of the LiveWell intervention, using the same framework to develop the technology's content and tools and its coaching support played an important role in enabling a clear division of labor to be defined between the technology and the coach. In addition, obtaining feedback from both users and coaches was critical in developing both the technology and the coaching support and assisting coaches without professional mental health training to learn and deliver the coach roles. We hope that providing a detailed explanation of the rationale and implementation of our coaching support roles for LiveWell will contribute to the development of more effective coaching support for MHT interventions.

Conflicts of Interest

DCM has accepted honoraria and consulting fees from Apple Inc, Otsuka Pharmaceuticals, Pear Therapeutics, and the One Mind Foundation, royalties from Oxford Press, and has an ownership interest in Adaptive Health Inc. EHG has accepted honoraria from Otsuka Pharmaceuticals. The other authors have no conflicts of interest to declare.

Multimedia Appendix 1
LiveWell coaching materials.
[PDF File (Adobe PDF File), 3326 KB - formative_v5i3e25810_app1.pdf]

Multimedia Appendix 2
Coding breakdown of the coaching scripts.
[XLSX File (Microsoft Excel File), 123 KB - formative_v5i3e25810_app2.xlsx]

References

1. Eaton WW, Martins SS, Nestadt G, Bienvenu OJ, Clarke D, Alexandre P. The burden of mental disorders. Epidemiol Rev 2008;30:1-14 [FREE Full text] [doi: 10.1093/epirev/mxn011] [Medline: 18806255]
2. Dilsaver SC. An estimate of the minimum economic burden of bipolar I and II disorders in the United States: 2009. J Affect Disord 2011 Mar;129(1-3):79-83. [doi: 10.1016/j.jad.2010.08.030] [Medline: 20888048]
3. World Health Organization. Burden of Mental and Behavioural Disorders. In: The World Health Report 2001: Mental health: new understanding, new hope. Geneva, Switzerland: The World Health Organization; 2001.
4. Crowe M, Whitehead L, Wilson L, Carlyle D, O'Brien A, Inder M, et al. Disorder-specific psychosocial interventions for bipolar disorder--a systematic review of the evidence for mental health nursing practice. Int J Nurs Stud 2010 Jul;47(7):896-908. [doi: 10.1016/j.ijnurstu.2010.02.012] [Medline: 20233615]
5. Keck PE, McElroy SL, Strakowski SM, West SA, Sax KW, Hawkins JM, et al. 12-month outcome of patients with bipolar disorder following hospitalization for a manic or mixed episode. Am J Psychiatry 1998 May;155(5):646-652. [doi: 10.1176/ajp.155.5.646] [Medline: 9585716]
6. Judd LL, Akiskal HS, Schettler PJ, Endicott J, Maser J, Solomon DA, et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. Arch Gen Psychiatry 2002 Jun;59(6):530-537. [doi: 10.1001/archpsyc.59.6.530] [Medline: 12044195]
7. Miklowitz DJ, Otto MW, Frank E, Reilly-Harrington NA, Wisniewski SR, Kogan JN, et al. Psychosocial treatments for bipolar depression: a 1-year randomized trial from the Systematic Treatment Enhancement Program. Arch Gen Psychiatry 2007 Apr;64(4):419-426 [FREE Full text] [doi: 10.1001/archpsyc.64.4.419] [Medline: 17404119]
8. Miklowitz DJ, Goodwin GM, Bauer MS, Geddes JR. Common and specific elements of psychosocial treatments for bipolar disorder: a survey of clinicians participating in randomized trials. J Psychiatr Pract 2008 Mar;14(2):77-85 [FREE Full text] [doi: 10.1097/01.pra.0000314314.94791.c9] [Medline: 18360193]
9. Miklowitz DJ. Adjunctive psychotherapy for bipolar disorder: state of the evidence. Am J Psychiatry 2008 Nov;165(11):1408-1419 [FREE Full text] [doi: 10.1176/appi.ajp.2008.08040488] [Medline: 18794208]
10. Busch AB, Ling D, Frank RG, Greenfield SF. Changes in the quality of care for bipolar I disorder during the 1990s. Psychiatr Serv 2007 Jan;58(1):27-33 [FREE Full text] [doi: 10.1176/ps.2007.58.1.27] [Medline: 17215409]
11. Wang PS, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC. Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):629-640. [doi: 10.1001/archpsyc.62.6.629] [Medline: 15939840]
12. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile Phone Ownership and Endorsement of "mHealth" Among People With Psychosis: A Meta-analysis of Cross-sectional Studies. Schizophr Bull 2016 Mar 22;42(2):448-455 [FREE Full text] [doi: 10.1093/schbul/sbv132] [Medline: 26400871]
13. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct 21;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
14. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
15. Proudfoot J, Parker G, Manicavasagar V, Hadzi-Pavlovic D, Whitton A, Nicholas J, et al. Effects of adjunctive peer support on perceptions of illness control and understanding in an online psychoeducation program for bipolar disorder: a randomised controlled trial. J Affect Disord 2012 Dec 15;142(1-3):98-105. [doi: 10.1016/j.jad.2012.04.007] [Medline: 22858215]
16. Hidalgo-Mazzei D, Mateu A, Reinares M, Matic A, Vieta E, Colom F. Internet-based psychological interventions for bipolar disorder: Review of the present and insights into the future. J Affect Disord 2015 Dec 01;188:1-13. [doi: 10.1016/j.jad.2015.08.005] [Medline: 26342885]
17. Hidalgo-Mazzei D, Mateu A, Reinares M, Undurraga J, Bonnín CDM, Sánchez-Moreno J, et al. Self-monitoring and psychoeducation in bipolar patients with a smart-phone application (SIMPLe) project: design, development and studies protocols. BMC Psychiatry 2015 Mar 20;15:52 [FREE Full text] [doi: 10.1186/s12888-015-0437-6] [Medline: 25884824]
18. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. Gen Hosp Psychiatry 2013;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]
19. Mohr D, Vella L, Hart S, Heckman T, Simon G. The Effect of Telephone-Administered Psychotherapy on Symptoms of Depression and Attrition: A Meta-Analysis. Clin Psychol (New York) 2008;15(3):243-253 [FREE Full text] [doi: 10.1111/j.1468-2850.2008.00134.x] [Medline: 21369344]
20. Schueller S, Tomasino K, Mohr D. Integrating Human Support Into Behavioral Intervention Technologies: The Efficiency Model of Support. Clin Psychol Sci Pract 2016 Nov 17;24(1):27-45. [doi: 10.1111/cpsp.12173]
21. Kelders SM, Bohlmeijer ET, Pots WTM, van Gemert-Pijnen JEWC. Comparing human and automated support for depression: Fractional factorial randomized controlled trial. Behav Res Ther 2015 Sep 17;72(1):72-80. [doi: 10.1016/j.brat.2015.06.014] [Medline: 26196078]

22. Kleiboer A, Donker T, Seekles W, van Straten A, Riper H, Cuijpers P. A randomized controlled trial on the role of support in Internet-based problem solving therapy for depression and anxiety. Behav Res Ther 2015 Sep;72(1):63-71. [doi: 10.1016/j.brat.2015.06.013] [Medline: 26188373]
23. Mira A, Bretón-López J, García-Palacios A, Quero S, Baños R, Botella C. An Internet-based program for depressive symptoms using human and automated support: a randomized controlled trial. NDT 2017 Mar;Volume 13:987-1006. [doi: 10.2147/ndt.s130994] [Medline: 26188373]
24. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009 Apr 24;11(2):e13-1006 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
25. Brendryen H, Kraft P, Schaalma H. Looking Inside the Black Box: Using Intervention Mapping to Describe the Development of the Automated Smoking Cessation Intervention 'Happy Ending'. J Smok Cessat 2012 Feb 21;5(1):29-56. [doi: 10.1375/jsc.5.1.29]
26. Brendryen H, Johansen A, Nesvåg S, Kok G, Duckert F. Constructing a Theory- and Evidence-Based Treatment Rationale for Complex eHealth Interventions: Development of an Online Alcohol Intervention Using an Intervention Mapping Approach. JMIR Res Protoc 2013 Jan 23;2(1):e6 [FREE Full text] [doi: 10.2196/resprot.2371] [Medline: 23612478]
27. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May 21;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
28. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
29. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
30. Bartholomew LK, Parcel GS, Kok G. Intervention mapping: a process for developing theory- and evidence-based health education programs. Health Educ Behav 1998 Oct;25(5):545-563. [doi: 10.1177/109019819802500502] [Medline: 9768376]
31. Kok G, Mesters I. Getting inside the black box of health promotion programmes using intervention Mapping. Chronic Illn 2011 Sep;7(3):176-180. [doi: 10.1177/1742395311403013] [Medline: 21900338]
32. Kok G, Schaalma H, Ruiter RAC, van Empelen P, Brug J. Intervention mapping: protocol for applying health psychology theory to prevention programmes. J Health Psychol 2004 Jan;9(1):85-98. [doi: 10.1177/1359105304038379] [Medline: 14683571]
33. Jonathan G, Dopke CA, McBride A, Babington P, Michaels T, Martin CM, et al. A Smartphone-Based Self-management Intervention for Bipolar Disorder (LiveWell): User-Centered Development Approach. JMIR Mental Health 2021 Jan 26:e24024 (forthcoming). [doi: 10.2196/20424]
34. O'Mahen H, Richards D, Woodford J, Wilkinson E, McGinley J, Taylor R, et al. Netmums: a phase II randomized controlled trial of a guided Internet behavioural activation treatment for postpartum depression. Psychol. Med 2013 Oct 09;44(8):1675-1689. [doi: 10.1017/s0033291713002092]
35. Robinson E, Titov N, Andrews G, McIntyre K, Schwencke G, Solley K. Internet treatment for generalized anxiety disorder: a randomized controlled trial comparing clinician vs. technician assistance. PLoS One 2010 Jun 03;5(6):e10942 [FREE Full text] [doi: 10.1371/journal.pone.0010942]
36. Titov N, Andrews G, Davies M, McIntyre K, Robinson E, Solley K. Internet treatment for depression: a randomized controlled trial comparing clinician vs. technician assistance. PLoS One 2010 Jun 08;5(6):e10939-e11689 [FREE Full text] [doi: 10.1371/journal.pone.0010939] [Medline: 20544030]
37. Clark NM. Management of chronic disease by patients. Annu Rev Public Health 2003 Jun 3;24(6):289-313. [doi: 10.1146/annurev.publhealth.24.100901.141021] [Medline: 12415147]
38. Clark NM, Gong M, Kaciroti N. A model of self-regulation for control of chronic disease. Health Educ Behav 2001 Dec;28(6):769-782 [FREE Full text] [doi: 10.1177/109019810102800608] [Medline: 11720277]
39. Clark NM, Zimmerman BJ. A social cognitive view of self-regulated learning about health. Health Educ Behav 2014 Oct;41(5):485-491. [doi: 10.1177/1090198114547512] [Medline: 25270173]
40. Von Korff M, Gruman J, Schaefer J, Curry SJ, Wagner EH. Collaborative management of chronic illness. Ann Intern Med 1997 Dec 15;127(12):1097-1102. [doi: 10.7326/0003-4819-127-12-199712150-00008] [Medline: 9412313]
41. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. Milbank Q 1996;74(4):511-544. [Medline: 8941260]
42. Clark NM, Gong M. Management of chronic disease by practitioners and patients: are we teaching the wrong things? BMJ 2000 Feb 26;320(7234):572-575 [FREE Full text] [doi: 10.1136/bmj.320.7234.572] [Medline: 10688569]
43. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/S15324796ABM2601_01] [Medline: 12867348]
44. Wagner EH. Chronic disease management: what will it take to improve care for chronic illness? Eff Clin Pract 1998;1(1):2-4. [Medline: 10345255]
45. Wagner EH. Counterpoint: chronic illness and primary care. Med Care 2011 Nov;49(11):973-975. [doi: 10.1097/MLR.0b013e318233a0d3] [Medline: 22005603]

46. Salyers MP, Godfrey JL, McGuire AB, Gearhart T, Rollins AL, Boyle C. Implementing the illness management and recovery program for consumers with severe mental illness. Psychiatr Serv 2009 Apr;60(4):483-490. [doi: 10.1176/appi.ps.60.4.483] [Medline: 19339323]
47. Lorig KR, Sobel DS, Stewart AL, Brown BW, Bandura A, Ritter P, et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. Med Care 1999 Jan;37(1):5-14. [doi: 10.1097/00005650-199901000-00003] [Medline: 10413387]
48. Bartels SJ, Pratt SI, Mueser KT, Naslund JA, Wolfe RS, Santos M, et al. Integrated IMR for psychiatric and general medical illness for adults aged 50 or older with serious mental illness. Psychiatr Serv 2014 Mar 01;65(3):330-337 [FREE Full text] [doi: 10.1176/appi.ps.201300023] [Medline: 24292559]
49. Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman Self-Rating Mania Scale. Biol Psychiatry 1997 Nov 15;42(10):948-955. [doi: 10.1016/S0006-3223(96)00548-3] [Medline: 9359982]
50. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
51. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
52. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions — A systematic review. Internet Interventions 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003] [Medline: 18752852]
53. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol. Med 2006 Nov 20;37(03):319-205. [doi: 10.1017/s0033291706008944] [Medline: 20183695]
54. Mohr DC, Duffecy J, Ho J, Kwasny M, Cai X, Burns MN, et al. A randomized controlled trial evaluating a manualized TeleCoaching protocol for improving adherence to a web-based intervention for the treatment of depression. PLoS One 2013;8(8):e70086 [FREE Full text] [doi: 10.1371/journal.pone.0070086] [Medline: 23990896]
55. Switsers L, Dauwe A, Vanhoudt A, Van Dyck H, Lombaerts K, Oldenburg J. Users' Perspectives on mHealth Self-Management of Bipolar Disorder: Qualitative Focus Group Study. JMIR Mhealth Uhealth 2018 May 02;6(5):e108 [FREE Full text] [doi: 10.2196/mhealth.9529] [Medline: 29720363]
56. Schöttle D, Huber CG, Bock T, Meyer TD. Psychotherapy for bipolar disorder: a review of the most recent studies. Curr Opin Psychiatry 2011 Nov;24(6):549-555. [doi: 10.1097/YCO.0b013e32834b7c5f] [Medline: 21918448]
57. Miklowitz DJ, Scott J. Psychosocial treatments for bipolar disorder: cost-effectiveness, mediating mechanisms, and future directions. Bipolar Disord 2009 Jun;11 Suppl 2:110-122. [doi: 10.1111/j.1399-5618.2009.00715.x] [Medline: 19538690]
58. Heaney CA. Intervention mapping and the new health promotion. Health Educ Behav 1998 Oct;25(5):564-8; discussion 569. [doi: 10.1177/109019819802500503] [Medline: 9768377]
59. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
60. Cane J, Richardson M, Johnston M, Ladha R, Michie S. From lists of behaviour change techniques (BCTs) to structured hierarchies: comparison of two methods of developing a hierarchy of BCTs. Br J Health Psychol 2015 Feb;20(1):130-150. [doi: 10.1111/bjhp.12102] [Medline: 24815766]
61. Michie S, Johnston M, Francis J, Hardeman W, Eccles M. From theory to intervention: Mapping theoretically derived behavioural determinants to behaviour change techniques. International Association of Applied Psychology 2008 Jul 08;57(4):660-680.
62. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
63. Miller WR, Lasiter S, Bartlett Ellis R, Buelow JM. Chronic disease self-management: a hybrid concept analysis. Nurs Outlook 2015 Oct;63(2):154-161 [FREE Full text] [doi: 10.1016/j.outlook.2014.07.005] [Medline: 25241136]
64. Salyers MP, Godfrey JL, Mueser KT, Labriola S. Measuring illness management outcomes: a psychometric study of clinician and consumer rating scales for illness self management and recovery. Community Ment Health J 2007 Oct;43(5):459-480. [doi: 10.1007/s10597-007-9087-6] [Medline: 17514504]
65. Scott J, Colom F, Vieta E. A meta-analysis of relapse rates with adjunctive psychological therapies compared to usual psychiatric treatment for bipolar disorders. Int J Neuropsychopharmacol 2007 Feb;10(1):123-129 [FREE Full text] [doi: 10.1017/S1461145706006900] [Medline: 16787554]
66. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
67. Michie S, Webb TL, Sniehotta FF. The importance of making explicit links between theoretical constructs and behaviour change techniques. Addiction 2010 Nov;105(11):1897-1898. [doi: 10.1111/j.1360-0443.2010.03161.x] [Medline: 21064254]
68. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. Health Psychol Rev 2016 Sep;10(3):277-296 [FREE Full text] [doi: 10.1080/17437199.2016.1151372] [Medline: 26854092]

69. Sheeran P, Klein WMP, Rothman AJ. Health Behavior Change: Moving from Observation to Intervention. Annu Rev Psychol 2017 Jan 03;68:573-600. [doi: 10.1146/annurev-psych-010416-044007] [Medline: 27618942]
70. Rollnick S, Butler CC, Kinnersley P, Gregory J, Mash B. Motivational interviewing. BMJ 2010 Apr 27;340(3):c1900-c1996 [FREE Full text] [doi: 10.1136/bmj.c1900] [Medline: 20423957]
71. Butler CC, Simpson SA, Hood K, Cohen D, Pickles T, Spanou C, et al. Training practitioners to deliver opportunistic multiple behaviour change counselling in primary care: a cluster randomised trial. BMJ 2013 Mar 19;346:f1191-f1600 [FREE Full text] [doi: 10.1136/bmj.f1191] [Medline: 23512758]
72. Adams J, Scott J. Predicting medication adherence in severe mental disorders. Acta Psychiatr Scand 2000 Feb 27;101(2):119-124. [doi: 10.1034/j.1600-0447.2000.90061.x] [Medline: 10706011]
73. Hardcastle SJ, Fortier M, Blake N, Hagger MS. Identifying content-based and relational techniques to change behaviour in motivational interviewing. Health Psychol Rev 2017 Mar;11(1):1-16. [doi: 10.1080/17437199.2016.1190659] [Medline: 27189713]
74. Blackburn I, James I, Milne D, Baker C, Standart S, Garland A, et al. The Revised Cognitive Therapy Scale (CTS-R): Psychometric Properties. Behav. Cogn. Psychother 2001 Oct 23;29(4):431-446. [doi: 10.1017/s1352465801004040]
75. Morriss RK, Faizal MA, Jones AP, Williamson PR, Bolton C, McCarthy JP. Interventions for helping people recognise early signs of recurrence in bipolar disorder. Cochrane Database Syst Rev 2007 Jan 24(1):CD004854 [FREE Full text] [doi: 10.1002/14651858.CD004854.pub2] [Medline: 17253526]
76. Sylvia LG, Hay A, Ostacher MJ, Miklowitz DJ, Nierenberg AA, Thase ME, et al. Association between therapeutic alliance, care satisfaction, and pharmacological adherence in bipolar disorder. J Clin Psychopharmacol 2013 Jun;33(3):343-350 [FREE Full text] [doi: 10.1097/JCP.0b013e3182900c6f] [Medline: 23609394]
77. Sajatovic M, Davies M, Bauer MS, McBride L, Hays RW, Safavi R, et al. Attitudes regarding the collaborative practice model and treatment adherence among individuals with bipolar disorder. Compr Psychiatry 2005;46(4):272-277. [doi: 10.1016/j.comppsych.2004.10.007] [Medline: 16175758]
78. Bauer MS, McBride L, Williford WO, Glick H, Kinosian B, Altshuler L, Cooperative Studies Program 430 Study Team. Collaborative care for bipolar disorder: part I. Intervention and implementation in a randomized effectiveness trial. Psychiatr Serv 2006 Jul;57(7):927-936. [doi: 10.1176/ps.2006.57.7.927] [Medline: 16816276]
79. Bauer MS, McBride L, Williford WO, Glick H, Kinosian B, Altshuler L, Cooperative Studies Program 430 Study Team. Collaborative care for bipolar disorder: Part II. Impact on clinical outcome, function, and costs. Psychiatr Serv 2006 Jul;57(7):937-945. [doi: 10.1176/ps.2006.57.7.937] [Medline: 16816277]
80. Bauer MS. The collaborative practice model for bipolar disorder: design and implementation in a multi-site randomized controlled trial. Bipolar Disord 2001 Oct;3(5):233-244. [doi: 10.1034/j.1399-5618.2001.30502.x] [Medline: 11903206]
81. Lorig KR, Ritter P, Stewart AL, Sobel DS, Brown BW, Bandura A, et al. Chronic disease self-management program: 2-year health status and health care utilization outcomes. Med Care 2001 Nov;39(11):1217-1223. [doi: 10.1097/00005650-200111000-00008] [Medline: 11606875]
82. Lorig K, Ritter PL, Pifer C, Werner P. Effectiveness of the chronic disease self-management program for persons with a serious mental illness: a translation study. Community Ment Health J 2014 Jan;50(1):96-103. [doi: 10.1007/s10597-013-9615-5] [Medline: 23748554]
83. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry 2011 Dec;168(12):1266-1277 [FREE Full text] [doi: 10.1176/appi.ajp.2011.10111704] [Medline: 22193671]
84. Hall RC. Global assessment of functioning. A modified scale. Psychosomatics 1995;36(3):267-275. [doi: 10.1016/S0033-3182(95)71666-8] [Medline: 7638314]
85. Lehman AF, Postrado LT, Rachuba LT. Convergent validation of quality of life assessments for persons with severe mental illnesses. Qual Life Res 1993 Oct;2(5):327-333. [doi: 10.1007/BF00449427] [Medline: 8136797]
86. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
87. Lane C, Huws-Thomas M, Hood K, Rollnick S, Edwards K, Robling M. Measuring adaptations of motivational interviewing: the development and validation of the behavior change counseling index (BECCI). Patient Educ Couns 2005 Feb;56(2):166-173. [doi: 10.1016/j.pec.2004.01.003] [Medline: 15653245]

Abbreviations

ASRM: Altman Self-Rating Mania Scale
BCT: behavior change technique
C-SSRS: Columbia Suicide Severity Rating Scale
FE: functional impairment evaluation
F2F: face-to-face
MHT: Mental health technology

PHQ-8: 8-item Patient Health Questionnaire
PP: percent of pages
SA: suicidality assessment
SAFE: suicidality assessment and functional impairment evaluation

Edited by G Eysenbach; submitted 16.11.20; peer-reviewed by F Mayoral, D Szinay; comments to author 08.12.20; revised version received 14.12.20; accepted 17.01.21; published 24.03.21.

Please cite as:
Dopke CA, McBride A, Babington P, Jonathan GK, Michaels T, Ryan C, Duffecy J, Mohr DC, Goulding EH
Development of Coaching Support for LiveWell: A Smartphone-Based Self-Management Intervention for Bipolar Disorder
JMIR Form Res 2021;5(3):e25810
URL: https://formative.jmir.org/2021/3/e25810
doi:10.2196/25810
PMID:33759798

Original Paper

System Architecture for "Support Through Mobile Messaging and Digital Health Technology for Diabetes" (SuMMiT-D): Design and Performance in Pilot and Randomized Controlled Feasibility Studies

Yuan Chi[1], BEng, PhD; Carmelo Velardo[1], MSc, PhD; Julie Allen[2], BA; Stephanie Robinson[2], BBA, JD; Evgenia Riga[2], MSc; David Judge[2], BA; Lionel Tarassenko[1], MA, DPhil; Andrew J Farmer[2], DM, FRCGP

[1]Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, United Kingdom

[2]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom

Corresponding Author:
Yuan Chi, BEng, PhD
Institute of Biomedical Engineering
Department of Engineering Science
University of Oxford
Old Road Campus Research Building
Headington
Oxford, OX3 7DQ
United Kingdom
Phone: 44 01865 617675
Email: yuan.chi@eng.ox.ac.uk

Related Article:

This is a corrected version. See correction statement: https://formative.jmir.org/2021/4/e29451/

Abstract

Background: Diabetes is a highly prevalent long-term condition with high morbidity and mortality rates. People with diabetes commonly worry about their diabetes medicines and do not always take them regularly as prescribed. This can lead to poor diabetes control. The Support Through Mobile Messaging and Digital Health Technology for Diabetes (SuMMiT-D) study aims to deliver brief messages as tailored interventions to support people with type 2 diabetes in better use of their diabetes medicines and to improve treatment adherence and health outcomes.

Objective: This paper describes the overall architecture of a tailored intervention delivery system used in the pilot and randomized controlled feasibility studies of SuMMiT-D and reports its performance.

Methods: The SuMMiT-D system includes several platforms and resources to recruit participants and deliver messages as tailored interventions. Its core component is called the *clinical system* and is responsible for interacting with the participants by receiving and sending SMS text messages from and to them. The personalization and tailoring of brief messages for each participant is based on a list of built-in commands that they can use.

Results: For the pilot study, a total of 48 participants were recruited; they had a median age of 64 years (first quartile, third quartile [Q_1, Q_3: 54.5, 69]). For the feasibility study, a total of 209 participants were recruited and randomly assigned to either the control or intervention group; they had a median age of 65 years (Q_1, Q_3: 56, 71), with 41.1% (86/209) being female. The participants used the SuMMiT-D system for up to 6 months (26 weeks) and had a wide range of different interactions with the SuMMiT-D system while tailored interventions were being delivered. For both studies, we had low withdrawal rates: only 4.2% and 5.3% for the pilot and feasibility studies, respectively.

Conclusions: A system was developed to successfully deliver brief messages as tailored health interventions to more than 250 people with type 2 diabetes via SMS text messages. On the basis of the low withdrawal rates and positive feedback received, it can be inferred that the SuMMiT-D system is robust, user-friendly, useful, and positive for most participants. From the two

studies, we found that online recruitment was more efficient than recruitment via postal mail; a regular SMS text reminder (eg, every 4 weeks) can potentially increase the participants' interactions with the system.

Trial Registration: ISRCTN Registry ISRCTN13404264; http://www.isrctn.com/ISRCTN13404264

(JMIR Form Res 2021;5(3):e18460) doi:10.2196/18460

KEYWORDS

type 2 diabetes; short message service; health-related behavior; mobile health; mHealth; mobile phone

Introduction

Diabetes is one of the most prevalent long-term conditions affecting the world population [1]. Currently, 4.7 million people in the United Kingdom have diabetes; more than 5 million people will have diabetes by 2025. Among people with diabetes, approximately 90% have type 2 diabetes [2]. People with type 2 diabetes are at high risk of developing serious complications, including cardiovascular disease, stroke, and chronic kidney disease, which in turn lead to an increase in the cost and resources needed for health care [1,3,4]. To avoid related comorbidities and complications, treatments with proven efficacy are needed alongside facilitating patients in the self-management of their condition [5-7].

People with diabetes commonly worry about their medicines and do not always take them regularly as prescribed [8]. This can lead to poor diabetes control. To support people with type 2 diabetes in better use of medicines, a diverse range of care programs are available. Many such services are standardized according to available evidence and based on guidelines; however, not all people equally benefit from such a *one-size-fits-all* approach [9-11]. People with poorly controlled diabetes benefit mostly from intensive, provider-driven management, whereas people with adequate glucose levels might maintain these levels independently. A more patient-centered approach is becoming a preferred strategy for improving patient outcomes, where the care is tailored according to individual patient needs and preferences [12-15].

Delivering brief messages that can address a wide range of concerns at a wide scale via digital health systems, in addition to usual care, is a promising approach to develop tailored interventions to improve treatment adherence and health outcomes [16-19]. It has been shown to be both effective and of low cost in improving health in other health conditions, including in reducing cardiovascular risk, lowering blood pressure, and stopping smoking [20-22]. In addition, as tailored interventions are more personally relevant for recipients, they would have a higher chance of being noticed, read, understood, and acted on [23].

This paper describes a SMS-based system developed for and evaluated in the Support Through Mobile Messaging and Digital Health Technology for Diabetes (SuMMiT-D) study. Through mobile health technology integrated with clinical care, this system delivers automated, tailored brief messages to support people with type 2 diabetes in effectively using medicines. The system has currently been used by more than 250 patients with type 2 diabetes across the United Kingdom and is planned to be used by another 1000 patients in the near future. The system is designed to be user-friendly: it is based on carefully developed SMS messages aimed at a large target population and accommodates users without previous experience of digital technologies. The SuMMiT-D system incorporates algorithms that identify messages to be tailored according to the user's response and characteristics (eg, medical history). This paper addresses an existing gap in the literature: the design of systems to deliver brief messages in the health care delivery system. The messages are tailored and personalized interventions to support people with type 2 diabetes. This paper presents the technical methodology used in the pilot and feasibility studies and reports on the system performance. A paper reporting the protocol for the SuMMiT-D feasibility study [24] has been published separately.

Methods

SuMMiT-D

SuMMiT-D is a program of work intended to develop and test the delivery of tailored and personalized brief messages, implemented in SMS text messages. Although implemented in a study, it is a demonstration of how this could be achieved when integrated within the National Health Service (NHS). The aim of the SuMMiT-D system is to provide enhanced and optimal support for people with type 2 diabetes. The program of work reported here consisted of two stages: a formative work package based around a pilot study to develop and test the system and a randomized controlled trial to evaluate feasibility. For the pilot study, 48 participants were recruited to evaluate and assess the system. During this phase, participants were asked to use the SuMMiT-D system for 13 weeks, with the possibility of continuing their participation for up to 26 weeks. This stage aimed to understand the feasibility of the technical aspects of the study.

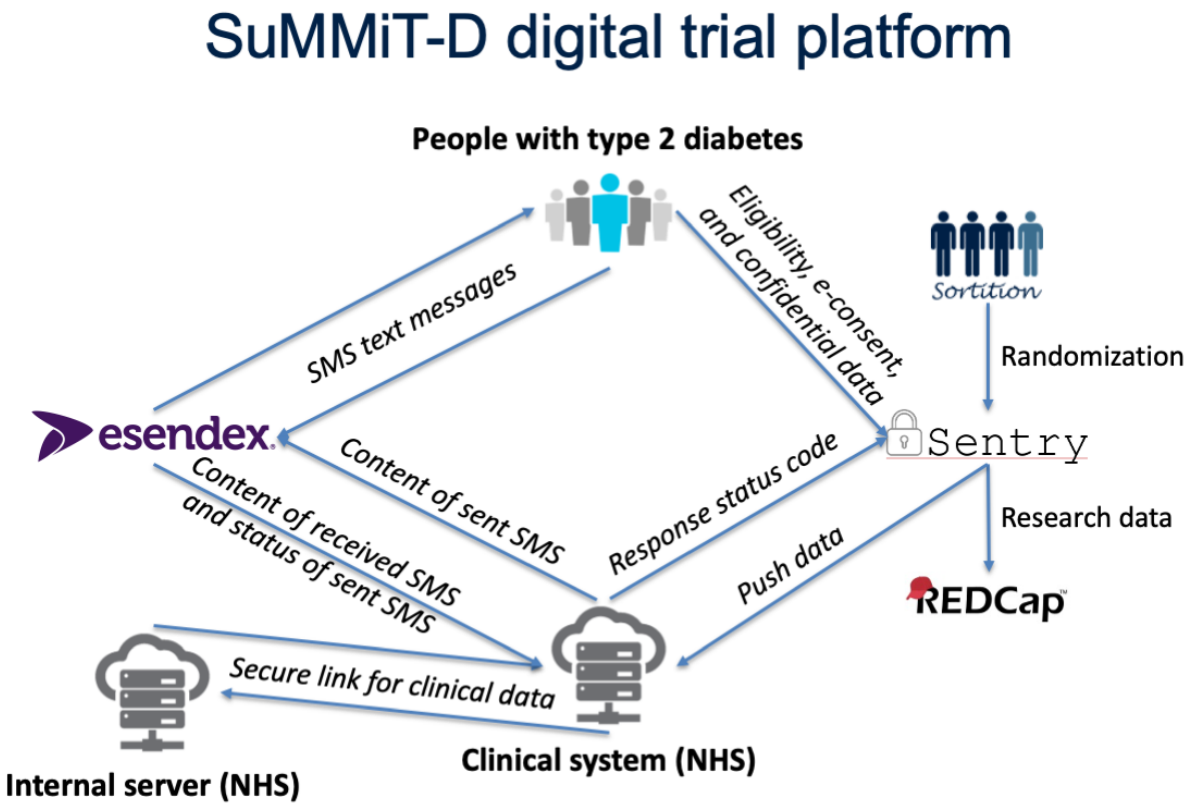
For the second phase, a feasibility study, 209 participants were recruited. This study evaluated the use of the system for a total of 6 months (26 weeks). Its purpose was to understand the participant recruitment rate and timing as well as the feasibility of integrating the developed system directly with general practitioner (GP) systems. This work was completed in October 2019. The developed methodology will be used for the main trail in 2020.

The SuMMiT-D System

Figure 1 shows the system architecture designed for the SuMMiT-D study. The design work required the coordination of the following platforms and resources: Sentry, an online-based recruitment system developed and managed by the Primary Care Clinical Trials Unit, University of Oxford;

the mobile phones used by the participants to send and receive SMS messages; a web interface accessible by the system users (ie, system administrators and research staff) to manage the clinical system; and Esendex [25], an SMS engine provider that enabled delivery and receipt of text messages to and from participants. The system was also designed to integrate alerts and personalization of messages based on primary care electronic health record (EHR) data. Platforms and protocols were required to meet the standards of data and security management [26]. Operating procedures were designed to achieve secure and independent management of the system as well as the storage of data.

Figure 1. Overview of systems developed for the Support through Mobile Messaging and Digital Health Technology for Diabetes study. Sentry: Secure entry; SuMMiT-D: Support Through Mobile Messaging and Digital Health Technology for Diabetes.



Clinical System

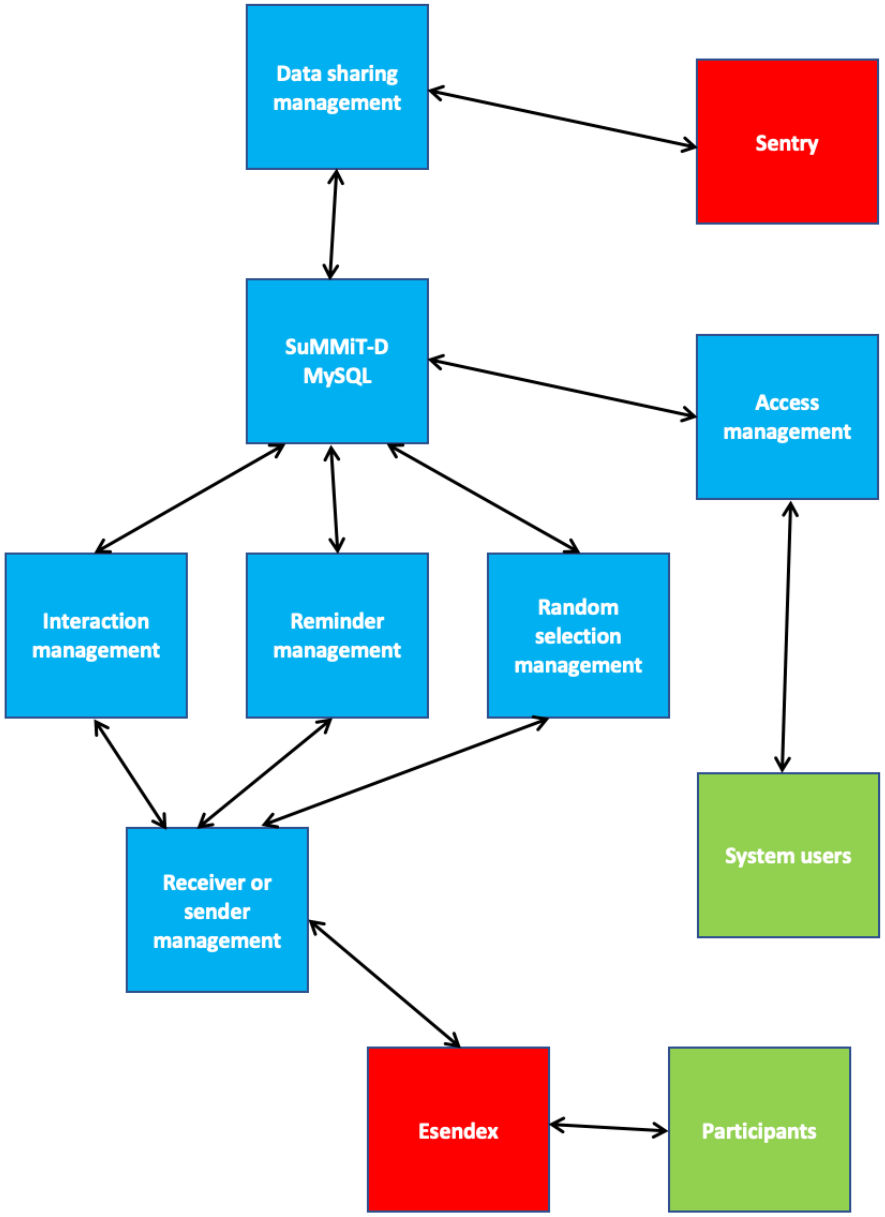
The back end (ie, the data access layer that provides access to data stored in a database) of the clinical system was developed using the CakePHP web framework [27], and the associated database tables were developed using the relational database management system MySQL [28], whereas the front end (ie, the presentation layer that formats and delivers information for further processing or display) was developed using three core technologies of the WWW [29]: HTML [30], Cascading Style Sheets [31], and Javascript [32]. These open-source technologies were chosen because of ease of use, fast iteration, and adherence to communication standards. The core component of the clinical system was implemented using CakePHP. The CakePHP web framework [27] has several advantages: it is an open-source platform that supports the PHP programming language [33]; it requires almost no preconfigurations, as most of the settings and options are auto-detected; it implements a model-view-controller pattern, which divides the developing user interfaces into 3 interconnected elements: model, view, and controller; it offers a wide range of built-in plug-ins (ie, software components that add specific features to an existing computer program) to enable customization; and it supports fast connectivity with database systems. All these characteristics make it particularly suitable for the reliable and extendible implementation of web-based services.

To meet both the trial and the NHS security and privacy requirements and to allow clinical implementation once the whole trial ends, the clinical system has been deployed in a server managed by the Oxford University Hospitals NHS Foundation Trust. The server is located behind the firewall managed by the local information management and technology team at the Oxford University Hospitals NHS Foundation Trust and hosted within the NHS Health and Social Care Network [34]. Access to this system is password protected and available only to registered users. All data (ie, database tables and participants' GP records) are stored separately on another internal server, which is inaccessible directly from the internet and communicates with the main front-end server via a

password-protected, Transport Layer Security (TLS) [35] encrypted channel.

Figure 2 shows the architecture of the clinical system with various key components that have been developed. The various components of the system include the following: *Access Management*, *Data Sharing Management*, *Receiver and Sender Management*, *Random Message Selection Management*, *Reminder Management*, *Interaction Management*.

Figure 2. Support through Mobile Messaging and Digital Health Technology for Diabetes architecture. Sentry: Secure entry; SuMMiT-D: Support Through Mobile Messaging and Digital Health Technology for Diabetes.



Participants can interact with the SuMMiT-D system by sending SMS messages from their mobile phones to a virtual mobile number (provided by Esendex) using one of the predefined commands. Esendex provides two ways of interacting with their SMS engines: (1) using the toll-free number *0800* or (2) using a virtual mobile number. For using the toll-free number *0800*, each command needs to start with a specified keyword (eg, *DIA*) for Esendex to identify the system with which the user wants to interact. In using a virtual mobile number, a virtual number is exclusively associated with one system, so that the routing keyword is not required. In this instance, the user bears the cost of sending SMS messages. Although the latter method represents a cost to the user, it has the following advantages: the cost of sending a text message is acceptable to participants; text messages from participants will always reach the SuMMiT-D system; and if the former method is used and the specified keyword has been forgotten or mistyped, the delivery of messages to the SuMMiT-D system will fail. System users (ie, health care professionals, research staff, and system administrators) engage with the clinical system using an internet-enabled computing device (eg, a smartphone, a tablet, or a computer).

Interaction Management

The *Interaction Management* module determines how the implemented algorithms in the SuMMiT-D system tailor messages according to each individual's preferences. Each time a participant sends a text via Esendex, the system (1) retrieves the original text from Esendex via the receiver or sender management; (2) parses and interprets the text referring to preferences; and (3) generates a response to be pushed back, for Esendex to finally send it to the corresponding participant.

The study participants can send predefined commands to the SuMMiT-D system. A command in the context of text messages refers to one or several words corresponding to one or more actions on the SuMMiT-D receiver end. In the SuMMiT-D system, the commands correspond to a list of settings that participants can change according to their preferences, for example, when to receive messages in a day and like or dislike the last received message to express interest in more or fewer messages of the same kind in the future. Other possible interactions include initial registration with the system, pausing messages for X weeks, restarting from being paused, completely stopping receiving messages, reviewing, and rating received messages, and requiring help instructions on how to interact with the system. The full list of commands is listed in Table 1.

Table 1. List of commands available in the Support through Mobile Messaging and Digital Health Technology for Diabetes system and the corresponding descriptions.

Study and commands	Description
Common to both studies	
Register *names*	Register with the SuMMiT-D system
Start	Start receiving messages again
Stop	Stop receiving messages completely (Need to contact the SuMMiT-D study team to restart the messages)
AM	Receive messages in AM (from 9 AM to 12 PM)
PM	Receive messages in PM (from 12 PM to 6 PM)
Help, Help 1, Help 2	Help messages for using the list of commands
Pilot	
Less or More	Receive fewer or more messages similar to the last received one
Like or Dislike	Like or Dislike the last received message
Easy or Hard	The last received message is easy or hard to understand
Useful or Not Useful	The last received message is useful or not useful
Help 3	Help messages for using the list of commands
Feasibility	
Like or Dislike	Receive fewer or more messages similar to the last received one

Reminder Management

Besides instantaneous interactions, a participant also received scheduled text messages during the study. The *reminder management* module schedules text reminders. During the pilot study, to encourage the participants to review and rate the randomly selected messages sent to them, a prompt text message was sent to each participant every week (ie, a reminder to use the commands *Like or Dislike*, *Easy or Hard*, and *Useful or Not Useful*). For the pilot study, after 4 weeks of participation, a reminder was sent on how to use the commands *More or Less*; whereas for the feasibility study, 4 weeks after randomization, a reminder was sent on how to use the commands *Like or Dislike*. In addition, at the beginning and in the middle of both studies, a participant would receive a text reminder on how to use the *Help* commands. Other types of engagement messages were a happy birthday text message and a reminder on the day of world diabetes day celebration (14th November).

Random Selection Management

The SuMMiT-D study participants received a number of brief messages sent periodically. Two types of brief SMS messages were used: one type is based on behavior change techniques (BCTs), and the other type is based on general lifestyle advice [36-38].

During the pilot study, in which all participants belonged to one group, these messages were sent three times a week. For every five continuous BCT-based messages, one general lifestyle message was sent.

During the feasibility study, participants were randomized to either the control or intervention groups. For those in the control group, only control messages were sent at a frequency of approximately once a month. The control messages here refer to general messages that are irrelevant to medicine management. For those in the intervention group, two types of messages (BCT and general lifestyle) were sent at a frequency of 3 messages per week (1 general lifestyle message, for every 2 continuous BCT-based messages). The increased frequency of lifestyle messages was based on the feedback of the participants in the pilot study. Examples of SMS text messages used in the SuMMiT-D feasibility study are shown in Table 2.

Table 2. Examples of SMS text messages used in the Support through Mobile Messaging and Digital Health Technology for Diabetes feasibility study.

Message category	Example messages
Behavior change technique	
1.4 Action Planning	Plan when, where and how you are going to take your medication
2.3 Selfmonitoring	Find a way to split your tablets into days so you notice when you have forgotten to take your tablets.
7.1 Prompts or cues	It can be difficult to remember to take your tablets. Why not set an alarm to remind you to take them?
G Health care system related concerns	Lots of questions? Check who the best person to see might be.
General lifestyle	
Signposting	Want to hear what other people with diabetes think? You can hear different people discuss their experiences at healthtalk online.

For the pilot study, there were a total of 157 unique BCT messages used, and they were from 30 different BCT groups, whereas there were 35 unique general lifestyle messages. On the basis of the feedback (ie, the message ratings) of the participants in the pilot study, the messages used in the feasibility study have been edited accordingly: there were a total of 170 unique BCT messages used, and they were from 30 different BCT groups, whereas there were 35 unique general lifestyle messages and 6 unique control messages. More detailed information on the messages used in the pilot and feasibility studies can be found in the studies by Farmer et al [24] and Bartlett et al [39].

The algorithm below describes the methodology for sending these SMS messages:

Suppose there are t different groups of BCTs as: $\{B_1, B_2, ..., B_t\}$, with the i-th ($1 \le i \le t$) BCT group having n_i different SMS texts, that is, . This results in n unique BCT SMS texts, where . There are also m different general lifestyle SMS texts . As a result, there are $n+m$ unique SMS messages in the system, which can be represented as $\{B_1, B_2, ..., B_t, G\}$. The SMS messages sent to each participant follow this rule: for each week (ie, 7 days), 3 messages are sent, and they are sent with an even time gap; in addition, after continuously sending q BCT-based SMS messages, 1 general lifestyle-based message should be sent (q=5 for the pilot study and q=2 for the feasibility study according to the participants' feedback from the pilot study).

The SMS messages sent to each participant were randomly selected as follows: for each participant, all t different BCT groups of messages have the same probability. To ensure that a different message was received each time, the system excluded all the previously sent messages and the most recently used BCT group and selected a message from the remaining messages and BCT groups. The same approach was used when sending general lifestyle-based SMS messages. The above process is enabled only for those participants who were scheduled to receive a message and only within the time slot selected by each participant.

In addition to the above process, the SuMMiT-D system tailors and personalizes the messages for each participant. Tailoring and personalization were performed according to the text responses from each participant. In the first instance, upon receiving *Pause X*, *Start*, and *STOP* from a participant, messages could be paused for X weeks, restarted, or completely stopped, respectively.

Upon receiving *More* and *Like* from a participant in the pilot and feasibility studies, respectively, after she or he receives a BCT-based message (ie, , where), another message from the same BCT group (ie, , where) is scheduled to be sent to this participant after 2 weeks. In addition, all the messages in the same BCT group (ie,) would have doubled the chance (ie, increase 100% probability) to be selected and sent to this participant in the future. Upon receiving *Dislike* from a participant in the feasibility study, after she or he receives a BCT-based message (ie, , where ,), all the messages in the same BCT group (ie,) would have halved the chance (ie, decrease 50% probability) to be selected and sent to this participant in the future. For details, please refer to the pseudocode listed in Multimedia Appendix 1.

Upon receiving *AM* and *PM* from a participant, the message delivery time slot to this participant was changed accordingly.

Receiver or Sender Management

The purpose of this component is to send and receive messages to and from participants. This management component employs a third-party mobile messaging system (Esendex [25]) to act as a bridge between the participants and the clinical system. To dispatch SMS messages to the participants' mobile phones, they are sent to the Esendex system's application program interface (API) by the clinical system. The information exchanged between the 2 systems includes the target mobile number, the message context to be sent, a unique message ID, and a key code to authorize this message to be sent. This communication is established over an encrypted HTTP channel over TLS [35]. After 5 hours, the clinical system again connects to Esendex's API and then checks the delivery status of the message using the unique message ID, which is returned by Esendex during the first connection. This process allows tracking of the delivery status of every sent message. Each message delivery status (ie, *Delivered*, *Sent*, *Failed*, *Expired*, *Cancelled*, and *Partially Delivered*) is stored by the clinical system, along with the corresponding message information.

For messages whose delivery fails (ie, a delivery status of *Failed*, *Expired*, *Cancelled*, and *Partially Delivered*), another attempt is scheduled immediately. After further 2 failed redelivery attempts, an automatic email is sent to the system users, whose role is to be either administrators or researchers, to notify this failed delivery and suggest the possibility that the participant may have changed this mobile number.

The clinical system checks Esendex's API at an interval of every 30 seconds. This very short period was selected to provide a sense of timeliness of the response to (potential) participants and to maintain a high engagement level.

Data Sharing Management

The eligible candidates were screened and then invited to the SuMMiT-D study (full details of the recruitment process have been reported elsewhere [24]). To participate in any SuMMiT-D study, patients had to express their interest with a text message (ie, Register *names*) to one of the virtual numbers provided by Esendex. Once the text was received, an automatic reply was sent back. This handshake process guaranteed that the mobile number was valid. For the pilot study, the recruitment process was completed over the phone, the paper consent form was sent and returned via post, and personal information was directly entered into the SuMMiT-D system via a web interface. For the feasibility study, to add another level of security and automation, the recruitment process was performed electronically via a system called Sentry (Secure entry) [40]. The consent form could then be completed electronically. Once the recruitment process for a participant was completed, all personal details were securely stored in REDCap [41]. To enable data sharing of personal details required to be shared with the clinical system, data were pushed from the Sentry system to the clinical system via HTTPs [42]. During this data transaction, the mobile number that the participant used to express interest was used as the matching key. The following personal details were shared by Sentry and stored in the clinical system: *first name, last name, mobile number, GP practice name, date of birth, preferred name, gender, NHS number, whether using a smartphone*, and *participant study ID*. As Sentry also implemented the randomization algorithm for determining the control or intervention group, it shared also the randomization result so that the SuMMiT-D system would know whether a participant data that had been pushed from Sentry to the clinical system belonged to the control or intervention group.

Access Management

System users can log in to the SuMMiT-D system using their personal usernames and passwords. Users are categorized according to 3 levels of privileges that determine their access rights to the information and functions available. The system administrators (*Admin*) get access to the entire system and all its functionalities; the study researchers (*Researcher*) get full access to the information of participants and GP practices; and the research nurses, trail manager, and coordinator (*User*) get only a read-only view of the information of participants and GP practices.

The web pages that a system user can access depend on the privilege level, which is determined by the log-in credentials. Table 3 provides the main views that a user can access according to their privilege level and a brief description for each of those views.

Table 3. Privilege levels and the corresponding functionality available at each of these privilege levels.

Functionality	Privilege level		
	Admin	Researcher	User
• Add system users (admin, researcher, user) • List of system users • Detailed view or edit of system users	✓[a]	—[b]	—
• Add participants • Add GP[c] practices • Detailed edit of participants • Detailed edit of GP practices • List of sent or received SMS • Detailed view of sent or received SMS	✓	✓	—
• List of participants • Search participants • List of GP practices • Detailed view of participants • Detailed view of GP practices • Detailed view or edit of own details	✓	✓	✓

[a]Function is available at this level.

[b]Function not available at this level.

[c]GP: general practitioner.

Development Approach

The SuMMiT-D system was developed in collaboration with health care professionals, the study team, the software development team, and a group of patient representatives. Both test GP practices and patients were set up during the early development process, so that the system could be evaluated for its reliability and to allow software bugs to be identified and fixed.

Defined Metrics

To evaluate the performance of the SuMMiT-D system, the following 3 metrics were defined: *response rate, keyword percentage, and rate versus prompt*.

Response Rate

To determine whether a participant was active in the study, the response rate of SMS messages for each participant was calculated according to:

(**1**)

where, for the pilot study,

(**2**)

whereas, for the feasibility study,

(**3**)

Keyword Percentage

In the pilot study, to evaluate the usability of a keyword-based system, a fictitious keyword was selected. Participants were asked to submit all commands after using the DIA keyword. The keyword (DIA) percentage is calculated according to:

(**4**)

Rate Versus Prompt

In the pilot study, to review whether the prompt messages could improve the total number of responses (ie, rating) received, the ratio of number of response messages to the number of prompt messages is calculated as:

(**5**)

Where rating is calculated according to equation 2.

Results

Pilot Study

For the pilot study, 48 participants were recruited, with a median age of 64 years (first quartile and third quartile [Q_1, Q_3: 54.5, 69]). Among the 48 participants, 19 (40%) were from the Greater Manchester region and 29 (60%) were from the Thames Valley region.

Duration of Study and Admission Process

The first participant expressed interest on April 24, 2018, and the last participant expressed interest on August 24, 2018. The first participant started to use the SuMMiT-D system on May 8, 2018, whereas the last participant started to use the system on September 17, 2018. The first participant stopped using the SuMMiT-D system (ie, either finished or withdrew from the study) on August 11, 2018, whereas the last one stopped on January 31, 2019.

An admission process duration refers to how long (in days) it took the participants from expressing interest (ie, sending the *Register* command to register with the system) to fully start the SuMMiT-D study (ie, to begin receiving regularly sent messages). The admission process duration here has a median value of 21.07 days (Q_1, Q_3: 17.31, 24.93).

Two participants withdrew from the study at days 23 and 67 (withdrawal rate of 4%). At 3 months (13 weeks) after using the system, 19% (9/48) participants chose to stop, whereas the remaining 77% (37/48) decided to continue.

In addition, a total of 11 *STOP* commands have been received, which are in line with the results above: 2 participants withdrew and 9 stopped at 3 months after using the system. Only 2 Pause *X* commands were received: one chose to pause the messages for 2 weeks and the other paused for 4 weeks.

SMS Message Response Rate of Participants

The metric here is defined as the *response rate*, which is calculated according to equations 1 and 2. The median *response rate* was 71% (Q_1, Q_3: 53%, 111%). This actual response rate is lower than what might have been expected, as there was a weekly reminder (ie, the prompt message encouraging participants to rate received messages).

Participants Including the Keyword DIA in Their Responses

Both toll-free and virtual numbers are available for SMS-based studies to interact with participants. The first approach generally shares the same toll-free number across different systems and, thus, requires the use of a keyword to identify the SMS texts received from users. However, the additional keyword may represent a challenge for some participants, as they may forget to add the keyword at the start of a message, and as a result, such a message will be missed at the corresponding system end completely.

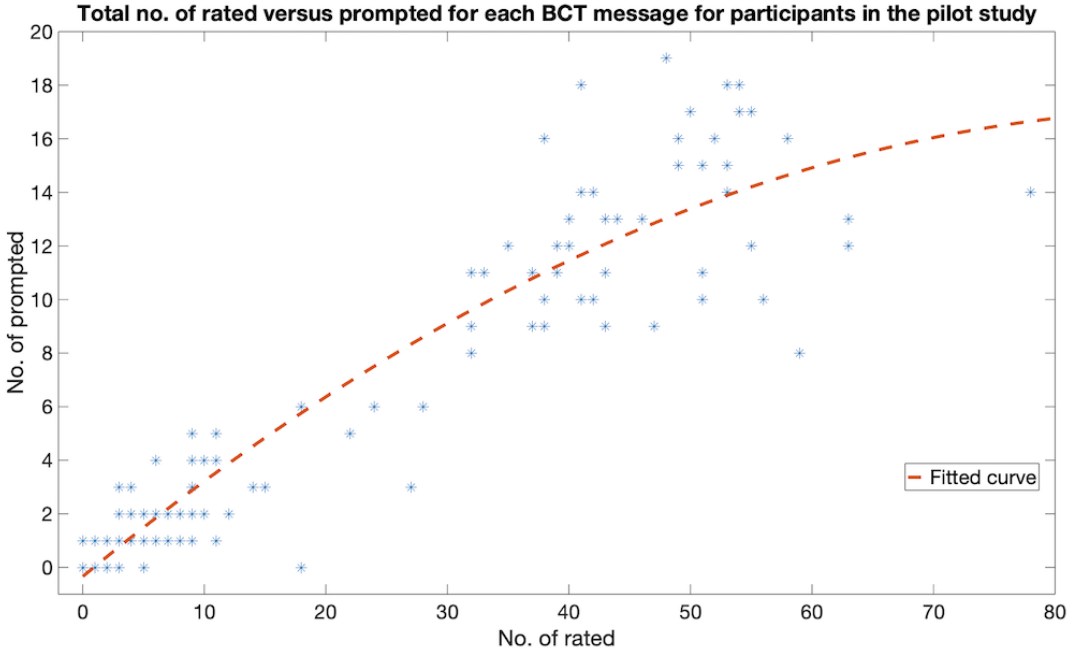
The metric here is defined as the *keyword percentage*, which is calculated according to equation 4. The median *keyword percentage* was 93% (Q_1, Q_3: 87%, 98%). The histogram for the keyword percentage is shown in Multimedia Appendix 2. Although the keyword (*DIA*) was used quite frequently, some messages would have still not been received on the SuMMiT-D system side.

Prompt Messages Versus Message Responses

A weekly reminder was sent to the participants in the pilot study in the form of a prompt message asking to rate the received messages. To review whether this could improve the total number of responses (ie, rating) received, the metric *rate versus prompt* is calculated as described above equation 5. The r*ate versus prompt* has a median value of 280% (Q_1, Q_3: 215%, 434%). The rate versus prompt is also lower than expected, which is in line with the results shown above.

In addition, for each BCT message, the total number rated against the total number prompted is summarized and shown in Figure 3, with a quadratic polynomial curve fitted to the data points. From the fitted curve, prompt messages can potentially improve the total number of responses (ie, rating) received.

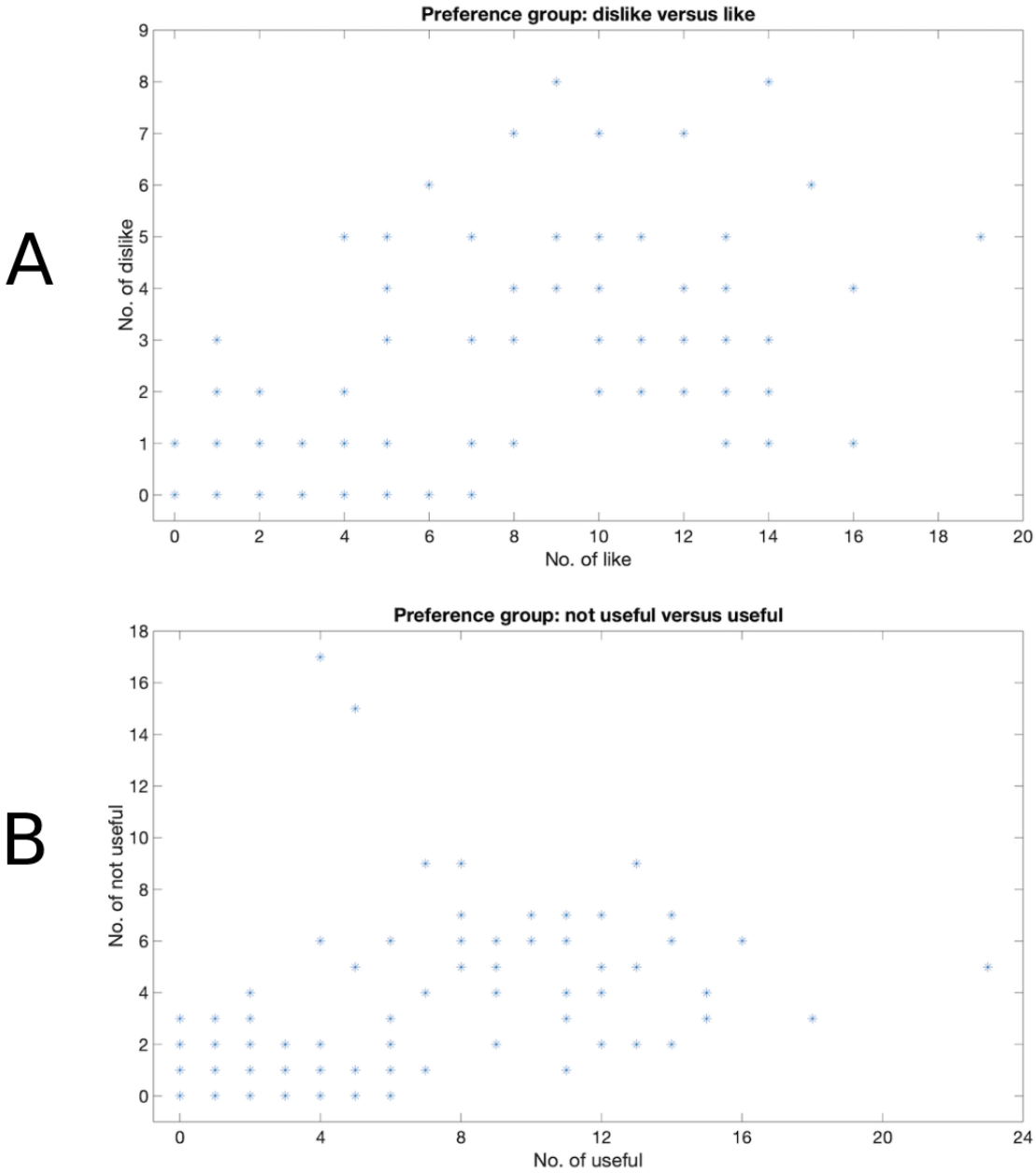
Figure 3. The total number of rated versus prompted for each behavioral change therapy message for participants in the pilot study.



Comparison Among the Preference Groups of Messages Responses

To compare which preference groups (ie, Like or Dislike, Easy or Hard, Useful or Not Useful, and More or Less) are mostly used by participants in the pilot study, the total numbers of Like or Dislike, Easy or Hard, Useful or Not Useful, and Less or More were counted. The results are shown in Figure 4. From the results, it can be concluded that the vast majority of the BCT messages sent in the pilot study received a positive rating (ie, Like and Useful).

Figure 4. Different preference groups among the SMS message response rating for participants in the pilot study. A: Preference group: Dislike vs Like; B: Preference group: Not useful vs Useful.



Feasibility Study

For the feasibility study, a total of 209 participants (with a median age of 65 years, Q_1, Q_3: 56, 71; with 86/209, 41.2% female) were recruited and randomized. Of these, 106 were randomized into the control group, whereas the remaining 103 were randomized into the intervention group. Of the 209 participants, 52 (24.8%) were from the Greater Manchester area and 157 (75.1%) were from the Thames Valley area.

Duration of Study and Admission Process

The first participant expressed interest on December 3, 2018, and the last one expressed interest on March 20, 2019. The first participant started to use the SuMMiT-D system on December 10, 2018, whereas the last participant started to use the system on April 16, 2019. The first participant stopped using the system (ie, either finished or withdrew from the study) on March 4, 2019, whereas the last one stopped using the system on October 16, 2019.

The admission process duration here has a median value of 8.76 days (Q_1, Q_3: 3.94, 15.07).

A total of 11 participants withdrew from the study (10 from the intervention group): 8 of them stopped the messages via sending the *STOP* commands, whereas the remaining 3 called the study team to ask for withdrawal. A total of 7 *Pause X* commands were received, and they were from 4 different participants.

SMS Message Response Rate of Participants

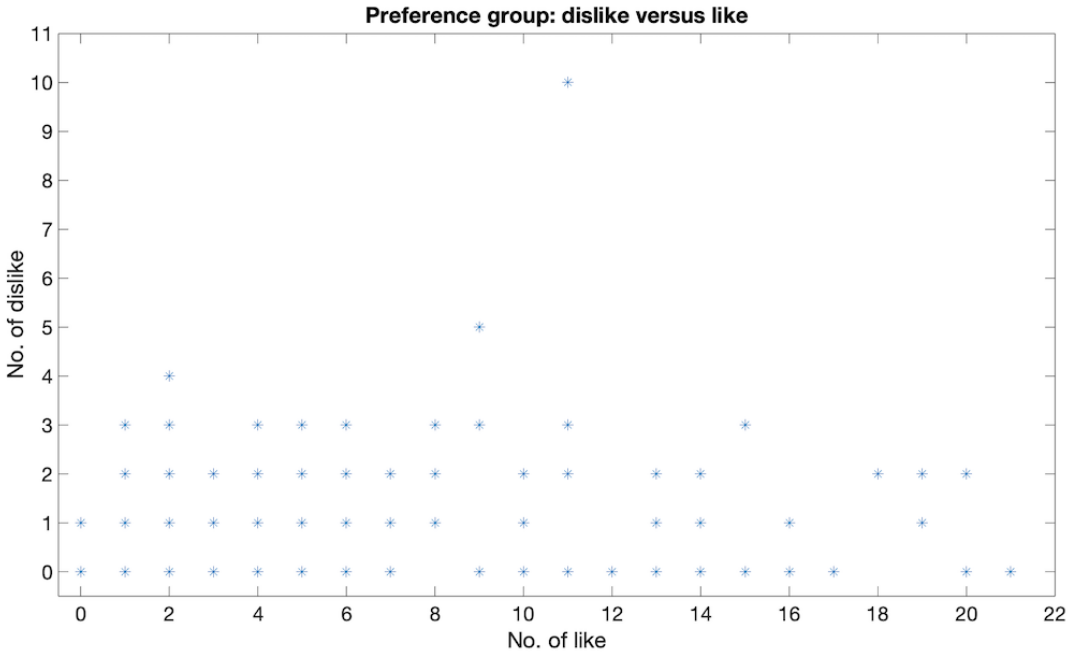
The metric here is the *response rate*, which is calculated according to equations 1 and 3. The median *response rate* was 17% (Q_1, Q_3: 13%, 22%). It can be concluded that there is a

much lower response rate for the participants in the feasibility study, as there was no weekly prompt message sent in the latter study.

Like Versus Dislike Among Message Responses

To compare which BCT message is mostly liked or disliked by participants in the feasibility study, the total number of Like or Dislike was counted for all the BCT messages, and the results are shown in Figure 5. In the same way as Figure 4, Figure 5 shows that the great majority of the BCT messages sent in the feasibility study receive a positive rating (ie, Like).

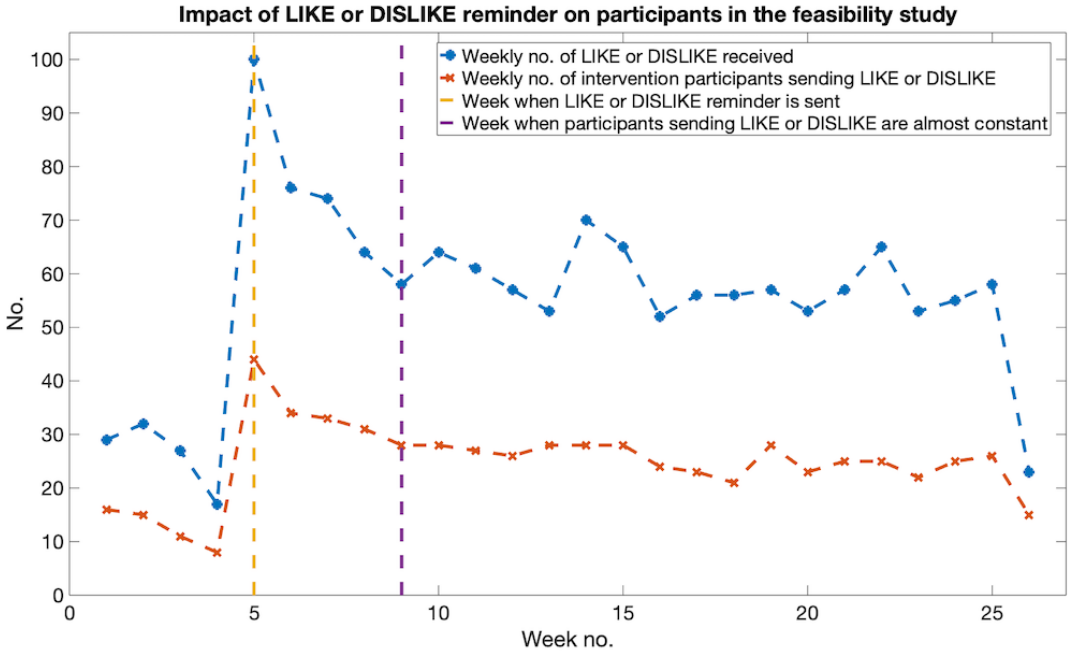
Figure 5. Like versus Dislike among the SMS message response ratings for participants in the feasibility study.



Impact of Like or Dislike Reminder

Unlike the pilot study, there were no weekly prompt SMS messages (ie, encouraging participants to rate all the received SMS messages) sent in the feasibility study. To remind the participants in the intervention group to use the Like or Dislike commands, a reminder of how to use these commands was sent to them when they had been in the study for 4 weeks. To determine the impact of Like or Dislike reminders (ie, how long the reminders will have an effect on the participants), the following 2 numbers have been plotted: the total number of Like or Dislike commands received per week and the number of participants sending Like or Dislike commands per week. The results are shown in Figure 6. From the results, it can be seen that in week 5 (ie, just after the reminders were sent), these 2 numbers reached their peak values. From week 9 (ie, 4 weeks after the reminders were sent), the number of participants sending Like or Dislike commands back per week remained almost constant. This suggests that the impact of Like or Dislike reminders persists for approximately 4 weeks.

Figure 6. Impact of Like/Dislike reminder on participants in the feasibility study.



Discussion

In these pilot and feasibility studies for the SuMMiT-D brief messaging, mobile phone–based intervention, participant withdrawal rates were 4.2% and 5.3%, respectively. In addition, from the message responses, it was concluded that most of the participants gave positive feedback to the BCT messages they received (ie, useful and like). Furthermore, when offered the opportunity to continue for another 3 months in the pilot study, 37 out of 46 participants chose to stay. These promising results suggest that the SuMMiT-D system is robust, user-friendly, useful, and positive for most participants.

Comparing the length of the admission process (registration and sign up), it can be seen that there is a significant improvement in the feasibility study compared with the pilot. This may be explained by changing the recruitment method from regular mail (ie, manually sending and receiving all forms: registration, questionnaires, and consent) to online recruitment.

One common issue in both studies is that the response rate of participants to the messages sent to them was lower than expected. This might have been associated with the demographics of those recruited to the studies (with a median age of about 65 years) where participants might not have been familiar with the use of mobile phones. In the pilot study, a weekly prompt message was sent to participants to encourage them to rate received messages. In the feasibility study, a one-time Like or Dislike reminder describing how to use the Like and Dislike commands was sent to participants. Our findings suggest that (1) prompt messages can potentially improve the total number of responses received and (2) a reminder encouraging participants to Like or Dislike messages has an effect persisting for about 4 weeks before a participant returns to their normal response rate. Potentially higher response rates may be obtained from participants if they were sent the reminder every 4 weeks.

The SuMMiT-D system uses brief messages as interventions to provide education and support self-management for people with type 2 diabetes. The brief messages are delivered through mobile phone–based SMS text messages. These patients are often concerned about starting new diabetes medicines or face difficulties in taking their medicines regularly. To enhance the support for patients' self-management, these brief messages were personalized and tailored for each participant individually, according to each patient's own preference and their feedback to previously received SMS text messages.

The SuMMiT-D system offers a model for technology-based self-management support [30,31], as it shows the clinical efficacy and cost-effectiveness of a text messaging intervention, compared with the usual care, for people with type 2 diabetes. In addition, the SuMMiT-D system is not only an effective tool for type 2 diabetes, but it could also be extended to other long-term conditions (eg, hypertension).

All the information and feedback that we gathered in each of the studies was carried out iteratively to improve the functionalities of the system. From the pilot study, we learned the following: (1) a recruitment process via regular post was time consuming and imposed a high workload from the study team, and in the feasibility study, we introduced the Sentry system to allow a digital (online) recruitment process, and (2) according to the participants' feedback, they would have liked to receive the general lifestyle messages more frequently, and we increased the ratio of the general lifestyle and BCT messages from 1 to 5 in the pilot study to 1 to 2 in the feasibility studies.

In the SuMMiT-D pilot and feasibility studies, we have demonstrated the successful delivery of interventions to more than 200 participants via SMS text messages. In the main SuMMiT-D study, to further enhance the support for patients'

self-management, their EHR data will be routinely shared and collected via a GP data provider. EHR data, such as medication possession ratio and glycated hemoglobin, will be used to tailor the intervention messages. In addition, more automation in the recruitment process will be introduced, such that the overall workload is reduced.

Acknowledgments

This project was funded by the National Institute for Health Research (NIHR; Programme Grants for Applied Research [project reference RP-PG-1214-20003]). LT and AF are supported by the NIHR Oxford Biomedical Research Centre. AF is an NIHR Senior Investigator. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Conflicts of Interest

ER supports a team receiving funds from the NIHR Health Technology Assessment (HTA) Programme.

Multimedia Appendix 1
Algorithm 1 pseudocode for receiving “More”/“Like” or “Dislike” after sending a behavior change technique message to a participant.
[DOCX File , 13 KB - formative_v5i3e18460_app1.docx]

Multimedia Appendix 2
Histogram plot of the keyword percentage keyword percentage among the SMS SMS message responses for participants in the pilot study.
[PNG File , 70 KB - formative_v5i3e18460_app2.png]

References

1. Tamayo T, Rosenbauer J, Wild S, Spijkerman A, Baan C, Forouhi N, et al. Diabetes in Europe: an update. Diabetes Res Clin Pract 2014 Feb;103(2):206-217. [doi: 10.1016/j.diabres.2013.11.007] [Medline: 24300019]
2. Us, diabetes and a lot of facts and stats. Diabetes UK. 2019. URL: https://www.diabetes.org.uk/resources-s3/2019-02/1362B_Facts%20and%20stats%20Update%20Jan%202019_LOW%20RES_EXTERNAL.pdf [accessed 2020-07-13]
3. U.K. Prospective Diabetes Study Group. U.K. prospective diabetes study 16. Overview of 6 years' therapy of type II diabetes: a progressive disease. U.K. Prospective Diabetes Study Group. Diabetes 1995 Nov 1;44(11):1249-1258. [doi: 10.2337/diab.44.11.1249] [PMC: 7589820]
4. Asche C, LaFleur J, Conner C. A review of diabetes treatment adherence and the association with clinical and economic outcomes. Clin Ther 2011 Jan;33(1):74-109. [doi: 10.1016/j.clinthera.2011.01.019] [Medline: 21397776]
5. UK Prospective Diabetes Study Group. Efficacy of atenolol and captopril in reducing risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 39. Br Med J 1998 Sep 12;317(7160):713-720. [doi: 10.1136/bmj.317.7160.713]
6. Norris SL, Lau J, Smith SJ, Schmid CH, Engelgau MM. Self-management education for adults with type 2 diabetes: a meta-analysis of the effect on glycemic control. Diabetes Care 2002 Jul 1;25(7):1159-1171. [doi: 10.2337/diacare.25.7.1159] [Medline: 12087014]
7. Colhoun HM, Betteridge DJ, Durrington PN, Hitman GA, W Neil HA, Livingstone SJ, et al. Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): multicentre randomised placebo-controlled trial. Lancet 2004 Aug;364(9435):685-696. [doi: 10.1016/s0140-6736(04)16895-5]
8. Salas M, Hughes D, Zuluaga A, Vardeva K, Lebmeier M. Costs of medication nonadherence in patients with diabetes mellitus: a systematic review and critical analysis of the literature. Value Health 2009 Sep;12(6):915-922 [FREE Full text] [doi: 10.1111/j.1524-4733.2009.00539.x] [Medline: 19402848]
9. Rothe U, Müller G, Schwarz PE, Seifert M, Kunath H, Koch R, et al. Evaluation of a diabetes management system based on practice guidelines, integrated care, and continuous quality management in a Federal State of Germany: a population-based approach to health care research. Diabetes Care 2008 May 10;31(5):863-868. [doi: 10.2337/dc07-0858] [Medline: 18332161]
10. Pimouguet C, Le Goff M, Thiébaut R, Dartigues JF, Helmer C. Effectiveness of disease-management programs for improving diabetes care: a meta-analysis. Cand Med Assoc J 2011 Feb 8;183(2):E115-E127 [FREE Full text] [doi: 10.1503/cmaj.091786] [Medline: 21149524]
11. Arianne E, Inge DP, Cor S, Marieke S, Hubertus V. Toward tailored disease management for type 2 diabetes. Am J Managed Care 2012 Oct;18(10):619-630.
12. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, American Diabetes Association (ADA), European Association for the Study of Diabetes (EASD). Management of hyperglycemia in type 2 diabetes: a patient-centered approach: position statement of the American Diabetes Association (ADA) and the European Association for the Study of

Diabetes (EASD). Diabetes Care 2012 Jun 19;35(6):1364-1379 [FREE Full text] [doi: 10.2337/dc12-0413] [Medline: 22517736]
13. Relman AS. The Institute of Medicine Report on The Quality of Health Care: Crossing the Quality Chasm: A New Health System for the 21st Century. N Engl J Med 2001 Aug 30;345(9):702-703. [doi: 10.1056/nejm200108303450917]
14. The American Geriatrics Society Expert Panel on Person-Centered Care. Person-centered care: a definition and essential elements. J Am Geriatr Soc 2016 Jan 2;64(1):15-18. [doi: 10.1111/jgs.13866] [Medline: 26626262]
15. Coulourides Kogan A, Wilber K, Mosqueda L. Moving toward implementation of person-centered care for older adults in community-based medical and social service settings: 'you only get things done when working in concert with clients'. J Am Geriatr Soc 2016 Jan 02;64(1):e8-14. [doi: 10.1111/jgs.13876] [Medline: 26626544]
16. Robby N, Nancy W, Tamara N, Nicholas H, Rebecca J, Arun K, et al. Interventions for enhancing medication adherence. Cochrane Database of Systematic Reviews 2014 Nov 20;2014(11). [doi: 10.1002/14651858.cd000011.pub4]
17. Ricci-Cabello I, Bobrow K, Islam SM, Chow CK, Maddison R, Whittaker R, et al. Examining development processes for text messaging interventions to prevent cardiovascular disease: systematic literature review. JMIR Mhealth Uhealth 2019 Mar 29;7(3):e12191 [FREE Full text] [doi: 10.2196/12191] [Medline: 30924790]
18. Farmer AJ, McSharry J, Rowbotham S, McGowan L, Ricci-Cabello I, French DP. Effects of interventions promoting monitoring of medication use and brief messaging on medication adherence for people with Type 2 diabetes: a systematic review of randomized trials. Diabet Med 2016 May 17;33(5):565-579. [doi: 10.1111/dme.12987] [Medline: 26470750]
19. Beratarrechea A, Lee AG, Willner JM, Jahangir E, Ciapponi A, Rubinstein A. The impact of mobile health interventions on chronic disease outcomes in developing countries: a systematic review. Telemed J E Health 2014 Jan;20(1):75-82 [FREE Full text] [doi: 10.1089/tmj.2012.0328] [Medline: 24205809]
20. 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: a randomized clinical trial. J Am Med Assoc 2015 Sep 22;314(12):1255-1263. [doi: 10.1001/jama.2015.10945] [Medline: 26393848]
21. Bobrow K, Farmer AJ, Springer D, Shanyinde M, Yu L, Brennan T, et al. Mobile phone text messages to support treatment adherence in adults with high blood pressure (star): a single-blind, randomized trial. Circulation 2016 Jan 14:CIRCULATIONAHA.115.017530. [doi: 10.1161/circulationaha.115.017530]
22. 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: a randomized clinical trial. J Am Med Assoc 2015 Sep 22;314(12):1255-1263. [doi: 10.1001/jama.2015.10945] [Medline: 26393848]
23. Kassavou A, Sutton S. Automated telecommunication interventions to promote adherence to cardio-metabolic medications: meta-analysis of effectiveness and meta-regression of behaviour change techniques. Health Psychol Rev 2018 Mar 12;12(1):25-42. [doi: 10.1080/17437199.2017.1365617] [Medline: 28805162]
24. Farmer A, Allen J, Bartlett K, Bower P, Chi Y, French D, SuMMiT-D Collaborative Group. Supporting people with type 2 diabetes in effective use of their medicine through mobile health technology integrated with clinical care (SuMMiT-D Feasibility): a randomised feasibility trial protocol. BMJ Open 2019 Dec 29;9(12):e033504 [FREE Full text] [doi: 10.1136/bmjopen-2019-033504] [Medline: 31888938]
25. Esendex. URL: https://www.esendex.co.uk [accessed 2020-07-13]
26. Chan A, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. Br Med J 2013 Jan 8;346(jan08 15):e7586 [FREE Full text] [doi: 10.1136/bmj.e7586] [Medline: 23303884]
27. CakePHP. URL: https://cakephp.org [accessed 2020-07-13]
28. MySQL. URL: https://www.mysql.com [accessed 2020-07-13]
29. World Wide Web. URL: https://www.w3.org [accessed 2020-07-13]
30. HTML. URL: https://www.w3.org/html [accessed 2020-07-13]
31. CSS. URL: https://www.w3.org/Style/CSS [accessed 2020-07-13]
32. Javascript. URL: https://www.w3.org/standards/webdesign/script [accessed 2020-07-13]
33. PHP. URL: https://www.php.net [accessed 2020-07-13]
34. Health and Social Care Network (HSCN). URL: https://digital.nhs.uk/services/health-and-social-care-network [accessed 2020-07-13]
35. The Transport Layer Security (TLS) Protocol. URL: https://tools.ietf.org/html/rfc524636 [accessed 2020-07-13]
36. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug 20;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
37. Lauren E C, Rachel N C, Marijn B, Alexander J R, Marie J, Michael P K, et al. Links between behavior change techniques and mechanisms of action: an expert consensus study. Annals of Behavioral Medicine 2019 Jul 17;53(8):708-720. [doi: 10.31234/osf.io/fge86]
38. Long H, Bartlett YK, Farmer AJ, French DP. Identifying brief message content for interventions delivered via mobile devices to improve medication adherence in people with type 2 diabetes mellitus: a rapid systematic review. J Med Internet Res 2019 Jan 9;21(1):e10421. [doi: 10.2196/10421]

39. Yvonne Kiera B, Andrew F, Rustam R, David P G. Use of brief messages based on behavior change techniques to encourage medication adherence in people with type 2 diabetes: developmental studies. J Med Internet Res 2020 May 13;22(5). [doi: 10.2196/preprints.15989]
40. Sentry (Secure Entry). URL: https://sentry.phc.ox.ac.uk/sentry [accessed 2020-07-13]
41. REDCap. URL: https://www.project-redcap.org [accessed 2020-07-13]
42. Secure Your Site with HTTPS. URL: https://support.google.com/webmasters/answer/6073543?hl=en [accessed 2020-07-13]

Abbreviations

API: application program interface
BCT: behavior change technique
EHR: electronic health record
GP: general practitioner
NHS: National Health Service
NIHR: National Institute for Health Research
Q_1, Q_3: first quartile and third quartile
SuMMiT-D: Support through Mobile Messaging and Digital Health Technology for Diabetes
TLS: transport layer security

Edited by G Eysenbach; submitted 27.02.20; peer-reviewed by D Wake, J Buekers; comments to author 29.06.20; revised version received 20.08.20; accepted 17.12.20; published 26.03.21.

Please cite as:
Chi Y, Velardo C, Allen J, Robinson S, Riga E, Judge D, Tarassenko L, Farmer AJ
System Architecture for "Support Through Mobile Messaging and Digital Health Technology for Diabetes" (SuMMiT-D): Design and Performance in Pilot and Randomized Controlled Feasibility Studies
JMIR Form Res 2021;5(3):e18460
URL: https://formative.jmir.org/2021/3/e18460
doi:10.2196/18460
PMID:33769299

Original Paper

The Dutch COVID-19 Contact Tracing App (the CoronaMelder): Usability Study

Britt Elise Bente[1], MSc; Jan Willem Jaap Roderick van 't Klooster[2], PhD; Maud Annemarie Schreijer[1], MSc; Lea Berkemeier[1], MSc; Joris Elmar van Gend[2], MSc; Peter Jan Hendrik Slijkhuis[2], MSc; Saskia Marion Kelders[1,3], PhD; Julia Elisabeth Wilhelmina Cornelia van Gemert-Pijnen[1], PhD

[1]Centre for eHealth and Wellbeing Research, Department of Psychology, Health and Technology, University of Twente, Enschede, Netherlands

[2]Behavioural Management and Social Sciences Lab, Faculty of Behavioral, Management and Social Sciences, University of Twente, Enschede, Netherlands

[3]Optentia Research Focus Area, North West University, Vanderbijlpark, South Africa

Corresponding Author:
Britt Elise Bente, MSc
Centre for eHealth and Wellbeing Research
Department of Psychology, Health and Technology
University of Twente
Drienerlolaan 5
Enschede, 7522 NB
Netherlands
Phone: 31 53 489 9660
Email: b.e.bente@utwente.nl

Abstract

Background: Adoption and evaluation of contact tracing tools based on information and communications technology may expand the reach and efficacy of traditional contact tracing methods in fighting COVID-19. The Dutch Ministry of Health, Welfare and Sports initiated and developed CoronaMelder, a COVID-19 contact tracing app. This app is based on a Google/Apple Exposure Notification approach and aims to combat the spread of the coronavirus among individuals by notifying those who are at increased risk of infection due to proximity to someone who later tests positive for COVID-19. The app should support traditional contact tracing by faster tracing and greater reach compared to regular contact tracing procedures.

Objective: The main goal of this study is to investigate whether the CoronaMelder is able to support traditional contact tracing employed by public health authorities. To achieve this, usability tests were conducted to answer the following question: is the CoronaMelder user-friendly, understandable, reliable and credible, and inclusive?

Methods: Participants (N=44) of different backgrounds were recruited: youth with varying educational levels, youth with an intellectual disability, migrants, adults (aged 40-64 years), and older adults (aged >65 years) via convenience sampling in the region of Twente in the Netherlands. The app was evaluated with scenario-based, think-aloud usability tests and additional interviews. Findings were recorded via voice recordings, observation notes, and the Dutch User Experience Questionnaire, and some participants wore eye trackers to measure gaze behavior.

Results: Our results showed that the app is easy to use, although problems occurred with understandability and accessibility. Older adults and youth with a lower education level did not understand why or under what circumstances they would receive notifications, why they must share their key (ie, their assigned identifier), and what happens after sharing. In particular, youth in the lower-education category did not trust or understand Bluetooth signals, or comprehend timing and follow-up activities after a risk exposure notification. Older adults had difficulties multitasking (speaking with a public health worker and simultaneously sharing the key in the app). Public health authorities appeared to be unprepared to receive support from the app during traditional contact tracing because their telephone conversation protocol lacks guidance, explanation, and empathy.

Conclusions: The study indicated that the CoronaMelder app is easy to use, but participants experienced misunderstandings about its functioning. The perceived lack of clarity led to misconceptions about the app, mostly regarding its usefulness and privacy-preserving mechanisms. Tailored and targeted communication through, for example, public campaigns or social media, is necessary to provide correct information about the app to residents in the Netherlands. Additionally, the app should be presented

as part of the national coronavirus measures instead of as a stand-alone app offered to the public. Public health workers should be trained to effectively and empathetically instruct users on how to use the CoronaMelder app.

(JMIR Form Res 2021;5(3):e27882) doi:10.2196/27882

KEYWORDS

usability testing; user evaluation; user experience; contact tracing apps; CoronaMelder; COVID-19; pandemic; mobile apps; mHealth; public health

Introduction

After the World Health Organization officially declared the COVID-19 outbreak as a pandemic, countries all over the world were urged to implement strict interventions in order to limit viral spread and to prevent overload of health care systems [1]. The key essentials of these interventions focus on reducing the risk of COVID-19 transmission and consist of a package of instruments that are implemented worldwide and are based on responses to earlier pandemics. They include behavioral measures (social distancing, handwashing, personal protective equipment), adequate resources (personnel and materials for massive-scale testing, contact tracing and supported isolation), monitoring symptoms (contact tracing of possibly infected persons), and the use of digital tools [2,3].

In traditional contact tracing approaches, public health authorities (PHA) follow protocols that aim "to interrupt transmission chains by ensuring that persons who have been in contact with an infected individual are notified that they are at increased risk of infection and how to take action to prevent passing the infection to others" [2]. This is important because although the coronavirus incubation period ranges between 1 and 14 days [4], an infected individual can transmit the virus up to 48 hours before the onset of symptoms [5]. In addition, some infected individuals do not develop symptoms but are still infectious [6]. According to the contact tracing protocol, PHA (1) contact positive-tested individuals, (2) advise them about measures, (3) identify together with them how or by whom they were infected, (4) list and contact all persons they have been in contact with, and (5) arrange for individuals' contacts to be tested [7]. Despite being successful, traditional contact tracing by public health staff is labor-intensive, slow, and error-prone because people do not remember all their contacts [8,9]. Hence, the European Centre for Disease Prevention and Control has recommended the usage of digital tools, such as mobile tracing apps, to enhance and optimize traditional contact tracing [2].

Contact tracing apps could potentially provide several benefits as they (1) do not rely on the memory of the user (reminding users who they have had contact with), (2) allow contacts unknown to the user to be notified, (3) can speed up and enhance the tracing process, and (4) may facilitate further follow-up of contacts by PHA [2,6]. However, there are also some limitations in using these apps—not everyone has a smartphone or is able to carry their phone with them at all times; older smartphones or operating systems may not support newer apps (eg, newly developed apps can only operate on smartphones with operating systems iOS 13.5 or Android version 6, or later); the tracing technology inherently produces false positives and false negatives; and there are privacy concerns [2]. Furthermore, not everyone will be capable of or willing to use these apps (eg, older adults or vulnerable populations) [10]. These apps may complement but can never replace conventional contact tracing systems coordinated by PHA [10].

The Dutch Ministry of Health, Welfare and Sports (VWS) created conditions for the implementation of such an app. These conditions are listed in a *Plan of Requirements* [11] and mandate that the app should (1) be anonymous and voluntary to use; (2) be developed as open source (co-designed in an open Figma design platform); (3) notify users when they are at increased risk; (4) be in line with Guidelines for Infection Control [12]; (5) operate in addition to manual contact tracing (individuals should not receive help through the app that they would not receive without the app); (6) be inclusive (aimed at the largest possible relevant target group through explicit attention to language, literacy, and general/digital limitations); (7) aim to prevent reporting of false positives; and (8) involve international cooperation (eg, the app should be available on all phones operating on iOS and Android systems; connections between app users are made via Bluetooth; protection of privacy should be guaranteed [the app should be in line with common security standards and Web Content Accessibility guidelines, and a Data Privacy Impact Assessment should be performed]; and calculation of risks [distance, duration, and date of exposure] should be performed by the Google/Apple Exposure Notification framework [13]).

Contact tracing apps from other countries were examined by experts [14] to evaluate if these apps could also be implemented in the Netherlands. However, none of the evaluated apps met the above-mentioned criteria. The VWS therefore decided to develop a COVID-19 contact tracing app using the Google/Apple Exposure Notification framework that would be interoperable (to facilitate cross-border use): the CoronaMelder app. A development team, supported by an advisory committee and four task forces, was assigned to develop and test the CoronaMelder. The design of the app followed a privacy-by-design approach to minimize privacy invasion. During the development of the CoronaMelder, the app was tested with a variety of end users in field tests [15], think-aloud usability tests, practical tests, and ethical tests [16]. Additionally, the digital security of the app was tested using penetration tests and data privacy impact assessments. The findings of these tests led to continuous evaluation of the development and implementation of the CoronaMelder [16].

This paper focuses on the added value of the CoronaMelder app to support contact tracing. Think-aloud usability tests were conducted from June 29 to July 3, 2020, in the selected test region of Twente in the Netherlands. Twente was chosen for the current study because of its willingness to participate,

available expertise, and the coronavirus-proof test infrastructure of the University of Twente. The usability tests aimed to evaluate the user-friendliness, understandability, reliability and credibility, inclusiveness, and user experience of the CoronaMelder. These criteria were chosen because the CoronaMelder could only support traditional contact tracing by PHA if the criteria are satisfied. As the app is to be used by the public, it needs to be understandable and usable by users varying in terms of digital literacy, educational background, and ethnicity. To this end, various target groups were involved in the usability tests. The findings of this study will contribute improvements in the design of the app and support the VWS in its decision whether to launch the app. This paper aims to answer the following question: is the CoronaMelder user-friendly, understandable, reliable and credible, and inclusive?

Methods

Setting

The study consisted of scenario-based usability tests with additional interview questions and the Dutch User Experience Questionnaire (UEQ-Dutch) [17]. The usability tests were conducted using a scenario-based, think-aloud method [18], captured by researcher observations and voice recordings, and took place from June 29 to July 3, 2020. A beta version of the app was evaluated using test iOS (version 0.1, build 172) and Android (version 0.3.1, build 107) phones. Mock-ups (Figma, version 0.7.1) were used for the scenarios that could not be tested in the beta version due to the current stage of development (ie, being able to download the app from the App Store). The study was conducted at the University of Twente's DesignLab and in the Experivan mobile lab [19], which was used to visit participants with an intellectual disability. The BMS Lab protocol for coronavirus-safe research on humans has been approved by the Executive Board of the University of Twente. Hands and equipment were disinfected before and after the tests, and national measures were followed (per the National Institute for Public Health and the Environment [RIVM] [20]).

Participants

Participants were recruited using convenience sampling in Twente.

To test whether the app is inclusive, participants from the following target groups were included:

- Youth (<21 years) with a lower level of education (n=14); this included primary education and prevocational secondary education;
- Youth (<21 years) with a higher level of education (n=5); this included senior general secondary education, preuniversity senior secondary vocational education, higher vocational education, and university education;
- Youth with an intellectual disability (n=4);
- Migrants (n=2);
- Adults (40-64 years) (n=5);
- Older adults (>65 years) (n=14).

Ethics Approval and Consent to Participate

The study was approved by the university's Ethical Committee (BCE200953).

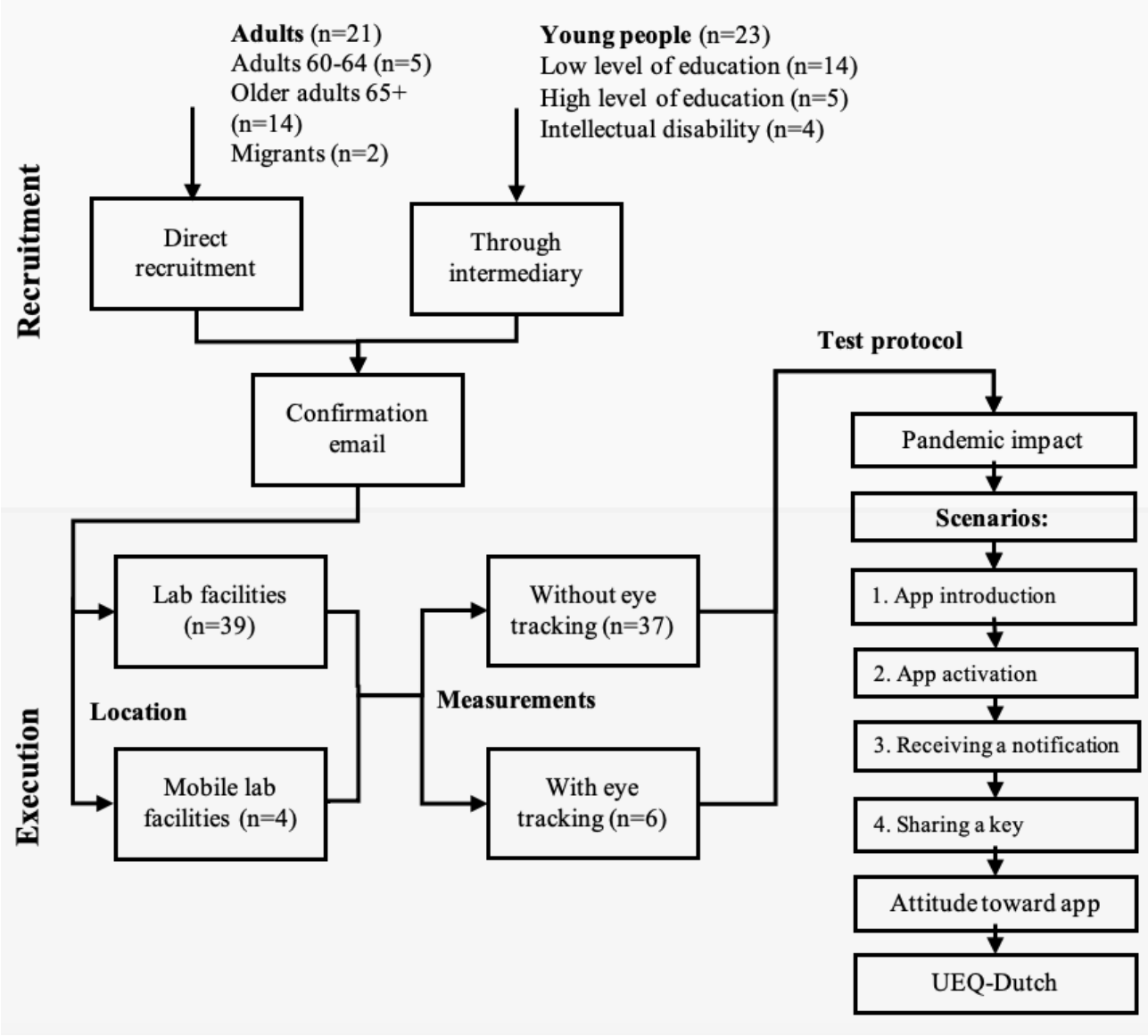
Participants were informed of the voluntary nature of their participation and confidentiality was guaranteed. All participants signed an informed consent, and a parent or guardian signed the informed consent for underaged participants.

Procedure

Adults (aged 40-64 years), older adults (aged >65 years), and migrants were contacted by the researchers by telephone. In this call, an explanation was provided about the nature and purpose of the study and an appointment was scheduled for the test. Recruitment of the younger participants was arranged through an intermediary (eg, school principal, mentor). The parents or guardians of minors were contacted by the intermediary via telephone and asked for their permission to let their children participate in the study. After the call with the researcher, all participants (or intermediaries) received email confirmation of the appointment, which also included additional information on participation. Figure 1 provides an overview of the procedure for participants.

The usability tests were conducted individually and in pairs in the case of youth in the lower-education category since research shows that minors are more capable of identifying a larger number of problems and details while working in pairs [21]. Before the test, the nature and purpose of the study were explained again, and permission for participation and audio recordings was given by signing the informed consent form. Parents or guardians had to sign the informed consent for minor youth. After providing additional consent, 6 participants additionally wore Tobii eye-tracker glasses (Tobii Gaming) for gaze analysis. An eye tracker was used due to people's strong urge to move their eyes, so that the fovea is pointed at whatever stimulus they are currently thinking about or processing. This is referred to as the so-called eye-mind link [22-24]. This theory makes eye tracking a reliable tool for exploring visual attention. These results are discussed elsewhere [25] and are outside of the scope of this paper.

Figure 1. Flowchart of study recruitment and procedure. The top part visualizes the recruitment methods, through both direct (n=21) and indirect (via intermediaries/representatives, n=23) channels. Included participants were tested in a stationary and mobile lab, with or without additional eye tracking, following the test protocol depicted on the right. UEQ-Dutch: Dutch User Experience Questionnaire.



The test protocol (Multimedia Appendix 1) started with general questions about the impact of the pandemic on the participants' lives and about what they had already heard about contact tracing apps for COVID-19. Thereafter, each participant actioned four scenarios on the app, which represented actual use of the app (1-hour test) while simultaneously thinking aloud. Before the usability test started, participants could choose between an iOS and Android test smartphone, per their preference. The four scenarios were:

1. Introduction to the app: in this scenario, the app was shown in the App Store, and additional information about the app could be read. Researchers focused on whether participants understood how to download the app and where they could find additional information, as well as whether they read the information.
2. Onboarding and activation of the app: in this scenario, the app's operation was explained through onboarding steps, in which participants learned about the content of the app and confirmed the right settings for app use (allow the app to use Bluetooth and to send notifications). After onboarding was completed, the app was activated and participants had the opportunity to explore the app independently. Researchers focused on whether the explanation of the app's function was clear, how participants acted, and whether they understood the permissions they consented to.
3. Receiving notifications: in this scenario, participants receive a notification from the app about their increased risk of infection because they have been in close contact with an individual who has tested positive for COVID-19. Researchers focused on whether participants understood what a notification entailed, whether it was clear to them why they received a notification, and what an increased risk status was based on. It was also examined whether it was clear to the participants what actions they should take after receiving a notification.
4. Sharing keys (telephone conversation with PHA): in this scenario, participants were asked to imagine they had recently been tested for COVID-19. During the scenario,

participants received a phone call from a PHA worker, informing them of a positive test result. The PHA worker followed the Dutch Gemeentelijke Gezondheidsdienst (Regional Public Health Services) telephone script (Multimedia Appendix 2), in which the participant was asked about their symptoms and received help with sharing their key (ie, their assigned identifier). First, the participant had to provide the key to the PHA worker by phone. Next, the participant had to click a button on the app interface to share the key with other app users (to warn the people they were in contact with).

After completing the scenarios, closing interview questions were asked about the participants' attitude toward the app and their willingness to use it. Additionally, a general questionnaire was administered (Multimedia Appendix 3), collecting data on gender, age, highest completed level of education, physical limitations in using apps, and a self-reported digital skills assessment, combined with the UEQ-Dutch. Within the UEQ-Dutch, participants had to assess different characteristics of the app (eg, whether the app is easy to use, visually attractive, supportive, and reliable) on a 7-point Likert scale. Researchers focused on whether the steps to be completed on the app were clear and easy to follow, and whether participants understood the utility and consequences of sharing their key. Differences between target groups were explored to investigate whether the app is inclusive. Whether the conversation with the PHA worker matched the steps that must be completed in the app was also examined.

Data Analysis

The recordings of the usability tests were pseudonymized (by BB) and stored on a data server at the University of Twente and were only accessible to the researchers involved. These storage and associated processes were certified according to ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) 27001 and NEN 7510 standards. Recordings of 3 participants (1 youth with a higher educational level and 2 youths with an intellectual disability) were not stored correctly due to technical errors, so these participants were excluded from the study. Recordings were transcribed verbatim (by MS, LB, JG), and all transcripts were analyzed (by BB and MS) to identify fragments on user-friendliness, understandability, reliability and credibility, and inclusiveness. Relevant fragments were labeled with the main codes "user-friendliness," "understandability," "reliability and credibility," and "inclusiveness" in Microsoft Excel. Next, the fragments within the main codes were analyzed axially to link fragments to each other and create new subcodes within each main code. Two researchers (BB and MS) coded 6 transcripts together to determine coding agreements. BB and MS each coded half of the other transcripts while considering the coding agreements. The coding scheme was revised several times by both researchers, and fragments were reread and recoded if necessary.

Descriptive statistics were generated for the general questionnaire using SPSS, version 20 (IBM Corp) [26]. The outcomes of the UEQ-Dutch were analyzed within the data analysis tool [27] offered by UEQ. This tool is benchmarked every year by UEQ to keep the validity of the tools as high as possible. The data gathered from the UEQ-Dutch questionnaire were input into the tool, which resulted in descriptive statistics and graphical representations of the gathered UEQ-Dutch data.

Availability of Data and Materials

The transcribed data are not publicly available due to privacy restrictions but are available from the corresponding author upon reasonable request. All screenshots of the CoronaMelder app are online, available via Figma.

Results

This section discusses participant demographics as well as the user-friendliness, understandability, reliability and credibility, and inclusiveness results of the CoronaMelder app.

Participant Characteristics

In total, data from 44 participants were included in this study. The sample comprised 31 males (70.5%). The mean age was 40 years, ranging from 13 to 79 years, and 26 (59%) participants reported having completed a higher education level. The majority indicated not having physical limitations to using apps in general. Nearly all rated their digital skills to be between "not handy, not clumsy" and "very handy." All participant characteristics are listed in detail in Table 1.

The following sections will focus on the research question: is the CoronaMelder user-friendly, understandable, reliable and credible, and inclusive? Table 2 presents an overview of how many arguments were mentioned by participants per theme, indicating their positive or negative opinion, or whether they understood the items related to the theme. The majority of the participants were positive about the user-friendliness, reliability, and credibility of the app. Participants from all target groups indicated more negative comments than positive ones regarding the understandability of the working mechanism of the CoronaMelder.

Multimedia Appendix 4 provides an overview of the number of positive and negative comments per topic, per target group.

Table 1. Demographic data of participants (N=44).

Characteristic	Value
Age (years), mean (range)	40 (13-79)
Gender, n (%)	
Male	31 (70.5)
Female	13 (29.5)
Highest completed education level, n (%)	
None or primary education	12 (27.3)
Preparatory secondary vocational education (practical)	3 (6.8)
Preparatory secondary vocational education (theoretical)	3 (6.8)
Secondary vocational education	0 (0)
General secondary education/secondary university education	6 (13.6)
Propedeutic (higher professional education or scientific education)	4 (9.1)
Bachelor's degree (higher professional education or scientific education	4 (9.1)
Master's or doctoral degree	12 (27.3)
Physical limitations in using apps in general, n (%)	
I have trouble reading	2 (4.5)
I am dyslectic	1 (2.3)
I am visually impaired	0 (0)
I have a motor disability	0 (0)
I am hard of hearing	1 (2.3)
I have limited digital skills	2 (4.5)
Other (please specify)	1 (2.3)
None	36 (81.8)
Did not state	1 (2.3)
Self-reported digital skills assessment, n (%)	
Very handy	9 (20.5)
Handy	22 (50.0)
Not handy, not clumsy	12 (27.3)
Clumsy	1 (2.3)
Very clumsy	9 (20.5)
I do not know, no opinion	0 (0)

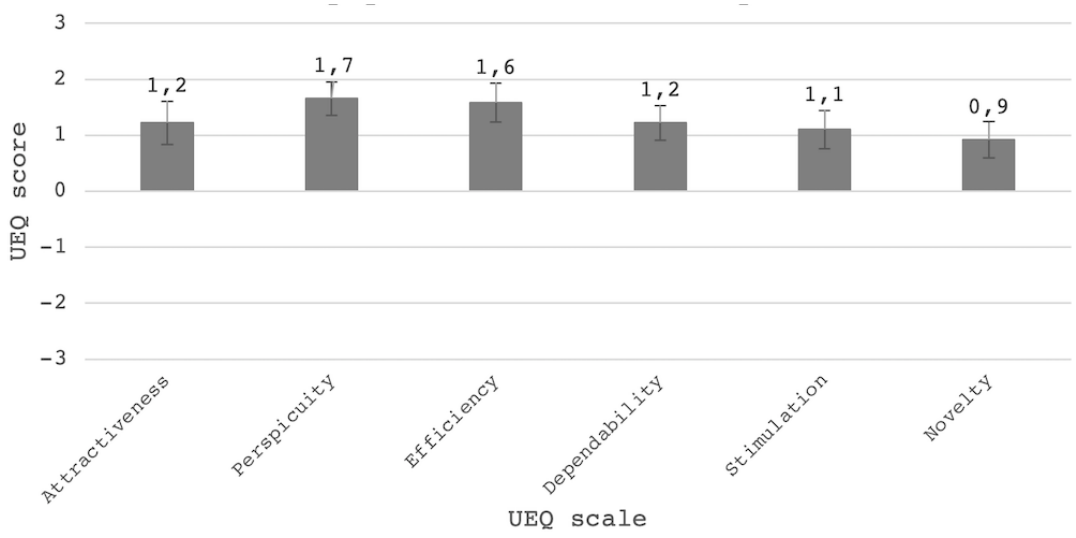
Table 2. Number of participants who stated a positive or negative comment about the CoronaMelder per topic (user-friendliness, understandability, reliability and credibility, and inclusiveness). For understandability of the notification and key-sharing system, the table demonstrates how many participants understood how the CoronaMelder app worked.

Theme	Positive/understand, n (%)	Negative/did not understand, n (%)
User-friendliness		
Layout	23 (53)	8 (18)
Navigation	33 (75)	9 (20)
Understandability		
Language	7 (16)	10 (23)
Receiving a notification	15 (34)	21 (48)
Sharing the key	15 (34)	19 (43)
Reliability and credibility	19 (43)	6 (14)
Inclusiveness	5 (11)	9 (20)

User-Friendliness

User-friendliness was assessed with the UEQ-Dutch and during the interview. In Figure 2, the outcomes of the UEQ-Dutch for the entire population are displayed for each of the six assessed scales. The scales include *attractiveness* (an overall impression of how much users like or dislike the app), *perspicuity* (how easy it is for participants to become familiar with the app), *efficiency* (the ability of users to use the app as intended), *dependability* (whether the user feels in control of the interaction with the app), *stimulation* (whether the app is exciting and motivating to use), and *novelty* (whether the app catches the interest of the user).

Figure 2. Boxplot with a mean score per User Experience Questionnaire (UEQ) scale for the total population, scoring from –3 (horribly bad) to +3 (extremely good). The error bars show confidence intervals.

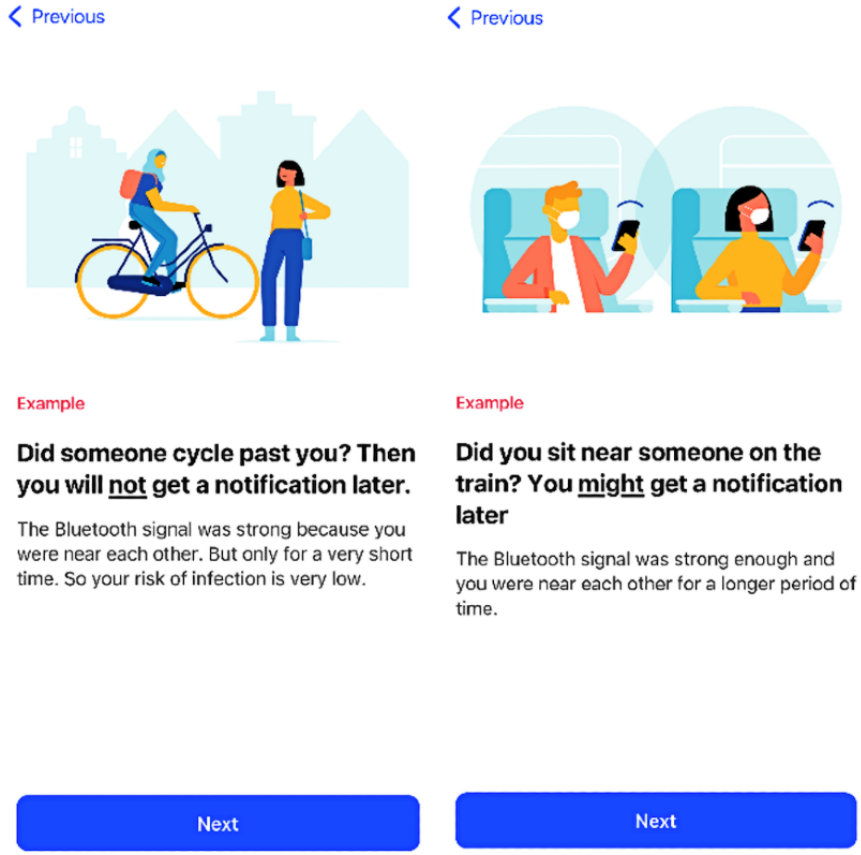


Participants rated the CoronaMelder app between 0.9 and 1.7 for all scales, which represents a positive evaluation (values >0.8) [28]. For perspicuity and efficiency, the app was rated relatively high, indicating the CoronaMelder is easy to become familiar with and participants were able to use the app as intended. The CoronaMelder scored relatively low on novelty, indicating that it was not proficient in capturing user interest.

Layout

The interviews showed that most of the participants liked the style of the app. The use of pictures within the app was appreciated, particularly by youth with an intellectual disability, who indicated having difficulties reading long texts. Representation and inclusiveness were achieved by displaying images of people from different cultures (Figure 3). The examples of when exposure leads to an increased risk (Figure 3) were commended since they helped participants better understand when they were exposed. Below is a sample response from participant 38, a youth with an intellectual disability:

Figure 3. App screenshots viewed during onboarding: (left) an example of when users are not exposed to increased risk and (right) an example of when users might be exposed to increased risk.



> *Question: Based on what you've read so far, what things do you like and recognize as being important for you?*
>
> *Participant: That it is explained in a simple way and that not too many words are getting used.*
>
> *Question: And do you think that this is addressed here?*
>
> *Participant: Yes.*

One negative aspect included participants' report of not reading long texts or only quickly scanning a text by reading the subheadings and words in bold to understand the most important information:

> *Well, I only read bold letters. I always want to be able to download an app quickly, so I am not going to read everything. The smaller, non-bold letters are then less important, I guess, so actually I skipped those.* [Participant 2, adult]

A few participants suggested that videos in which the information within the app is explained in more detail would be useful because not everyone likes to read, or some may not be able to read well. Additionally, using visualizations was recommended: youth with a lower educational level and those with an intellectual disability were not able to understand how the app worked by only reading the informational texts within the app:

> *It might have been useful to design a human with a mobile phone, standing with a group of people, one of whom appeared to be infected, and they all use the app. If there are 5 people and 1 have the virus, then if you make it visual how the app makes contact with other apps, you can explain what Bluetooth does. Because I don't know if everyone knows what Bluetooth is.* [Participant 20, adult]

Navigation

Most comments on navigation were positive. For example, participants considered the flow of information to be logical. Regardless, multiple issues with navigation were identified. First, it was unclear to participants whether an additional information page with more explanations about the app could be opened in the Google Play Store (Figure 4); hence, the page was not read by every participant. In particular, older participants did not know it was possible to consult additional information, although they reported wanting to read extra information.

Figure 4. The CoronaMelder app as seen in the Google Play Store. The app can be installed by clicking the *Install* button. The text below this button reads: "Receive a notification if you have been in close contact with someone who later tested positive for COVID-19. This app is authorized for use in the Netherlands and uses the Google/Apple Exposure Notification system." Additional information can be found in the *About this app* section. The text under this heading reads: "Help to stop the spread of the coronavirus in the Netherlands.".

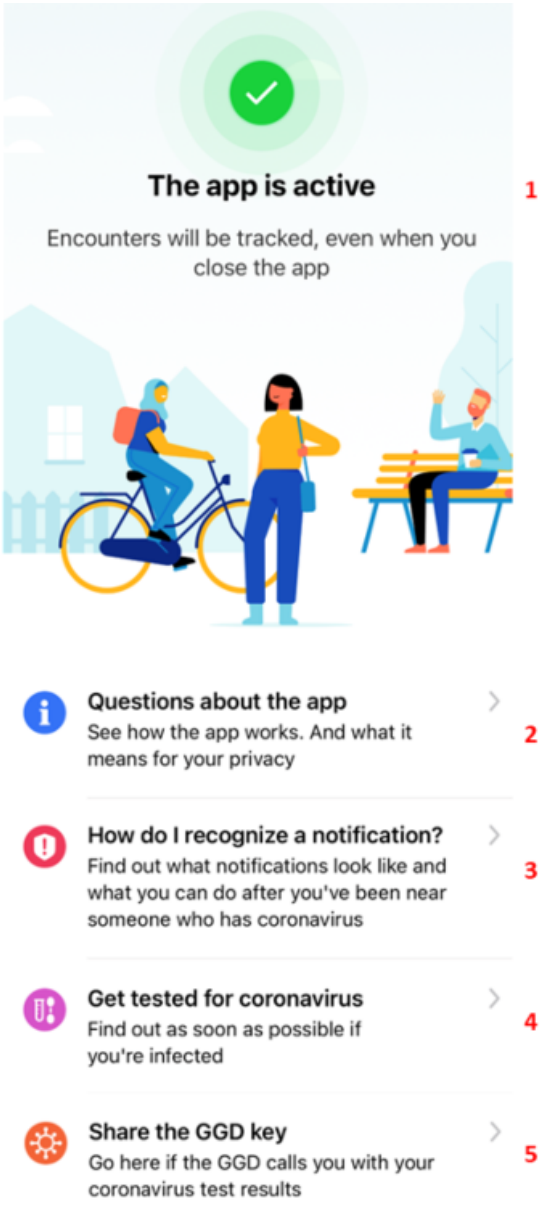
Second, within the app, some buttons (eg, *Sharing the Key* and *Request a Corona Test*) were only visible after scrolling, which was not clear for some participants, who therefore could not find the buttons:

> *I didn't know there was another fourth button […] Sometimes you think that what is visible here, is everything. Then I will not go scrolling automatically, but it could just be me.* [Participant 27, adult]

Third, the app did not provide clear expectations to the participants on what to do after activating the app. After activation, the home screen (Figure 5) is displayed, informing users that the app is activated and ready for use. Younger participants automatically closed the app, whereas older participants became disoriented and reported being unsure about what to do next:

> *You could also put that in the text 'the app is active and if you are sufficiently informed you don't have to do anything', because people will ask 'what now?'* [Participant 8, adult]

Figure 5. The app's home screen, which appears after the user has finished the onboarding and activation of the app. At the top (1), the app explains its activation for use. Below this are several information pages on (2) how the app works (frequently asked questions), (3) what to do if one receives a notification, (4) how to request a COVID-19 test, and (5) how to share one's key.



Positive remarks on navigation included the ease of onboarding and activating the app and the logical sequence flow of providing information about the app during onboarding (Multimedia Appendix 5). The majority of both older and younger participants indicated that they normally would click quickly through the screens and only read the information properly since they were participating in the study. The younger participants gave permission to receive notifications and use Bluetooth without delay, while the older participants thought carefully about granting permission:

> *Most of it is very logical. I now read the texts properly, but usually I would probably click and skip faster through it.* [Participant 4, older adult]

Understandability

Participants from all target groups provided more negative comments than positive ones on the understandability of several topics. Understandability problems occurred due to inconsistency in the terms used and not reading information on how the app works. Most problems concerned receiving a notification from the app and sharing the key when tested positive to support contact tracing. The app was inconsistent in the ways it referred to the coronavirus (eg, "coronavirus," "COVID-19," and "corona") and to the key that should be shared after testing positive (eg, "ID," "code," "control code"). Furthermore, a clear definition or explanation was lacking within the app about what it means to be exposed ("exposure") or at "increased risk" of a COVID-19 infection. The texts also include technical English vocabulary, such as "ID," "share," "enable," and "upload." Older participants in particular did not know what those words meant and were confused by them:

> *Now suddenly some English words are used. Well, that is a problem for some people. You should not do that. Or you should provide both languages. But now you have people who get stuck around here […] this can confuse people.* [Participant 4, older adult]

Understandability of Receiving Notifications When Exposed to an Increased Risk

The app sends a notification to the phone, explaining that the user is at increased risk because on a certain date they were close for more than 10 minutes to another app user who later tested positive. Opening the notification brings the user to an information page in the app that explains what the user should do.

Being at Risk: When and How a Notification Will Be Received

The test showed that the majority of participants, but in particular youth in the lower-education category, did not understand under what circumstances they would receive a notification; for example, participants believed they would receive an alarm immediately after exposure:

> *But I am not quite sure how it works exactly, whether it's anonymous or when you receive a notification… If I understand correctly, you will receive a notification if you have been with someone for more than 10 minutes, but then you do not know whether the person is infected or not? And if they have been tested, you will receive a notification that they were infected. Can I figure that out based on this? Well, I don't think so.* [Participant 32, older adult]

Youth in the higher-education category, adults, and older adults also reported that the time between exposure and notification (within 14 days) is too long, although they can imagine why this is the case. A few participants even labeled the app as useless when they did not receive a notification immediately after being exposed because by the time they receive the notification, according to them, it is too late to take appropriate action:

> *It would be nice if, for example, someone has corona, then if you walk by, your phone will beep at once, like a message will be given […] Because after five days, it is already too late. If someone has the coronavirus and you immediately get a message, then you know, oh I must keep my distance.* [Participant 14, youth, lower-education category]

None of the participants, regardless of age and education, understood when they were at increased risk for possible infection. For example, they were not aware of how long and how close they must be to someone who turned out to be infected to receive an increased risk notification. Participants also appeared not to understand that the app only communicates with other apps, so they will only receive a notification if the infected person also uses the app. It was unclear to participants that exposure detection was based on Bluetooth connection between different smartphones and based on actually being exposed to the infected persons themselves. The exposure threshold level of 10 minutes was questioned by participants. Participants (across all target groups) believed that they can also be infected when they are in close proximity to someone for less than 10 minutes.

> *But 10 minutes … isn't that quite long? Suppose he has coronavirus and I stand near him for 2 minutes and then leave […] 2 minutes is enough anyway.* [Participant 15, youth, lower-education category]

What to Do After Receiving a Notification

While youth with both lower and higher educational levels indicated that the app adequately explained what to do after they receive a notification, both adults and older adults mentioned that the app does not provide clear advice and even gives contradicting advice. For example, the app advises staying at home but also continuing daily life while being aware of symptoms.

> *But what I do want to know – and I miss that in here – is: what should I do now? I would like to know very specifically: what should I do? What options do I have? […] Or even more socially democratic: we recommend the following […] Suppose if you receive a notification, I want to know what now? What should I do? Then I don't need to know about symptoms or about a corona test...* [Participant 12, adult]

Participants emphasized that the app should clarify how it notifies users, and what users are expected to do after receiving this notification. A few participants reported that they did not find the information about symptoms of the coronavirus and the possibility of requesting a test useful; they preferred to read advice on what they should do at that moment:

> *If you get a notification, 1) I want to know how I get this notification, 2) suppose I have received a notification, what now? What am I supposed to do? I don't need to know about symptoms or about a corona test then.* [Participant 12, adult]

Understandability of Sharing a Key After Testing Positive

Various understandability issues were identified in the scenario of a positive test, as well as some positive remarks. The issues revolved around not understanding what the key is, where to find it, when to share it, and how to share it. For example, some participants mentioned they did not know what the key involved or what they were expected to do:

> *Well, I see now that it [the app] works through a key and I haven't read about that anywhere yet. So, I don't know what that key involves.* [Participant 12, adult]

Participants were also unaware that a PHA worker would call the participant to initiate key sharing:

> *Now they say 'then the PHA worker will ask you in the telephone conversation to share the key from the app and then upload the keys of the telephones you've been in contact with.' What are they talking about? Which key? I don't know what key they are talking about.* [Participant 12, adult]

A few youths in the lower-education category or with an intellectual disability did not understand how the app knew that one had received a positive test result:

> *But how does the app know you have corona? Do you have to type that in the app?* [Participant 15, youth, lower-education category]

Participants expressed different opinions regarding the text about sharing the key. Youth in the lower education category appeared to easily share the key and did not want information about what was expected from them. A few adults mentioned that the symptoms of the coronavirus and the implemented measurements are repeated too often within the app. The texts were considered to be too long, and it was reported that an overload of information should be prevented. Adults in particular expressed their indignation about the choice to share one's key; they mentioned that if people were not willing to share the key, they should not have downloaded the app:

> *If I have the app, wouldn't that obligate me to share the key? I think you should share the key, that this choice option doesn't have to be in there. Otherwise, I wouldn't have to install the app. We try to help get this virus under control, together. Together, therefore, means that you have to share this information with others. So, I think that this choice to share or not is ridiculous.* [Participant 5, adult]

The steps participants must perform to share the key were clear to most youth and adults, although some older adults did not understand which process starts after they click on the *Share Your Code* button. It was unclear to them why they must point the key to the PHA worker first and whether they had to share the key afterward with other users by clicking on the button. Some thought they must send the key to their contacts via other communication channels:

> *My question is whether if it says 'share codes', whether this relates to the person I've been in contact with […] or whether codes are shared with the PHA, and the PHA then warns people I've been in contact with. That's unclear to me.* [Participant 18, older adult]

Some youths in the lower-education category pointed out that when the key is mentioned to a PHA worker, the app is not anonymous anymore since the PHA worker knows which key belongs to which person. Below is a sample exchange between participant 13 (youth, lower-education category) and the interviewer:

> *Participant: Well, now they suddenly have my number at the PHA?*
>
> *Interviewer: Yes, that's right, the corona test is not anonymous. But sharing the key is.*
>
> *Participant: But if I share my key with them, then it is not completely anonymous, right?*

In addition to this, some adults were irritated by the notification (Figure 6) that comes after sharing the key. They reported that one only used the app if one wanted to warn others and assessed the extra permission notification for sharing as unnecessary:

> *In my opinion, this is a strange choice. […] It was clearly emphasized at the beginning [onboarding] that the app is anonymous. I also think now, since I downloaded the app, I have to warn others. That's not a choice, it's a logical consequence of the fact that I installed the app. I specifically downloaded the app because it's anonymized. This is just part of the deal. I have no idea who I'm sending it to [the key], I don't know where I was [during the possible infection], I don't know anything, but I do know that others will receive a signal just like I received a signal. Then I should no longer have the choice of sharing or not sharing.* [Participant 5, adult]

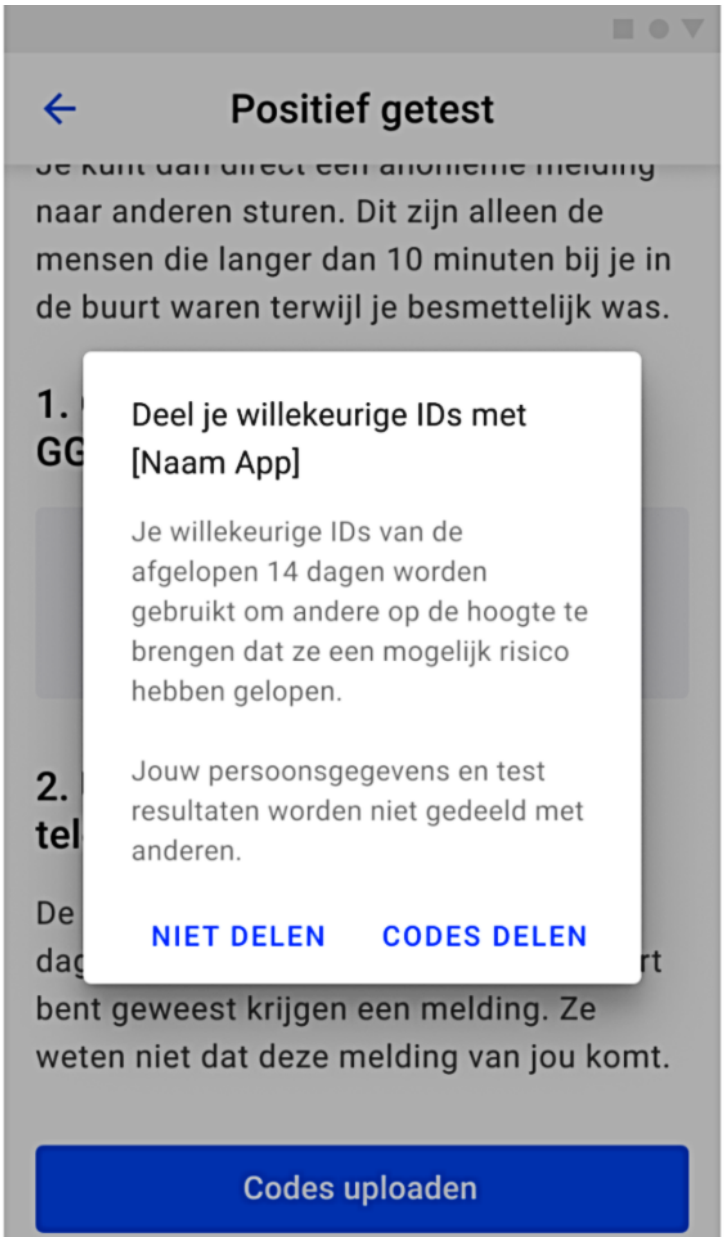
Participants, with the exception of older adults, considered the steps provided by the PHA worker for sharing the key on the app to be clear and easy to follow:

> *It was super easy for me. Even a young child can do this on their own.* [Participant 2, adult]
>
> *This is unclear, how to do that, because there is nothing, the screen says only 'close' to me […] Yes, because, what you said to me, I cannot carry out […] No, I only have to close the blue bar […] Maybe I can do that, maybe that will work […] No, I returned to the previous screen…* [Participant 4, older adult]

Participants did express that the simulated telephone conversation with the PHA worker lacks guided assistance, further explanations, and empathy. However, a telephone conversation was seen as a more personal approach, although youth with a lower educational level or with an intellectual disability preferred not to be called, and a few said they will not answer the phone if the PHA worker calls from a number with no calling ID. Additionally, too little attention was paid to the emotions that the message of testing positive can convey to participants. Participants suggested that PHA workers should make time to help them through the scenario, especially older adults, and should identify themselves to ensure reliability. Participants reported that the PHA worker should calmly go through the steps with the app users and not ask them to perform certain steps without providing an explanation:

> *I think it is pleasant if someone calls you and goes calmly through the app with you. That they don't just say, well you have to do this and that, but that they really explain step by step what to do and where I have to click, in the app. That might be useful for older people or people who do not know or understand how the app works. I think that would be useful.* [Participant 25, youth, higher-education category]

Figure 6. The notification that pops up after sharing the key, asking the user for permission to share their information and notify others after they click on *Share the Key*. The text reads: “Share your random IDs with [name of the app]. Your randomly generated IDs for the past 14 days will be used to notify others that they have been at increased risk. Your personal data and test results will not be shared with others.” The choice options are translated as “do not share” and “share codes.”.



Reliability and Credibility

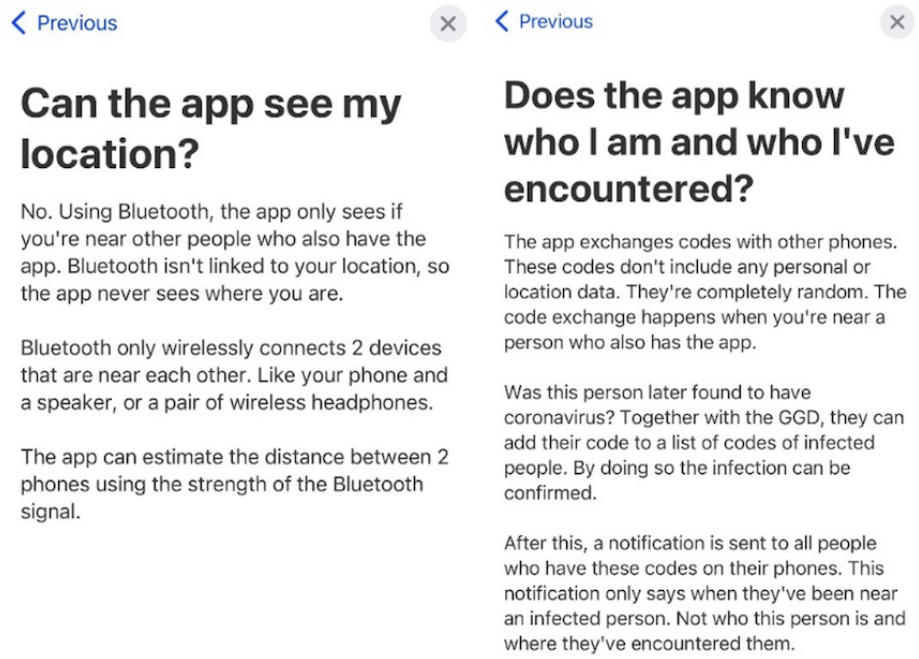
The app was assumed to be reliable because it was presented as a government app, and participants trust the VWS:

> *[Participant is reading additional information in the App Store] ‘Released by the Ministry of Health, Welfare and Sports’. Well, that seems confidential to me.* [Participant 19, youth, lower-education category]

The explanation about data storage and anonymity (Figure 7) earns trust, and adult participants in particular applauded the fact that the app does not require personal data.

> *I think it’s good that they clearly state what makes the app safe and anonymous, I think that’s strong […] Otherwise if these weren’t listed it would scare people off.* [Participant 21, youth, higher-education category]

Figure 7. Screenshot of frequently asked questions, with responses (left) indicating that the app does not track users' GPS and (right) how the app is anonymous.



Although adults assessed the app as trustworthy, youth in the lower-education group did not understand how the app guarantees privacy. Participants (from different target groups) believed that they would be tracked and that others would know their name and address when they reported a positive diagnosis. These misconceptions are not only caused by a lack of explanation or by participants not reading the information provided, but also because participants mistrust the use of Bluetooth—some participants may think it will still track their location and that the app will connect with people who are not directly in close proximity to them (eg, separated by a wall):

> *It says that the app knows via Bluetooth whether you were close to someone […] The app doesn't know where you were and who you are. But that's nonsense, it has to be. If you turn on Bluetooth, you immediately see where someone is […] That's through the Apple satellite, same for the Samsung satellite. They can always track your phone, it doesn't matter if you have turned it [GPS tracking] off. That's why it's nonsense, and they should add that. But well, if that is the case, if I already know that someone will use my location, I will immediately delete the app.* [Participant 15, youth, lower-education category]

Other participants were less doubtful about the use of Bluetooth, reporting that they thought the acquisition and storage of data was safe. Additionally, if privacy is guaranteed, multiple participants (mostly youth in the higher-education category and adults) mentioned their willingness to use the app:

> *Of course, in relation to privacy, you always check who monitors what data, but that will undoubtedly also be properly secured and your GPS location data will not be used or stored. That sounds safe, and I assume it is.* [Participant 4, older adult]

Inclusiveness

Differences in inclusiveness were found across age groups. In general, participants from all target groups with the exception of older adults did not have any problems using the app. Among older adults, there was a dichotomy between those who had digital skills and those who had trouble finding specific information within the app, opening and closing specific screens, and attending a telephone call while simultaneously opening and using the app:

> *Turn on speaker, that will be interesting. I'm going to see if I can do that. Speaker, yes, I did it! I succeeded. Well, to the Corona app, let me ask, how do I get there? These are things I'm not handy with. I have to go to the app. It works on my own phone, but now it won't. Close everything … no, I shouldn't do that. Ah, this one. Yes, I'm in the Corona app now.* [Participant 29, older adult]

Older adults were able to perform the steps to share the key under the guidance of the PHA worker but had difficulty performing these steps while talking on their mobile phone. For example, older adults were not aware of how they could turn on the speaker or close the call screen and open the app. However, they signaled willingness to learn how to use the app:

> *Oh that's difficult, then I have to make a phone call and look something up in the app. I don't know how to do that. Normally I can't even answer my phone when I'm doing something else on my phone.* [Participant 1, older adult]

Additionally, youth with an intellectual disability appeared to lean on the researcher while conducting the test. They were in doubt and asked for confirmation each time before clicking on a screen or button. They appeared to not know what they were doing or why they must do it. For example, while sharing the

key, they followed the steps the PHA worker provided them, without seeming to understand what they were doing or what would happen next.

Regarding the language used in the app, participants reacted differently according to their educational or cultural background. For example, youth in the higher-education category and both adult and older adults thought the language was clear and easy to understand and that appropriate words were used to express the purpose of the app:

> *Of course not everyone can read properly, that can be a bottleneck. The information should actually be as simple as possible. I think it is easy to read, but I don't know if that applies to everyone.* [Participant 11, adult]

Youth with a lower educational level, youth with an intellectual disability, and migrants reported that the words used were too difficult to understand and texts were too long. The latter indicated that the app should work in other languages, such as English and Arabic.

> *P1: It has really difficult words…P2: I agree, and difficult words are annoying to read* [Paired participants 16.1 and 16.2, youth, lower-education category]

> *Some words I don't understand so well. It would be easier for me if I could choose another language.* [Participant 41, migrant]

Discussion

Principal Findings

This study aims to answer the research question: is the CoronaMelder user-friendly, understandable, reliable and credible, and inclusive? Based on the findings, we can conclude that the CoronaMelder is easy to use. The app was seen by most as a good initiative because it warns them about possible infections, protects them, and could help avoid a second viral wave. The app was considered reliable because it is an initiative from the government (VWS). After participants read the information in the App Store, they indicated understanding how the app operates, and many were curious to become familiar with the app and expressed their intention to download it. Several general reasons why participants were willing to use the app were indicated, such as protecting themselves and their loved ones, creating sufficient support for the app, and helping to get COVID-19 under control and ease nationwide measures. However, it appeared that essential parts of the app were not understood by the participants, such as the notification system, the sharing of the authorization key via PHA, or how the app guarantees privacy.

Doubts and fears were expressed regarding privacy, usefulness, and consequences of the CoronaMelder. Among the reasons for these negative attitudes were fewer positive arguments in the media and the number of false positives. Reasons not to use the CoronaMelder were expressed, such as perceiving the app as useless, thinking the coronavirus and corresponding measures were overrated, not wanting to be in quarantine (without confirmed risk), and limited phone memory or battery capacity.

In terms of inclusiveness, it appears the CoronaMelder is not accessible to various target groups. Youth at a lower educational level or those with disabilities had difficulties using the app due to low literacy and language problems, and the older adults experienced difficulties related to limited digital skills.

Whether the app will be effective in supporting traditional contact tracing is a concern since the majority of participants did not understand how the app operates or why there is a delay between exposure/increased risk and receiving a notification. The app provides complex information and lacks explanations; therefore, users find it unclear what actions the app expects from them. This lack of clarity has led to misconceptions about the app in terms of operation, privacy, and usefulness, thus affecting participants' willingness to use it. This also affects the adoption of the app and adherence. Additionally, the protocol of PHA workers lacks guidance, explanation, and empathy, revealing that PHA staff are unprepared to support users with the app to the fullest extent during the pandemic (eg, with key-sharing procedures) in addition to their other responsibilities.

Comparison With Studies on Other COVID-19 Contact Tracing Apps

To the best of our knowledge, this is the first study to pretest the usability of the Dutch CoronaMelder app. Studies of other countries' contact tracing apps that operate similarly to the CoronaMelder reported comparable findings on participants' attitudes toward these kinds of apps. In a study by Horstmann et al [29] on the German Corona-Warn-App, participants indicated that there were no reasons not to use the app, that the benefits would outweigh the risks, and that they believed the app would contribute to slowing down the pandemic [29]. On the other hand, in studies of apps in Germany, Switzerland, and France, the most frequently mentioned reasons not to use the app were privacy concerns [29-35], doubts about usefulness [29,34,36], and lack of technical equipment (eg, not all smartphone operating systems can access the apps) [29,34,37]. In both of the studies on the Corona-Warn-App (Germany) [29] and the SwissCovid app (Switzerland) [34], as well as in a longitudinal survey about the Dutch CoronaMelder [38], privacy concerns appeared to be associated with a lack of trust in the national government (or in PHA).

In addition, the StopCovid app (France) showed low uptake [36,39] because the app appeared uninteresting and ineffective [36]. The same study also reported that 71% of their participants suggested that better communication strategies would increase uptake of the app. Our study stated that the reasons mentioned for not using the app were already expressed before contact tracing apps were launched. However, it was reported [29,40] that these concerns were still raised after app implementation, which indicates that public health campaigns promoting contact tracing apps were not able to eliminate these concerns.

Lessons Learned and Recommendations

Adequate and Targeted Communication

The tests showed that lack of clarity led to misconceptions about the app, affecting participants' willingness to use it. Communication about the app is therefore essential for

acceptance [3,29,40]. Targeted and tailored group-specific communication should happen through channels such as public campaigns, animations, social media, and ambassadors or influencers [3,41]. Communication should be customized according to the aim of the app in relation to other national COVID-19 measures (eg, testing, quarantine, social distancing) [3] and in collaboration with PHA and family physicians.

Providing more explanations, plus emphasizing the advantages of the app in comparison to regular contact tracing, increases the likelihood that, when individuals download the app, they will know what to expect as well as what is expected of them [29,35]. Walrave et al [35] had already reported that the intention to adopt a contact tracing app increases if people know how to use the app. Hence, it would be important to have a helpdesk where people can ask questions, for example, about the purpose of contact tracing and how the CoronaMelder contributes to this, how the app operates abroad (ie, outside of the Netherlands), how anonymity is guaranteed, how data is stored, what the role of PHA will be, and who to approach in case of uncertainty or fear regarding possible risks of infection.

Embedment in Traditional Contact Tracing by PHA

Our study suggests that it is important for PHA workers to be well prepared in guiding users to share their keys in case of positive test results. It is important to consider the app not as a stand-alone tool but as part of the pandemic infrastructure [3,35,41] and to embed it within PHA workers' workflow. Hence, it is crucial to provide access to tests, regardless of symptoms, but dependent on the contact date with an infected person; complete testing quickly and deliver tests results within 24 hours; clarify the scope of the app compared to other digital resources (eg, Dashboard, Thuisarts.nl) or apps to be developed; arrange international agreements about interoperability with contact tracing apps in other countries; facilitate effective and efficient interaction between PHA and the CoronaMelder; and evaluate the effects of the CoronaMelder on contact tracing, individuals' behavior, and society.

PHA should coordinate how their health care workers can guide individuals through the steps of sharing the key with the app. PHA workers should be trained to properly and empathetically explain which steps people must follow on the app. After all, PHA workers are responsible for both conducting the conversation about the test result and instructing users on the app. This means that they should be well educated about the aim and operation of the app and about their task and role during the phone call. It is therefore recommended to examine ways PHA workers can proceed to effectively and empathetically interact with app users.

Strengths and Limitations

The first strength of this study is its focus on participants of different backgrounds (age, education, etc) to test whether the CoronaMelder is accessible to all residents of the Netherlands. The second strength is the real-time pretesting of the key parts of the CoronaMelder to enable revisions before the definitive launch. The findings of the study were communicated with the software development team to exchange feedback on adjustments to the app and to revise the app during the test days. During development, minor adjustments in the app's design were made, which means that participants who tested the app later in the study may have viewed some screens that were different to those tested by participants earlier, even though the essential parts of the app were the same. Based on the findings of this study, the VWS decided to launch the app. However, the definitive launch (October 10, 2020) was postponed due to changes in testing policy. The premise "test without symptoms," which is an important driver for using the app, was changed due to a lack of testing capacity.

Future Research

To fulfil the requirements of the CoronaMelder [11], the design of the app can be improved. The accessibility and understandability of the app should be customized to differences in literacy and digital skills. Think-aloud, real-world–based scenarios and eye tracking should be designed to involve end users with different literacy levels and digital skills to test use of the app in real time.

Evaluation of the CoronaMelder app should focus on the key essentials of the app to support early and better contact tracing. Data should therefore be collected on use and adherence regarding follow-up actions after a notification, sharing a key to inform PHA and other users (via the app), and going into isolation when testing positive. The privacy-by-design policy could complicate gaining insight into the added value of the CoronaMelder app since it hinders data collection. A critical view is needed on how to find a balance between user-centered design and the privacy-by-design policy.

Future studies should also focus on how communication campaigns can be best targeted and customized to reduce uncertainties and misconceptions, thereby improving the understanding of digital contact tracing apps as well as adoption and adherence. Overall, an adequate infrastructure (resources, personnel, capacities, etc) and powerful management tools are needed to implement the CoronaMelder and other digital tools to facilitate and optimize contact tracing to fight a pandemic. The COVID-19 pandemic has had an adverse global impact and requires an interdisciplinary approach. Future studies of the CoronaMelder app should consider the app not as a stand-alone device but as part of a coherent package of anti–COVID-19 measures to fight the pandemic, considering the impact on users, stakeholders, and testing and tracing procedures.

Acknowledgments

We want to thank Jeanique Wegdam (premaster student, Industrial Engineering and Management, University of Twente) for her help with transcribing the interviews. The Dutch Ministry of Health, Welfare and Sports funded the usability study. The DesignLab

of the University of Twente facilitated the research by means of a corona proof usability test environment. The Behavioural Management and Social Sciences Lab provided the test infrastructure and the ExperiVan to enable mobile testing.

Authors' Contributions

BB and JGP were responsible for the coordination of the usability tests. SK, BB, and JGP were involved in the creation of the scenarios and interview scheme. JK, JG, and PS delivered the materials. MS, LB, JG, and PS performed the usability tests. BB and MS analyzed the data and discussed the findings with all authors. BB was a major contributor in writing the manuscript, JGP, JK, and SK contributed to the manuscript by providing feedback and discussing the interpretation of results. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Test protocol.
[DOCX File , 31 KB - formative_v5i3e27882_app1.docx]

Multimedia Appendix 2
PHA telephone script for positive test results.
[DOCX File , 14 KB - formative_v5i3e27882_app2.docx]

Multimedia Appendix 3
The UEQ-Dutch questionnaire.
[DOCX File , 151 KB - formative_v5i3e27882_app3.docx]

Multimedia Appendix 4
Number of participants per target group who stated a positive or negative comment about the CoronaMelder, per topic (user-friendliness, understandability, reliability and credibility, and inclusiveness). In the case of the understandability of the notification system and sharing of the key, the table shows how many participants understood how the CoronaMelder app worked.
[DOCX File , 17 KB - formative_v5i3e27882_app4.docx]

Multimedia Appendix 5
Steps to onboard and activate the CoronaMelder app.
[DOCX File , 470 KB - formative_v5i3e27882_app5.docx]

References

1. WHO announces COVID-19 outbreak a pandemic. World Health Organization. 2020 Mar 12. URL: https://tinyurl.com/nfsz9hkr [accessed 2020-07-22]
2. Mobile applications in support of contact tracing for COVID-19 - A guidance for EU EEA Member States. European Centre for Disease Prevention and Control. 2020 Jun 10. URL: https://tinyurl.com/y7wx2pzy [accessed 2020-07-22]
3. Allen D, Block S, Cohen J, Eckersley P, Eifler M, Gostin L, et al. Roadmap to Pandemic Resilience. Edmond J Safra Center for Ethics. 2020 Apr 20. URL: https://ethics.harvard.edu/Covid-Roadmap [accessed 2020-07-22]
4. When to Quarantine. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/if-you-are-sick/quarantine.html [accessed 2020-10-29]
5. Rapid risk assessment: Coronavirus disease 2019 (COVID-19) pandemic: increased transmission in the EU/EEA and the UK – seventh update. European Centre for Disease Prevention and Control. 2020 May 25. URL: https://tinyurl.com/3xmd7kt7 [accessed 2020-07-22]
6. Kretzschmar ME, Rozhnova G, Bootsma MCJ, van Boven M, van de Wijgert JHHM, Bonten MJM. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. The Lancet Public Health 2020 Aug;5(8):e452-e459. [doi: 10.1016/s2468-2667(20)30157-2]
7. Factsheet: Coronavirus: how does contact tracing work? Government of the Netherlands. 2020 Aug 18. URL: https://www.government.nl/documents/publications/2020/06/08/coronavirus-how-does-contact-tracing-work [accessed 2020-09-10]
8. Bi Q, Wu Y, Mei S, Ye M, Zou X, Zhang Z, et al. Epidemiology and Transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts. medRxiv. Preprint posted online March 27, 2020 [FREE Full text] [doi: 10.1101/2020.03.03.20028423]

9. Chen W, Wang Q, Li YQ, Yu HL, Xia YY, Zhang ML, et al. [Early containment strategies and core measures for prevention and control of novel coronavirus pneumonia in China]. Zhonghua Yu Fang Yi Xue Za Zhi 2020 Mar 06;54(3):239-244. [doi: 10.3760/cma.j.issn.0253-9624.2020.03.003] [Medline: 32064856]
10. Contact tracing for COVID-19: current evidence, options for scale-up and an assessment of resources needed. European Centre for Disease Prevention and Control. 2020 May 05. URL: https://www.ecdc.europa.eu/en/publications-data/contact-tracing-covid-19-evidence-scale-up-assessment-resources [accessed 2020-09-10]
11. Programma van Eisen voor een digitale oplossing ter aanvulling op bron- en contactonderzoek van de GGD. Rijksoverheid. 2020 Aug 25. URL: https://www.rijksoverheid.nl/documenten/richtlijnen/2020/08/25/programma-van-eisen-digitale-aanvulling-bron--en-contactonderzoek-ggd [accessed 2020-09-10]
12. COVID-19 Richtlijn. Landelijke Coördinatie Infectieziektenbestrijding (LCI). 2020. URL: https://lci.rivm.nl/richtlijnen/covid-19 [accessed 2020-11-13]
13. Google Ondersteuning van volksgezondheidsinstanties. Government of the Netherlands. 2020 Aug 28. URL: https://www.rijksoverheid.nl/documenten/publicaties/2020/08/28/google-ondersteuning-van-volksgezondheidsinstanties [accessed 2020-10-19]
14. KPMG. Securitytest potentiële Corona-apps. Rijksoverheid. 2020 Apr 19. URL: https://www.rijksoverheid.nl/documenten/publicaties/2020/04/19/rapportage-veiligheidstest-potentiele-corona-apps [accessed 2020-10-19]
15. Veldtest Bluetooth Validatie COVID-19 Notificatie APP. Rijksoverheid. 2020 Jun 08. URL: https://tinyurl.com/fa4cnz3k [accessed 2020-10-19]
16. Hoe is CoronaMelder getest? Rijksoverheid. URL: https://tinyurl.com/3zp6fkjb [accessed 2020-10-19]
17. Hinderks A, Schrepp M, Domínguez Mayo FJ, Escalona MJ, Thomaschewski J. Developing a UX KPI based on the user experience questionnaire. Computer Standards & Interfaces 2019 Jul;65:38-44. [doi: 10.1016/j.csi.2019.01.007]
18. van Someren M, Barnard Y, Sandberg J. The Think Aloud Method - A Practical Guide to Modelling Cognitive Processes. London, UK: Academic Press; 1994.
19. van 't Klooster JW, Slijkhuis P, Langener S, Nijen Es L, Beerlage-de Jong N, van Gemert-Pijnen L. Research on the road using the ExperiVan. In: Abstract Persuasive Health Conference. 2021 Presented at: Health by Tech Conference; Jun 11-12; Enschede, Netherlands p. 40-44 URL: https://research.utwente.nl/en/publications/research-on-the-road-using-the-experivan
20. COVID-19 (Nieuwe Coronavirus). RIVM. URL: https://www.rivm.nl/coronavirus-covid-19 [accessed 2020-11-06]
21. Als BS, Jensen JJ, Skov MB. Comparison of think-aloud and constructive interaction in usability testing with children. In: Proceedings of the 2005 Conference on Interaction Design and Children. 2005 Presented at: IDC05: Interaction Design and Children; June; Boulder, CO p. 9-16. [doi: 10.1145/1109540.1109542]
22. Just MA, Carpenter PA. A theory of reading: From eye fixations to comprehension. Psychological Review 1980;87(4):329-354 [FREE Full text] [doi: 10.1037/0033-295x.87.4.329]
23. Rayner K. Eye movements and attention in reading, scene perception, and visual search. Q J Exp Psychol (Hove) 2009 Aug;62(8):1457-1506 [FREE Full text] [doi: 10.1080/17470210902816461] [Medline: 19449261]
24. Rayner K, Reingold E. Evidence for direct cognitive control of fixation durations during reading. Current Opinion in Behavioral Sciences 2015 Feb;1:107-112 [FREE Full text] [doi: 10.1016/j.cobeha.2014.10.008]
25. van 't Klooster JWJR, Slijkhuis PJH, van Gend JE, Bente BE, van Gemert-Pijnen JEWC. First Eyetracking Results of Dutch CoronaMelder Contact Tracing and Notification App. In: Singh M, Kang DK, Lee JH, Tiwary US, Singh D, Chung WY, editors. Intelligent Human Computer Interaction. IHCI 2020. Lecture Notes in Computer Science, vol 12616. Cham: Springer; 2021:199-207.
26. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp; 2011.
27. UEQ Data Analysis Tool, Version 8. 2020. URL: https://www.ueq-online.org/ [accessed 2020-10-19]
28. Schrepp DM. All you need to know to apply the UEQ successfully in your projects. User Experience Questionnaire Handbook. 2019 Dec 31. URL: https://www.ueq-online.org/Material/Handbook.pdf [accessed 2021-03-24]
29. Horstmann KT, Buecker S, Krasko J, Kritzler S, Terwiel S. Who does or does not use the 'Corona-Warn-App' and why? Eur J Public Health 2021 Feb 01;31(1):49-51 [FREE Full text] [doi: 10.1093/eurpub/ckaa239] [Medline: 33340328]
30. Jansen-Kosterink S, Hurmuz M, den Ouden M, van Velsen L. Predictors to use mobile apps for monitoring COVID-19 symptoms and contact tracing: A survey among Dutch citizens. medRxiv. Preprint posted online June 02, 2020 [FREE Full text] [doi: 10.1101/2020.06.02.20113423]
31. Joo J, Shin MM. Resolving the tension between full utilization of contact tracing app services and user stress as an effort to control the COVID-19 pandemic. Serv Bus 2020 Sep 01;14(4):461-478. [doi: 10.1007/s11628-020-00424-7]
32. Rheault L, Musulan A. Explaining Support for COVID-19 Cell Phone Contact Tracing. SOCArXiv. Preprint posted online July 3, 2020 [FREE Full text] [doi: 10.31235/osf.io/8wcgz]
33. Thomas R, Michaleff ZA, Greenwood H, Abukmail E, Glasziou P. Concerns and Misconceptions About the Australian Government's COVIDSafe App: Cross-Sectional Survey Study. JMIR Public Health Surveill 2020 Nov 04;6(4):e23081 [FREE Full text] [doi: 10.2196/23081] [Medline: 33048826]
34. von Wyl V, Höglinger M, Sieber C, Kaufmann M, Moser A, Serra-Burriel M, et al. Drivers of Acceptance of COVID-19 Proximity Tracing Apps in Switzerland: Panel Survey Analysis. JMIR Public Health Surveill 2021 Jan 06;7(1):e25701 [FREE Full text] [doi: 10.2196/25701] [Medline: 33326411]

35. Walrave M, Waeterloos C, Ponnet K. Adoption of a Contact Tracing App for Containing COVID-19: A Health Belief Model Approach. JMIR Public Health Surveill 2020 Sep 01;6(3):e20572 [FREE Full text] [doi: 10.2196/20572] [Medline: 32755882]
36. Ilaria MN, Roussel N, Thiébaut R, Tzourio C. The French Covid-19 contact tracing app: knowledge, attitudes, beliefs and practices of students in the health domain. medRxiv. Preprint posted online November 12, 2020 [FREE Full text] [doi: 10.1101/2020.10.23.20218214]
37. Blom A, Wenz A, Cornesse C, Rettig T, Fikel M, Friedel S, et al. Barriers to the Large-Scale Adoption of a COVID-19 Contact Tracing App in Germany: Survey Study. J Med Internet Res 2021 Mar 02;23(3):e23362 [FREE Full text] [doi: 10.2196/23362] [Medline: 33577466]
38. Proszowska DK, Jansen G, de Vries PW. COVID-19 en de “intelligente lockdown" in de ogen van de burgers: deel 3: Meningen over de corona-app. Universiteit Twente. 2020 Oct. URL: https://coronapapers.nl/uploads/media_item/media_item/146/86/Rapport-Lockdown-in-de-ogen-van-burgers-deel-3-Corona-app-1603965497.pdf [accessed 2020-11-19]
39. Jacob S, Lawarée J. The adoption of contact tracing applications of COVID-19 by European governments. Policy Design and Practice 2020 Nov 28:1-15. [doi: 10.1080/25741292.2020.1850404]
40. van der Laan LN, van der Waal NE, de Wit JMS. Eindrapportage - CoronaMelder Evaluatie. Rijksoverheid. 2020 Dec 09. URL: https://tinyurl.com/xr5p6n6y [accessed 2021-03-23]
41. Mobile applications to support contact tracing in the EU’s fight against COVID-19: Common EU Toolbox for Member States. European Commission. 2020. URL: https://ec.europa.eu/health/sites/health/files/ehealth/docs/covid-19_apps_en.pdf [accessed 2020-11-19]

Abbreviations

IEC: International Electrotechnical Commission
ISO: International Organization for Standardization
PHA: public health authorities
RIVM: National Institute for Public Health and the Environment
UEQ-Dutch: Dutch User Experience Questionnaire
VWS: Ministry of Health, Welfare and Sports

Edited by G Eysenbach; submitted 11.02.21; peer-reviewed by S Hallberg, A Benis; comments to author 25.02.21; revised version received 12.03.21; accepted 15.03.21; published 26.03.21.

Please cite as:
Bente BE, van 't Klooster JWJR, Schreijer MA, Berkemeier L, van Gend JE, Slijkhuis PJH, Kelders SM, van Gemert-Pijnen JEWC
The Dutch COVID-19 Contact Tracing App (the CoronaMelder): Usability Study
JMIR Form Res 2021;5(3):e27882
URL: https://formative.jmir.org/2021/3/e27882
doi:10.2196/27882
PMID:33724198

Original Paper

Global Collaborative Social Network (Share4Rare) to Promote Citizen Science in Rare Disease Research: Platform Development Study

Roxana Radu[1], MS; Sara Hernández-Ortega[2], PhD; Oriol Borrega[3], BA; Avril Palmeri[4], BBA; Dimitrios Athanasiou[5], MBA; Nicholas Brooke[1], MBA; Inma Chapí[3], MA; Anaïs Le Corvec[6], BA; Michela Guglieri[4], MD, PhD; Alexandre Perera-Lluna[7], PhD; Jon Garrido-Aguirre[7], MSc; Bettina Ryll[8], MD, PhD; Begonya Nafria Escalera[2], BA

[1]The Synergist, Brussels, Belgium

[2]Sant Joan de Déu Research Institute, Esplugues de Llobregat, Spain

[3]Omada Interactiva SL, Barcelona, Spain

[4]John Walton Muscular Dystrophy Research Centre, Newcastle University, Newcastle, United Kingdom

[5]Stichting United Parent Projects Muscular Dystrophy, World Duchenne Organization, Amsterdam, Netherlands

[6]ClicLab, Barcelona, Spain

[7]Center for Biomedical Engineering Research, Universitat Politècnica de Catalunya, Barcelona, Spain

[8]Melanoma Patient Network Europe, Uppsala, Sweden

Corresponding Author:
Begonya Nafria Escalera, BA
Sant Joan de Déu Research Institute
Santa Rosa 39-57
Esplugues de Llobregat, 08950
Spain
Phone: 34 669167771
Email: bnafria@sjdhospitalbarcelona.org

Abstract

Background: Rare disease communities are spread around the globe and segmented by their condition. Little research has been performed on the majority of rare diseases. Most patients who are affected by a rare disease have no research on their condition because of a lack of knowledge due to absence of common groups in the research community.

Objective: We aimed to develop a safe and secure community of rare disease patients, without geographic or language barriers, to promote research.

Methods: Cocreation design methodology was applied to build Share4Rare, with consultation and input through workshops from a variety of stakeholders (patients, caregivers, clinicians, and researchers).

Results: The workshops allowed us to develop a layered version of the platform based on educating patients and caregivers with publicly accessible information, a secure community for the patients and caregivers, and a research section with the purpose of collecting patient information for analysis, which was the core and final value of the platform.

Conclusions: Rare disease research requires global collaboration in which patients and caregivers have key roles. Collective intelligence methods implemented in digital platforms reduce geographic and language boundaries and involve patients in a unique and universal project. Their contributions are essential to increase the amount of scientific knowledge that experts have on rare diseases. Share4Rare has been designed as a global platform to facilitate the donation of clinical information to foster research that matters to patients with rare conditions. The codesign methods with patients have been essential to create a patient-centric design.

(JMIR Form Res 2021;5(3):e22695) doi:10.2196/22695

KEYWORDS

Share4Rare; rare disease; citizen science; participatory medicine; natural history; genotype; phenotype

Introduction

Rare diseases are characterized by their low prevalences, and often, for being chronic, debilitating, or life-threatening conditions [1]. Despite this, over 400 million people worldwide are estimated to be affected by a rare disease [2]. In Europe, a disease is defined as rare when it affects less than 1 in 2000 citizens, accounting for an estimated 30 million people in the European Union [3]. Between 6000 and 8000 different rare conditions have been identified to date [4].

Most rare diseases (80%) have a genetic origin, and between 50% and 75% have an onset at birth or during the first years of life [5], with a wide variety of symptoms and signs that vary, not only from disease to disease, but also from patient to patient. Indeed, relatively common symptoms can hide underlying rare diseases, leading to misdiagnosis. It is estimated that approximately 25% of patients with a rare disease wait between 5 and 30 years to obtain a diagnosis (if they ever obtain one), and it is estimated that, during that time, 40% receive an incorrect diagnosis [6]. This heterogeneity in diseases and symptoms, along with geographic dispersion, makes it more difficult to access reliable information (if it exists) or to gather a significant number of patients for research to better understand the cause of each condition. It is also more difficult to describe the natural history of the disease in order to facilitate patient care and treatment. Furthermore, numerous countries do not have mandatory registries of patients, making it almost impossible to know the epidemiology of the diseases and the incidence in their populations.

From a medical point of view, ultrarare diseases, also called *orphan diseases*, often do not have any treatment options. This leads to use of off-label drugs without scientific evidence for 73% of cases, and wasteful or harmful treatment may occur [7]. By providing professionals with information about frequency, signs, symptoms, age of onset, and disease progression, potential outcome measures can be identified for later therapeutic trials. Once the variability and rate of progression of a specific disease-related sign or symptom are known, the information can be utilized as a control in the design of a clinical trial. The natural history description of a disease helps clinicians with early diagnosis and aids in the counseling of patients and families.

The internet has increased public access to health information and transformed patient behaviors. New services are offered in the health field based on the power of the interconnection of this network. For example, an overwhelming majority of parents (89%) accessed the internet before meeting with genetic providers at metabolic treatment centers [8]. Another interesting statistic shows that 18% of users connected to others with their disease through the internet [9]. However, few clinicians recommend websites to parents at the time of diagnosis as there are no trusted sources of information available for rare disease patients and caregivers or virtual meeting spaces.

Participatory health is a growing area, with individuals using health social networks, crowdsourced studies [10], smartphone health apps, and personal health records to organize their own research studies through health collaboration communities created especially for the purpose of self-experimentation and the investigation of health-related concerns.

In this landscape, Share4Rare is a project supported by the European Commission through the Horizon 2020 program (GA780262). The project aims to (1) develop an online platform to collect information about patients affected by rare diseases in order to build a global community, (2) connect people and patient organizations and communities, and (3) promote research by collecting clinical information to describe the natural history of rare diseases.

The aim of this paper is to describe the methodology that was followed in the Share4Rare project to build a collaborative social network with the 3 previous objectives—educate, connect, and foster research based on the donation of medical information from the users of the platform. To cocreate the platform, several co-design sessions were organized with the main stakeholders: patients, patient representatives, clinicians, and researchers.

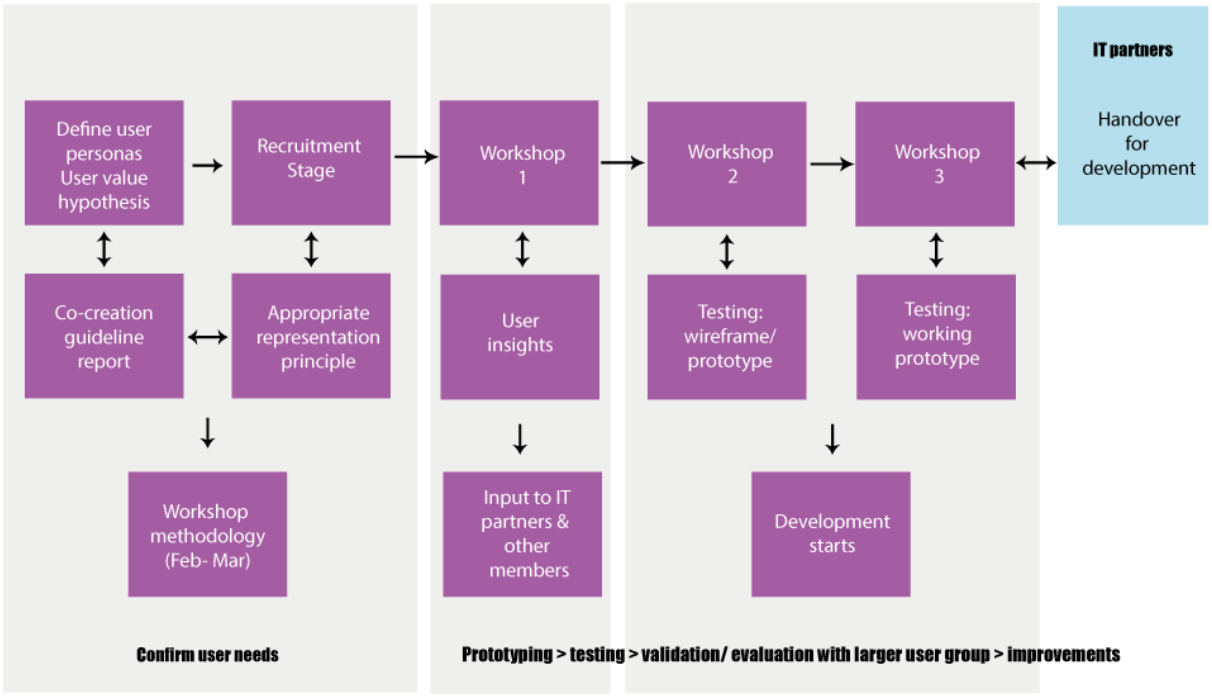
Altogether, the methods and analysis of the sessions were used to build our platform based on (1) including a patient-centered perspective in the design and development of research projects about rare diseases through clinical information donation, (2) promoting a supportive collaborative social network to enhance synergies between families so they can interact and improve disease management, (3) connecting patients to their respective patient organizations to grow and empower all rare disease communities, (4) generating a sense of community to help reduce the isolation of and stress levels in families living with a rare disease and to improve their quality of life, and (5) providing a virtual space for patients and caregivers to connect and learn from each other as well as from health care professionals.

Methods

Cocreation Process

As a participatory platform centered around the needs of patients and caregivers affected by rare diseases, Share4Rare integrated various stakeholders' perspectives in the initial design of the platform through a cocreation process. The adoption of cocreation frameworks (Figure 1) aimed to ensure that end users were able to directly influence how the platform would take shape, and thus, facilitate their joint ownership of the end product.

Figure 1. The diagram illustrates the key steps in the cocreation process, highlighting our iterative approach to the design of the platform.



Stakeholders, whose input was critical to the delivery of a useful platform and representing the target audience who could benefit from the platform, were invited to participate. The target audience included specific patients and patient groups along with the caregivers of minors, as well as clinicians and researchers. The inclusion criteria specified that participants needed to be able to travel to the workshop locations or participate via video or teleconference and be able to understand English or Spanish.

The recruitment process began with creating user profiles for approval by the organizations involved in the project. Then, an invitation was created along with an engagement plan. The plan was reviewed along with the target of users represented in different workshops. The consortium of organizations involved in the project reached out to our networks (other organizations with which we collaborate) to suggest participants and to organize all the practicalities and accessibility considerations if any were needed.

Theoretical Framework

Theoretical frameworks were used to guide the design process. The Human Centered Design approach by IDEO [11] was adopted as a framework to guide the cocreation workshops (Figure 2). This approach aims to ensure an outcome that balances meeting users' desires, offering a viable business model and technical feasibility.

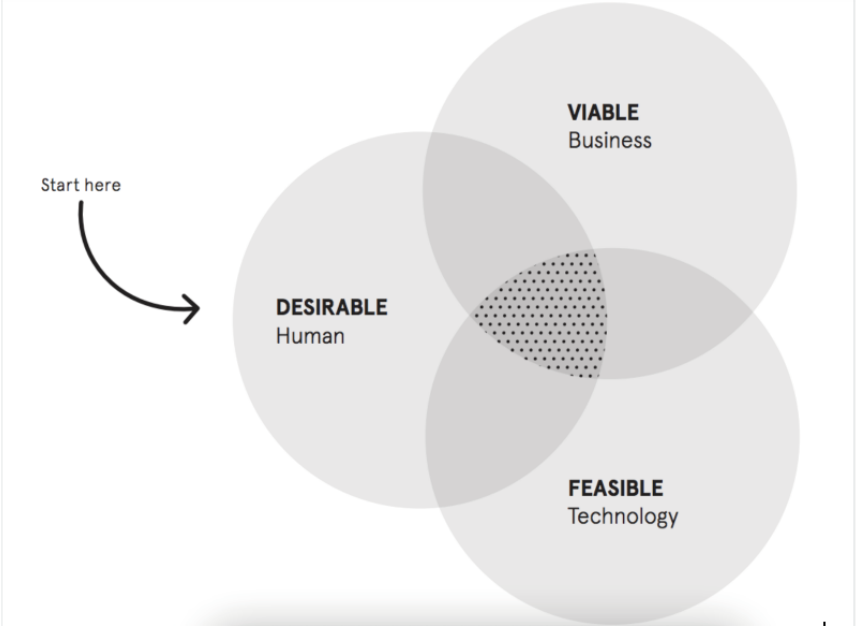
It provides a dedicated space and facilitation for cocreation activities while ensuring a balanced and diverse representation of stakeholders. It is broken down into 3 phases: Inspiration, Ideation, and Implementation:

The *Inspiration* phase is an information-gathering phase aimed at understanding users. Observing users' everyday lives and challenges provides a deep understanding of their needs. Research and analysis of other sources form a foundation to build upon.

The *Ideation* phase assesses the observations gathered in the Inspiration phase to identify possible solutions that fit with business or project needs, user desires and technical feasibilities. Prototypes are developed, shared then iterated based on the feedback received.

Finally, the cocreated prototypes are built during the *Implementation* phase; for Share4Rare, this refers to the development of the platform. Again, an iterative process allowed for multiple rounds of feedback.

Figure 2. Human Centered Design model from [9].



Preworkshop Phase

Prior to the workshops, information was used within the initial project proposal as the basis for a collaborative document describing the *personas* of the expected primary users of the platform and how these users might behave on the platform, based on their individual needs.

To validate and improve this information, an online survey was used to gather feedback initially from patient advocacy groups closest to patients and caregivers, and later, from the rest of the consortium. Members were asked to check, prioritize, and complete the user personas, user scenarios, and feature types identified. The outcomes from this first step helped us decide the activities and group interview questions for Workshop 1.

Cocreation Workshops

The workshops were structured in a successive order following the Human Centered Design approach [11]. Workshop 1, the first cocreation workshop, focused on the Inspiration phase. It aimed to understand the needs of patients, caregivers, and clinicians to ensure that the platform was designed with their needs in mind. Workshop 2, the second cocreation workshop, focused on the Ideation and Implementation phases, validating the initial conclusions drawn in Workshop 1. Workshop 3, the third cocreation workshop, focused on testing the hypotheses for the unique value proposition to enable us to confirm the main features and direction the platform development should take.

Patients and caregivers (sample size target: 10 to 12) were included in each of the 3 cocreation workshops. We specifically looked for patients without a diagnosis, newly diagnosed patients, long-term care patients, and palliative care patients. Caregivers attended the workshops on behalf of pediatric patients. Patient advocates were involved in the selection of participants, and we had more than 3 European countries represented in every workshop.

During these workshops, we employed a simplified Delphi methodology [12], which helped build consensus while ensuring all participants had an equal voice in decisions. Following the Human Centered Design [11] ensured that all participants had opportunities to contribute throughout the design process. This allowed end users to directly influence how the platform would take shape, ensuring joint ownership of the end product.

Results

Workshop 1: Methodology and Conclusions

The first session, with patients and caregivers, began with group interviews. The group was divided into smaller, diverse groups of caregivers and patients who, supported by a moderator, were asked to answer a series of questions to understand their needs.

This provided key information on main hurdles and challenges the possible users are facing in their day to day life, habits regarding the usage of digital tools, and some generic expectations from a platform similar to Share4Rare (Textbox 1).

The second session, with clinicians and researchers, also began with group interviews. Participants were asked a series of questions to understand their specific needs in the Share4Rare community. The clinicians and researchers were then asked to prioritize needs previously identified in the preworkshop phase.

In the second part of the patient and caregiver session, a game helped us to prioritize user needs previously identified in the preworkshop phase among the focus group. Each participant was given an equal amount of false money and presented with several boxes representing banks and labeled with a need type. They were asked to distribute their money according to how important they considered that need. At the end of the exercise, the amount of money in each bank was counted. A group discussion explored why each need was or was not chosen.

When prioritizing some proposed features, the patients and caregivers clearly preferred having the ability to contribute to

research, closely followed by connecting with similar people to them, as well as finding information about the diseases (Table 1).

The first iteration of the Share4Rare platform design and interaction model was drafted after both sessions.

The most needed features or feature types (Table 2) required further development and prototyping to be tested with a wider group of users.

Textbox 1. Insights from Workshop 1 with different stakeholders.

Patients and caregivers

- Almost all patients or caregivers use the internet to reach out to other people in the same situation
- Patients or caregivers would like an easy way to interact; notifications are also important to keep them up to date
- Everyone needs a space where they can find various levels of information on the disease; the language used needs to be appropriate
- It would be interesting to have information about symptoms and what they could mean (there are a lot of patients without diagnoses)
- There needs to be a space for open dialogue between patients and doctors
- Parents/patients would like to have some kind of job board adapted to their needs

Clinicians and researchers

- Ability to connect with similar experts in the field
- An exercise to map global research projects in rare diseases and identify gaps would greatly help their work
- The primary caregiver needs to be educated on rare diseases and how to refer patients to bigger diagnostic centers
- Integration with European Reference Networks is essential
- Collaborations with other institutions seem to be a main issue for the clinicians
- Having access to the latest news in the field
- Essential to have a section to help educate general public, patient associations, funders on the procedures of the research process, main needs, etc
- Tackle lack of funding in some way through the platform

Table 1. Features prioritized after Workshop 1.

Feature	Score
Contribute to research efforts	15
Connect with other people	10
Find information	6
Access latest news	5
Mental health support	4
Get support during the diagnostic journey	3
Connect with my doctor or medical facility	2
Read about other people like me	1

Table 2. Summary of the most needed features or feature types identified following Workshop 1.

Feature	Needs
Level 1 (medical, open information)	
Wiki space—easy to read medical information (easily grouped by types of conditions, diseases, pediatric or not; information to be constantly kept up to date); important—information on genetic risk and prevention measures	• Easy to read, in lay language • Easy to find relevant topics • Easy to add or modify by dedicated professionals (not by community) • Should link to external resources • Consider "verified" button or something indicating quality
Dedicated section to diagnosis (here we can have a grouping on symptoms as the patients suggested)	Same as above
List of specialized centers	• An easy to navigate map • Should have a way to visualize simple details for each center
Level 2 (communities)	
Support groups on various topics (Based on the content interaction model: patient/caregiver can decide on topics he/she is interested in)	• Very easy to quickly find relevant topics or discussions • Have tags or other system to label conversations with more than one keyword to make them easier to search for • Browsing feature or conversation feed for those who do not have a particular question but just want to see what is there • Easy to join a discussion • Quality of the conversation is important • Easy to navigate to direct messaging if they want to make the conversation private
Direct messaging option	• Safe, private, easy to use • Easy to connect with the forum
Maps with various specialists to help parents (mental health specialists, lawyers, social workers, etc)	• An easy to navigate map • Ideally, we should have a "sign up to help" call to action for the professionals mentioned above
A method to connect with expert clinicians	• Ability to view lists with vetted clinicians that could answer the needs of the various users. • The platform should provide the first point of contact and should not replace in-person clinical assessments
Level 3 (medical data)	
Search engine (transversal) for clinical trials, drugs, therapeutic methods	• Intuitive way to search for various subjects (new/ongoing clinical trials, disease based, etc)

Workshop 2: Methodology and Conclusions

The workshop began with a *gallery walk exercise* in which participants were presented a gallery of Share4Rare platform screenshots and features. They were then invited to ask questions and give their first impressions and feedback. A guided presentation of the Share4Rare platform prototype followed with a group discussion on each aspect was presented. Targeted questions explored the points raised by participants during the gallery walk in greater depth.

The group gave feedback on various specific features proposed by the information technology partner. For the storytelling feature, participants prioritized the idea of filtering stories by disease type and role, as well as the ability to add links to owned materials & resources. A key question that was raised was "how are children protected on the platform?" Parents could be happy to share lots of information, but as their child grows, they might be less happy having the information visible.

Regarding the question and answer feature, participants discussed how users should be listed in relation to others (based on interest, active topics, etc), as well as the ability to add or remove topics of interest. For the participants, it was important to note how we would manage the answers since some patients or caregivers may give answers as if they were doctors or experts that could be inaccurate or misleading. In smaller communities (such as Facebook groups), administrators step in and manage the situation. The participants also decided that only patients and caregivers should be allowed in the forum. Suggestions for content were also made, especially around the quality of life topic and language concerning children.

Additional conclusions from the group further helped guide platform design. For example, the group was against a popularity ranking in the platform and raised various questions around privacy, especially when it came to the option of navigating as an anonymous avatar (since the topics of interest can still serve as an identifier).

The undiagnosed group raised the point that it can be difficult to know where to start when navigating the platform. Users also required the ability to use filters or navigate based on disease types and symptoms and to implement the ranking of topics or content pieces based on an algorithm that takes into account topics the user has mentioned they might be interested in and navigation habits.

After the conclusion of the second cocreation workshop, participants felt that the unique value proposition of the platform was not strong enough and that there was a risk that Share4Rare would not succeed in attracting enough users.

As a result, participants decided to modify the methodology for the third workshop. The platform design was revised to articulate the platform features and unique value proposition more clearly.

Workshop 3: Methodology and Conclusions

The final cocreation workshop focused on the unique values of the Share4Rare platform and the specific feature needs of patients and caregivers. Participants were asked about their perceptions, opinions, beliefs, and attitudes regarding the Share4Rare platform prototype. A full day of interviews, role-playing creativity games, and design challenges helped to determine the subsequent development steps: (1) Welcoming activities introduced participants and facilitators. (2) Participants interacted with and tested the prototype Share4Rare platform, allowing us to identify areas for revision and validate its unique selling point. (3) Predetermined questions were asked to encourage participants to share their opinions and ideas as well as listen to and engage with the opinions and ideas of others in a small and safe group setting. (4) A cocreation activity, in which participants built their own version of the Share4Rare platform, brought participants into the design process by rapidly prototyping their own solutions to the problems. (5) Clinicians and researchers received targeted surveys aimed at validating the platform's development and unique value proposition.

The workshop provided some key information that helped prioritize key aspects—from feature creation to the value proposition and event content creation.

We realized that it was important to deliver the right information at the right time to the user, as a patient or caregiver may not feel comfortable getting in contact with people or receiving information from a later stage of the disease. That is why the key aspect was the ability to find other people with similar problems, someone to talk to or to help them understand the condition and procedures. Users wanted to know about the cause and the progression of the disease when there is no treatment. Each disease is different; therefore, they would like to see specific filters for each disease and other filters such as age or level of clinical understanding. Anonymous navigation—the ability to "watch from a distance"—was mentioned many times. The future users mentioned many times that they need to know what is located close to them in terms of facilities and expertise (keeping in mind that there are different types of experts, such as patient experts). We also understood from the discussions that we would need different levels of access to allow the users decide how much they share and how much they interact with other people. Content-wise, patients and caregivers mentioned the need to discuss comorbidities, mental health, physical therapy, palliative care, and other general topics.

The patients and caregivers were given a survey and asked to prioritize various content topics and subtopics previously agreed upon by the consortium. Keeping in mind that this type of feedback is usually biased as users prioritize according to their needs, we reorganized the topic and subtopics to reflect prioritization (Textbox 2).

The participants also had the opportunity to add topics that they believed were missing, which included symptom control; database with various specialists (neurology, orthopedics, cardiology, respiratory); standards of care; advocacy; legal issues in trials; end-of-life care—advanced care planning (resuscitations, wills, funeral planning), power of attorney, and capacity; independent living, equipment, and professional caregivers; complementary nutrition (eg, supplements); disability models; and existing tools (for example, integration with Orphanet, European Reference Networks, and local tools).

Textbox 2. Topic prioritization.

1. Understand your disease
- Find your rare disease
- Reference centers
- Rare diseases in facts
- Undiagnosed diseases
- Genetic counseling
- Genetics
- Basic research

2. Treatments
- Clinical trials
- Medicines
- Psychology
- Physiotherapy
- Orthopedics
- Nutrition
- Palliative care
- Speech therapy

3. Quality of life
- Emotional support
- Education
- Patient associations
- Caregivers
- Accessibility and disability issues
- Leisure and sport

4. Legal issues
- Grants and subsidies
- Ethics
- Regulations

Cocreation Beyond the Workshops

The workshops provided significant and meaningful insights into the needs and requirements of the users, but further collaborative work was needed with the consortium. First, to advance the discussions around the unique value proposition of the platform, and second, to prioritize the extensive feature list which came out of the user cocreation activities.

Consortium members identified and ranked the primary value propositions of the platform in an online survey based on the outcomes of the previous cocreation activities (Textbox 3). A decision-making matrix assisted the consortium in debating the platform's final development plan (Table 3).

Textbox 3. Value proposition survey results, answering the question "how will we distinguish ourselves from other solutions so people actually use Share4Rare?"

Unique value proposition

1. One network showing the 360-degree picture, not just the medical perspective
2. Get targeted, quality information faster and using fresh, innovative delivery from peers and/or expert people
3. Capture common pathways/experiences to gain time and better management of disease
4. Give an active role to patient organizations to support their members
5. Unique connection point between patients and health care professionals
6. Provide practical support to health care professionals across rare diseases, including access to medical information generated by expert doctors and updated with real data from patients.
7. Easier for patients and caregivers to access or request information that can allow a second opinion

Table 3. Top features and feature types—which features will make the unique value proposition a reality?

Rank	Feature type	Top ranked features (top 50% within each type)
1	Connect patients, caregivers and clinicians through data	1. Donate clinical data to facilitate the generation of new and qualified knowledge about the disease
		2. Check information that can facilitate the access to a second opinion using clinical data
2	Providing education and support	1. Content reviewed by an editorial board/ volunteering community management team will receive a "trusted source" stamp. All other content will still appear in the platform but with a quality warning (eg, "This information has not been verified yet")
		2. Filter the search per disease type
		3. Access medical chapters—books with relevant information about diseases
		4. Search engine across all content types (medical information, forums, questions, etc)
		5. Upload information and materials on diseases
		6. Search for a symptom
		7. Filter by language
		8. Filter per disease stage (certain diseases only)
3	Find a mentor who can support the experience with the disease	1. Find a local disease ambassador, patient advocate or mentor
		2. Ask for specific information that can worry them and that can be provided or curated by the mentor
		3. Mentors have a dashboard to easily manage mentorship tasks
4	Community support	1. Forum feature
		2. Private forums, only accessible by invitation
		3. User can follow tags and forum threads, appearing on the user's dashboard
		4. Allow users to contact through private messaging inside the platform
5	Resources for disease management & support	1. Navigate a world map & find health care facilities (diagnostic centers, treatment centers, etc)
		2. Navigate a world map to find various experts. Expertise filters would help navigate the map
		3. Find other users by using profile filters (caregiver, patient, doctors...), interest or geographic filters
6	Interaction with a health care professional	1. Health care professionals will be able to share their expertise with other clinicians that can have patients from the two pilot groups of conditions
		2. Health care professionals will be able to sign up in the platform and volunteer to support the rare disease community

Discussion

Analysis of the Co-design Findings

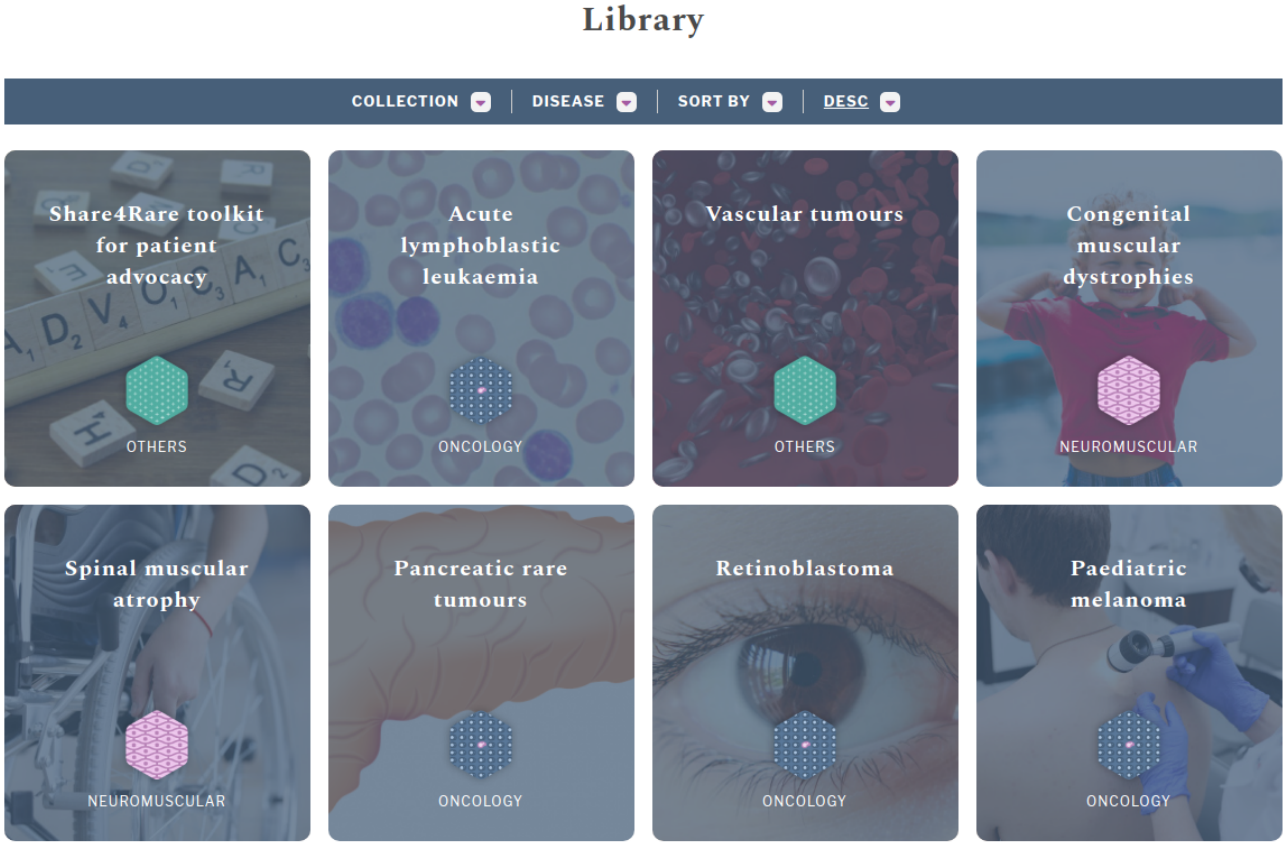
After these analyses and cocreation workshops were completed, a proposal of the platform was developed: (1) Education layer: Existing curated content will be linked to the platform. At the same time, for diseases where a minimum of information is not fully covered, new curated material will be developed in order to reach all of the diseases included in the pilot of the project. (2) Share layer: A sense of community will be created by a network of interactions that will allow patients to connect to peers or clinicians based on the answers they give through different questionnaires. These data will be collected, analyzed through machine learning methods and the network around the user will be updated daily, allowing him or her to adjust and find the correct person to contact in each case. (3) Research layer: This is the gold value of the platform. Data donation will be accomplished through several questionnaires. In the case of adult patients or relatives (legal guardians) of a pediatric patient in the pilot disease areas, the number of questionnaires will be aligned with the complexity of the disorder.

All of this information was connected and a smooth design has been created in order to facilitate the navigation of users through the platform and encourage them to return.

Education Layer

The education layer will be a source of medical content for patients and caregivers within the rare disease community. During the platform development period, the platform will pilot over 3 groups of diseases (pediatric rare tumors, neuromuscular diseases, and undiagnosed patients), content for some of these diseases will be presented in the platform and offered to users. This material, as well as blogs and toolkits, will be publicly available (Figure 3).

Figure 3. Screenshot showing some of the public medical materials available on Share4Rare.



Reciprocity and Digital Karma

The basic functions of the community and research layers are based on the principles of reciprocity and digital karma, whose goals are to promote active and voluntary participation on the part of the users and to reward them for it.

The principle of reciprocity governs the relationship between each user and the rest of the platform and is bidirectional. In one direction, all contributions will always entail a reward (for instance, if a user completes a questionnaire, the user will gain access to enriched information related to it). In the reverse direction, to gain a benefit from the platform it is necessary to actively contribute to it (for instance, researchers who wants to register a questionnaire for their research must commit to being available to communicate actively with the rest of the users).

The principle of digital karma governs the profile of each user and their level of access to the functionalities that the platform implements. Essentially, the profile of each of the users of the platform is determined by their actions within Share4Rare: a user who contributes more information or who is more active in the community will see that this behavior is reflected on their profile and will be able to benefit from more functionalities.

The Core of the Research Layer: Questionnaires

Medical and psychosocial questionnaires are central to Share4Rare. It is through them that the project will gather data, ensure the engagement of users and, most importantly, be able to generate new knowledge about rare diseases. The questionnaires will be designed by clinicians and other experts with expertise in the diseases covered by Share4Rare and will be web-based (on the platform). A subset of the answers given will be analyzed using statistical methods in an external server with the help of tools developed by a team of data scientists.

It is extremely important to obtain accurate and reliable information. Because of this, questionnaires use validation mechanisms for what is required, allowed values, minimum or maximum values and allowed formats for each question. Some of the questionnaires will be specific to a disease, while others may be applicable to several diseases or related to specific symptomatology.

Questionnaires will be provided to users (either patients or caregivers) progressively, so they can have control over which information they are sharing at every step and can decide whether to share it or not. As users advance through the questionnaires, they will have increasing feedback from the platform: (1) Each completed questionnaire will allow access to up-to-date comparative information about the answers that the rest of the community has given to the questionnaire and other topics related to it. (2) Each completed questionnaire will disclose more functionalities for the user (for instance, by giving them access to ask questions to the community or to open private messages with other users), and (3)The relevant data from each questionnaire will be incorporated to the user's profile, which will increase their karma level and will allow the platform to better characterize the user in order to be able to present information of greater relevance.

All tools that the platform uses to collect data from the users (eg, registration forms, medical questionnaires, etc) included in the private environment of the platform require access using a user and password. The overall data collection and custody process of the platform is described in the Data management plan, a mandatory document in all Horizon 2020 European projects that is based the General Data Protection Regulation (mandatory and common across all European countries).

The Share4Rare platform and every pilot research project that will be performed until the end of the European Commission Grant have the approval of the Ethics Committee of *Sant Joan de Déu* Research Foundation. It is mandatory for all users of the platform to sign the informed consent document that regulates the use of clinical data and medical information for research purposes. If any use of these data occur in the future, the platform will ask for the re-consent of users that will be affected by this new use. Every year an external audit will be performed to oversee the use, gathering, and security of the data in Share4Rare.

Conclusions

1. The platform aims to gather meaningful data from the patients who participate. Once duly treated and stored to guarantee patient anonymity and to ensure its statistical and clinical validity, said data will become part of a highly valuable repository that will be available for scientists and researchers conducting investigations in the field of rare diseases.

2. Share4Rare will develop a procedural framework that includes a set of validated tools that will allow researchers to gather much-needed data in a standardized manner, preserving its structural integrity, allowing for cross-comparison and cross-reference between studies and promoting the engagement of patients.

In parallel, the solutions that Share4Rare has devised to implement these 2 goals in the platform will allow clinicians and researchers to create a patient registry and will allow patients to be part of a community that makes it easier to connect with other people in their situation. At the same time, the platform will supply its participants with dynamic state-of-the-art content about the conditions, research being conducted, and other topics of interest to encourage active user involvement.

Acknowledgments

This work has been funded by the European Research Council under the European Union Horizon 2020 research and innovation program (grant agreement number H2020-780262-SHARE4RARE). BN acknowledges the support from *Sant Joan de Déu* Barcelona Children's Hospital.

Conflicts of Interest

None declared.

References

1. Baldovino S, Moliner AM, Taruscio D, Daina E, Roccatello D. Rare diseases in Europe: from a wide to a local perspective. Isr Med Assoc J 2016 Jun;18(6):359-363 [FREE Full text] [Medline: 27468531]
2. Jia J, Shi T. Towards efficiency in rare disease research: what is distinctive and important? Sci China Life Sci 2017 Jul;60(7):686-691. [doi: 10.1007/s11427-017-9099-3] [Medline: 28639105]
3. Orphan designation: overview. European Medicines Agency. URL: https://www.ema.europa.eu/en/human-regulatory/overview/

orphan-designation-overview#:~:text=About%2030%20million%20people%20living,medicines'%20in%20the%20medical%20world [accessed 2021-02-24]
4. Rare diseases. European Commission. URL: https://ec.europa.eu/info/publications/rare-diseases_en [accessed 2021-03-03]
5. Bavisetty S, Grody WW, Yazdani S. Emergence of pediatric rare diseases: review of present policies and opportunities for improvement. Rare Dis 2013 Oct 20;1(1):e23579 [FREE Full text] [doi: 10.4161/rdis.23579] [Medline: 25002987]
6. Black N, Martineau F, Manacorda T. Diagnostic odyssey for rare diseases: exploration of potential indicators. Policy Innovation Research Unit. London, United Kingdom: Department of Health Services Research & Policy London School of Hygiene and Tropical Medicine; 2015. URL: https://piru.ac.uk/assets/files/Rare%20diseases%20Final%20report.pdf [accessed 2021-03-03]
7. Frost J, Okun S, Vaughan T, Heywood J, Wicks P. Patient-reported outcomes as a source of evidence in off-label prescribing: analysis of data from PatientsLikeMe. J Med Internet Res 2011 Jan 21;13(1):e6 [FREE Full text] [doi: 10.2196/jmir.1643] [Medline: 21252034]
8. DeLuca JM, Kearney MH, Norton SA, Arnold GL. Internet use by parents of infants with positive newborn screens. J Inherit Metab Dis 2012 Sep 2;35(5):879-884. [doi: 10.1007/s10545-011-9449-7] [Medline: 22297410]
9. Fox S. The social life of health information, 2011. Pew Research Center. 2011 May 11. URL: https://www.pewresearch.org/internet/2011/05/12/the-social-life-of-health-information-2011/ [accessed 2021-02-21]
10. Swan M. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. J Med Internet Res 2012 Mar 07;14(2):e46 [FREE Full text] [doi: 10.2196/jmir.1988] [Medline: 22397809]
11. Field guide to human centered design. IDEO. 2015. URL: https://d1r3w4d5z5a88i.cloudfront.net/assets/guide/Field%20Guide%20to%20Human-Centered%20Design_IDEOorg_English-0f60d33bce6b870e7d80f9cc1642c8e7.pdf [accessed 2021-02-21]
12. The delphi method: techniques and applications. Wayback Machine. 1975. URL: https://web.archive.org/web/20080520015240/http:/is.njit.edu/pubs/delphibook/ [accessed 2021-02-24]

Edited by G Eysenbach; submitted 21.07.20; peer-reviewed by S Steingrimsson, A Thunström; comments to author 30.09.20; revised version received 22.11.20; accepted 20.12.20; published 29.03.21.

Please cite as:
Radu R, Hernández-Ortega S, Borrega O, Palmeri A, Athanasiou D, Brooke N, Chapí I, Le Corvec A, Guglieri M, Perera-Lluna A, Garrido-Aguirre J, Ryll B, Nafria Escalera B
Global Collaborative Social Network (Share4Rare) to Promote Citizen Science in Rare Disease Research: Platform Development Study
JMIR Form Res 2021;5(3):e22695
URL: https://formative.jmir.org/2021/3/e22695
doi:10.2196/22695
PMID:33779572

Original Paper

Online Tool for the Assessment of the Burden of COVID-19 in Patients: Development Study

Esther M J van Noort[1*], MSc; Danny Claessens[2*], MSc; Catharina C Moor[3*], MSc, PhD; Carlijn A L Van Den Berg[4], BSc; Marise J Kasteleyn[1*], PhD; Johannes C C M in 't Veen[4,5*], PhD; Onno C P Van Schayck[2*], Prof Dr; Niels H Chavannes[1*], Prof Dr

[1]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, Netherlands

[2]Department of Family Medicine, Care and Public Health Research Institute School for Public Health and Primary Care, Maastricht University Medical Centre, Maastricht, Netherlands

[3]Department of Respiratory Medicine, Erasmus Medical Center, Rotterdam, Netherlands

[4]Erasmus Medical Center, Rotterdam, Netherlands

[5]Department of Pulmonology, Franciscus Gasthuis and Vlietland Hospital, Rotterdam, Netherlands

[*]these authors contributed equally

Corresponding Author:
Esther M J van Noort, MSc
Department of Public Health and Primary Care
Leiden University Medical Center
Albinusdreef 2
Leiden, 2333 ZA
Netherlands
Phone: 31 162520571
Email: e.m.j.van_noort@lumc.nl

Abstract

Background: The impact of COVID-19 has been felt worldwide, yet we are still unsure about its full impact. One of the gaps in our current knowledge relates to the long-term mental and physical impact of the infection on affected individuals. The COVID-19 pandemic hit the Netherlands at the end of February 2020, resulting in over 900,000 people testing positive for the virus, over 24,000 hospitalizations, and over 13,000 deaths by the end of January 2021. Although many patients recover from the acute phase of the disease, experience with other virus outbreaks has raised concerns regarding possible late sequelae of the infection.

Objective: This study aims to develop an online tool to assess the long-term burden of COVID-19 in patients.

Methods: In this paper, we describe the process of development, assessment, programming, implementation, and use of this new tool: the assessment of burden of COVID-19 (ABCoV) tool. This new tool is based on the well-validated assessment of burden of chronic obstructive pulmonary disease tool.

Results: As of January 2021, the new ABCoV tool has been used in an online patient platform by more than 2100 self-registered patients and another 400 patients in a hospital setting, resulting in over 2500 patients. These patients have submitted the ABCoV questionnaire 3926 times. Among the self-registered patients who agreed to have their data analyzed (n=1898), the number of females was high (n=1153, 60.7%), many were medically diagnosed with COVID-19 (n=892, 47.0%), and many were relatively young with only 7.4% (n=141) being older than 60 years. Of all patients that actually used the tool (n=1517), almost one-quarter (n=356, 23.5%) used the tool twice, and only a small group (n=76, 5.0%) used the tool 6 times.

Conclusions: This new ABCoV tool has been broadly and repeatedly used, and may provide insight into the perceived burden of disease, provide direction for personalized aftercare for people post COVID-19, and help us to be prepared for possible future recurrences.

(JMIR Form Res 2021;5(3):e22603) doi:10.2196/22603

KEYWORDS

COVID-19; patient-reported outcomes; ABCoV tool; monitoring; patient outcome; long-term impact; tool; assessment; online patient platform

Introduction

The COVID-19 pandemic caused by the new coronavirus SARS-CoV-2 swept through the Netherlands from the end of February 2020 and caused over 24,000 hospitalizations and over 13,000 deaths by the end of January 2021 [1]. Although many recovered from the acute infection, damage to the lungparenchyma (portion of the lungs involved in gas exchange) was observed in computed tomography scans of patients who were hospitalized [2], with a subsequent risk of long-term lung damage. Experiences with other coronaviruses also raised a serious concern for long-term *sequelae*. As an example, the Q fever epidemic, which had 4026 cases (in the period 2007-2010) [3], had an extensive aftermath: 20% of patients with acute symptoms subsequently experienced fatigue long after the initial infection was resolved [3]. Adequate follow-up of patients with COVID-19 may reduce the long-term consequences by means of early detection and symptom management.

Monitoring patients who have had a COVID-19 infection is therefore pivotal. In early March 2020, the Lung Foundation Netherlands (a Dutch patient advocacy group) became aware of the need for a better understanding of COVID-19 in the public [4,5]. During the peak of the pandemic (first wave), there was only a limited number of tests for COVID-19 in the Netherlands. Hence, there was a rapid growing requirement for information in those that experienced COVID-19 symptoms but were never tested or medically diagnosed.

The Lung Foundation Netherlands offers a help desk and an online forum for the public for all lung-related questions [6]. In the beginning of the pandemic, Lung Foundation Netherlands was confronted with all kinds of concerns and questions related to COVID-19. To handle these concerns, Lung Foundation Netherlands decided on a structured approach. Since Lung Foundation Netherlands is involved in scientific research [7], it was a logical decision to exploit patient-reported outcome measurements (PROMs) for a better understanding of the symptoms and long-term impacts of COVID-19. However, no tools to report on symptoms and long-term effects of COVID-19 were available yet. On the other hand, PROMs to assess the patient-experienced burden of disease do exist for other chronic conditions. Therefore, we decided upon the development of a COVID-19–oriented tool.

Methods

Team

A team of leading medical, technical, and process experts was formed, consisting of the chief executive officer (CEO) of the Lung Foundation Netherlands, the CEO of the Dutch Lung Alliance, experts from the Care and Public Health Research Institute (CAPHRI), pulmonologists from a training hospital (Franciscus Gasthuis and Vlietland) and an academic hospital (Erasmus MC), and the CEO of an eHealth company (Curavista). Because patients infected by COVID-19 are primarily at risk of developing lung damage, the selection of potential tools was narrowed down to tools assessing pulmonary symptoms. It was also considered best to adapt an existing, validated, and well-known tool to be launched as quickly as possible.

The Assessment of Burden of Chronic Obstructive Pulmonary Disease Tool

The assessment of burden of chronic obstructive pulmonary disease (ABC) tool, used in the monitoring and care for people with chronic obstructive pulmonary disease (COPD) [8-10], was selected as the best-suited tool. The ABC tool was preferred over other tools such as the COPD assessment tool or the COPD control questionnaire for four reasons. First, in the management of chronic conditions such as COPD, there is a paradigm shift from doctor-driven care to patient-centered integrated care with active involvement of and self-management by the patient. The original ABC tool was designed to be used in this transition of facilitating self-management support and shared decision making. As such, it considers symptoms and lifestyle. Second, the ABC tool offers a strong graphical visualization (the status per symptom being scored in *colored balloons*). The balloons are easy to understand, enhancing the long-term patient participation necessary to collect individual long-term data. Third, the ABC tool had previously been adapted for other conditions such as asthma, diabetes, and heart failure [11]. Finally, the tool is already widely used in the Netherlands by both general practitioners (GPs) and hospital specialists in monitoring patients with COPD, and the tool is integrated in the guidelines for regular GP care of patients with COPD [12].

The ABC tool, developed in 2014, measures the integrated health status of an individual patient with COPD [9]. The self-administered questionnaire consists of 14 statements that evaluate the burden of COPD experienced by patients in five domains (symptoms, functional state, mental state, emotions, and fatigue) and with some objective parameters. An algorithmic computer program visualizes outcomes and provides treatment advice and an index score for future health care costs [10]. Each domain is visualized using balloons: a high green balloon indicates a good score on a particular item, while a low red balloon indicates difficulties on that item (Figure 1).

Figure 1. Visualization of the assessment of burden of chronic obstructive pulmonary disease tool.

The previous score appears as a gray balloon and indicates improvement or deterioration. The ABC tool has been validated and proven to be effective and reliable for people with COPD, is proven to be effective in improving quality of life, is perceived as easy to use by both health care providers and patients [8], and is adaptable for other chronic conditions [8-13].

To adjust the ABC tool for use in post–COVID-19 cases, a focused literature search was performed in early April 2020 by the CAPHRI Institute of Maastricht University. This search covered symptoms, complaints, and burden of disease in cases of COVID-19 reported earlier and in reports of the previous coronavirus outbreaks of severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS).

The most prominent long-term effects of SARS and MERS are chronic fatigue and lung fibrosis, resulting in shortness of breath, dry cough, and decreased physical and mental health [14-19]. Patients with COVID-19 show similar symptoms as well as fever, headache, and chest pain [20-23]. The ABC tool includes all of those symptoms. The aforementioned nonpulmonary symptoms were added to the new tool.

Some patients (29%) with COVID-19 are admitted to the intensive care unit due to acute respiratory distress syndrome (ARDS) [22]. Quite often these patients show signs of a restrictive lung disease (25%) and, in fewer cases (4.5%), obstructive lung disease [21]. The ABC tool includes the most important symptoms of obstructive and restrictive lung disease (Table 1). Furthermore, ARDS can lead to decreased physical and mental health, decreased attention and concentration, and muscle weakness [24]. This warranted the inclusion of symptoms related to postintensive care syndrome and ARDS [24,25]. Physical and mental health are addressed in the ABC tool, but decreased attention and concentration and muscle weakness are not. Since only a subgroup of patients developed ARDS, the expert group considered these three additional symptoms out of scope, and these items were not integrated in the assessment of burden of COVID-19 (ABCoV) tool.

Table 1. Modification of symptoms from the ABC tool to the ABCoV tool.

Symptoms	ABC[a] tool	ABCoV[b] tool
Smoking	Included	Included
Exacerbation	Included	Excluded
Shortness of breath	Included	Included
BMI	Included	Included
FEV_1[c]	Included	Included
Physical exercise	Included	Included
Physical well-being: walking, stairs, dishwashing	Included	Included
Symptoms like cough or phlegm production	Included	Included
Mental health due to lung problems	Included	Included
Fatigue	Included	Included
Social well-being, engagement in social activities	Included	Included
Chest pain	—[d]	Added
Dizziness	—	Added
Headache	—	Added
PTSD[e] screening	—	Added
Open text	—	Added

[a]ABC: assessment of burden of chronic obstructive pulmonary disease.

[b]ABCoV: assessment of burden of COVID-19.

[c]FEV_1: forced expiratory volume in the first second of expiration.

[d]Item was not in the original ABC tool.

[e]PTSD: posttraumatic stress disorder.

To gain an understanding of the impact of COVID-19 on mental health [24], an additional screening question was added to determine whether patients had traumatic experiences related to COVID-19 infection. If indicated so, the first five questions from the Global Psychotrauma Screen (GPS) [26] are presented to evaluate the risk of developing posttraumatic stress disorder. If not, all questions from the GPS are left out.

Because of the limited information on the long-term sequelae of COVID-19 at that time, it was decided to be overly inclusive. Therefore, only one item (exacerbation) from the original ABC tool for COPD was excluded, as it was not described at all in the literature on COVID-19 at the time of development. The lifestyle items were included as well because lifestyle seemed to be an important risk factor for hospitalization [27] and in influencing outcomes [28]. We included all items from the literature search and additionally offered an *open text field* for all other possible symptoms. This offers the best possibility to get better insight in the incidence of the different symptoms in a new disease like COVID-19. A full overview of the original ABC tool and new ABCoV tool is presented in Table 1. Multimedia Appendix 1 shows the questionnaire, known as the ABCoV tool, that is currently in use.

Algorithm

Patient responses to the ABCoV tool are translated into balloons with a certain height and color (Figure 2). The height and color of each balloon is determined by the answers to questions. Answers that point toward *no burden of disease* generate high-floating green balloons, and answers that point toward *heavy burden of disease* generate a red balloon down to the ground. Orange, yellow, and light green balloons represent intermediate answers to *burden of disease*. The original algorithm for COPD was unaltered for every item except BMI; optimal BMI ranges differentiate for people with COPD from that of the general population. For the ABCoV tool, the optimal BMI range of the general population was used (<18.5 underweight, 18.5-25 normal weight, 25-30 overweight, >35 obese). A 7-point Likert scale was used for all items in the ABCoV tool other than risk factors. Newly included questions in the ABCoV tool are scored on a visual analog scale ranging from 0 to 10. The ABC tool generates treatment recommendations and an index score assigning the overall burden to one of three categories: low, medium, or high. More research is needed to implement both of these in the ABCoV tool, so they have been omitted for now. The expert group decided not to impose a specific frequency of use or time window since the use of the ABCoV tool is patient-driven. Patients can fill out the ABCoV tool once a day, and there are no reminders.

Figure 2. Visualization of the ABCoV tool. ABCoV: assessment of burden of COVID-19; FEV1: forced expiratory volume in the first second of expiration.

Online eHealth Platform

The expert group chose to roll out the tool using the Curavista.health platform, a certified modular platform (NEN7510, ISO27001, CE class I MDD, Health Insurance Portability and Accountability Act compliant), which has the ABC tool already in place and is used by patients, GP practices, and hospitals. The platform is available in six languages, and other languages are being implemented. As such, the tool is scalable.

The ABCoV tool was incorporated, the algorithm validated and tested, and its content cleared by the CAPHRI institute. The preproduction environment was tested by 3 patients. They suggested to add an explanation that describes the purpose of the ABCoV tool and why this was adopted.

Enrollment

During the first months of the pandemic, only limited testing was available for COVID-19 in the Netherlands. The Lung Foundation Netherlands received many questions from people who experienced COVID-19 symptoms, with or without having been tested or medically diagnosed [4]. Both may benefit from the ABCoV tool to monitor their progression and to get tested if their symptoms worsen or do not subside. Therefore, it was decided to enable registration for everyone for the tool and self-monitoring via the website *Coronalungsquare* [29]. Registered persons can invite their physician to join them online. A *medical route* is also available by which doctors can invite their patients to join and monitor their post–COVID-19 experience using the ABCoV tool online.

Results

The ABCoV tool was launched on May 7, 2020. A total of 2162 people (each identified by a unique email address) had downloaded the app and self-registered by January 20, 2021. A total of 1898 (87.8%) people gave permission to analyze their data. The majority of participants were female and between 18-60 years of age (Table 2). A subtotal of 1690 (78.1%) people submitted the first screening questions on how and when they were diagnosed with COVID-19. A total of 892 (47.0%) people responded that the diagnosis was medically confirmed either by testing (n=628) or by a medical doctor without testing (n=262). Only 2 people could not recollect how they were diagnosed. A total of 472 people did not answer this question (Table 2). Additionally, 456 patients, diagnosed by pulmonologists at a large teaching hospital (Franciscus Gasthuis and Vlietland) or at one of three academic hospitals (Erasmus

Medical Center, Leiden University Medical Center, or Amsterdam University Medical Center), were enrolled for the use of the ABCoV tool, and more GPs and hospitals will do so in the coming months.

Participants can log in with their personal account and submit the ABCoV tool again. Participants do not receive reminders. The majority of the participants submitted the questionnaire once. After submitting the ABCoV 19 tool once, 23.5% (356/1517) of participants submitted the ABCoV tool again (see Table 3).

Table 2. Profile of registered users (self-registration).

Demographics	Registered users (N=2162), n (%)
Permission to analyze data	1898 (87.8)
Gender[a]	
Male	364 (19.2)
Female	1153 (60.7)
No answer	381 (20.1)
Age group (years)[a]	
<18	5 (0.3)
18-40	358 (18.9)
41-60	507 (26.7)
61-80	141 (7.4)
>80	1 (0.0)
Unknown	886 (46.7)
Medical diagnosis?[a]	
Yes	892 (47.0)
No	534 (28.1)
No answer	472 (24.9)
If yes, how were you diagnosed?[b]	
By test	628 (70.4)
By doctor	262 (294)
Unknown	2 (0.2)

[a]Percentages are based off those who consented (n=1898).

[b]Percentages are based off those who were medically diagnosed (n=892).

Table 3. Number of individual submissions in the assessment of burden of COVID-19 tool.

Number of times submitted	Individuals (n=1517), n (%)
One time	1517 (100)
Two times	356 (23.5)
Three times	206 (13.6)
Four times	134 (8.8)
Five times	98 (6.5)
Six times	76 (5.0)
Seven times	57 (3.8)
Other	441 (29.1)

Discussion

The ABCoV tool was created by an expert team to monitor patients with a (suspected) COVID-19 infection. This tool may detect problems with patients' physical and psychological health, their social life, and their lifestyle risk factors at an early stage. The tool can be used to study the long-term patient-experienced burden of COVID-19 and provide the insights needed to drive

optimal treatment. Obviously, this approach has its limitations. People may not always be aware of a COVID-19 symptom or differentiate it from other causes. This bias can potentially be reduced substantially by asking participants how and when they were diagnosed with COVID-19, by including a large number of participants, and by asking them to submit the ABCoV tool multiple times. We tried to address all possible items; all items from the original ABC tool, except one (exacerbations), were included. We added all other items mentioned in the literature plus an *open text field* for unforeseen symptoms. The next step is to evaluate the ABCoV tool's validity, practical use, and user experience. Frequent feedback from patients and health care providers is needed to ensure its optimization.

Finally, the purpose of the ABCoV tool is to support patients with COVID-19. In its present presentation, it is a preliminary tool to be validated and evaluated in the future. Better understanding of COVID-19 symptoms obtained by these longitudinal patient-reported outcomes may enable more insights into the long-term impact and disease burden after an infection with COVID-19 and provide tailored health care in a digital patient-centered environment.

Acknowledgments

The Lung Foundation Netherlands supported the software development of the ABCoV tool and created the Coronalungsquare website [29].

Access to deidentified data or related documents can be requested through submission of a proposal with a valuable research question, necessary data protection plan, and ethical approvals. A contract will be signed. Data requests should be addressed to the corresponding author.

M Rutgers, Director of the Dutch Lung Foundation, created the opportunity to introduce the tool to a large public audience via the Coronalungsquare website. Anton Kool, Director of Curavista, provided the technical team that developed and validated the software for the tool in a short time. Dr J Landers, Senior Research Fellow, Hertford College (Oxford, United Kingdom), proofread the manuscript.

Authors' Contributions

EMJvN contributed toward the concept, analyzing process, literature search, and preliminary data interpretation and conclusion. DC contributed toward the literature research concerning the ABCoV tool, cowriting, and approval for submission. CM contributed toward the literature search, cowriting, and approval for submission. CVDB contributed toward the first draft of the article, the literature search, and Table 1 and Multimedia Appendix 1. MK, JiV, OVS, and NHC contributed toward coreading, suggesting improvements, and approval for submission.

Conflicts of Interest

EMJvN is a PhD candidate at the Department of Public Health and Primary Care, Leiden University Medical Centre, Leiden, and is the co-owner of Curavista BV.

Multimedia Appendix 1
The assessment of burden of COVID-19 tool (not validated in English).
[PDF File (Adobe PDF File), 48 KB - formative_v5i3e22603_app1.pdf]

References

1. Coronadashboard. Rijksoverheid. 2020. URL: https://coronadashboard.rijksoverheid.nl/ [accessed 2021-01-24]
2. Chua F, Armstrong-James D, Desai SR, Barnett J, Kouranos V, Kon OM, et al. The role of CT in case ascertainment and management of COVID-19 pneumonia in the UK: insights from high-incidence regions. Lancet Respir Med 2020 May;8(5):438-440 [FREE Full text] [doi: 10.1016/S2213-2600(20)30132-6] [Medline: 32220663]
3. Multidisciplinaire LCI-richtlijn Q-koortsvermoeidheidssyndroom. LCI richtlijnen. 2020. URL: https://lci.rivm.nl/sites/default/files/2019-07/Multidisciplinaire%20LCI-richtlijn%20QVS%20hoofdtekst%20%28zonder%20bijlagen%20en%20literatuurlijst%29.pdf [accessed 2020-02-12]
4. Onderzoek gestart naar CALD, de nieuwe chronische longziekte door corona. Trouw. 2020 May 07. URL: https://www.trouw.nl/nieuws/onderzoek-gestart-naar-cald-de-nieuwe-chronische-longziekte-door-corona~bbfe7011/ [accessed 2020-11-29]
5. Zorg en onderzoek. Coronalongplein. 2020. URL: https://coronalongplein.nl/informatie/zorg-en-onderzoek [accessed 2020-11-29]
6. Longfonds. URL: https://www.longfonds.nl [accessed 2020-11-24]
7. Lung Foundation Netherlands. Longfonds. URL: https://research.longfonds.nl/lung-foundation-netherlands [accessed 2020-11-24]

8. Slok AHM, in 't Veen JCCM, Chavannes NH, van der Molen T, Rutten-van Mölken MPMH, Kerstjens HAM, et al. Development of the Assessment of Burden of COPD tool: an integrated tool to measure the burden of COPD. NPJ Prim Care Respir Med 2014 Jul 10;24:14021. [doi: 10.1038/npjpcrm.2014.21] [Medline: 25010353]
9. Slok AHM, Kotz D, van Breukelen G, Chavannes NH, Rutten-van Mölken MPMH, Kerstjens HAM, et al. Effectiveness of the Assessment of Burden of COPD (ABC) tool on health-related quality of life in patients with COPD: a cluster randomised controlled trial in primary and hospital care. BMJ Open 2016 Jul 11;6(7):e011519 [FREE Full text] [doi: 10.1136/bmjopen-2016-011519] [Medline: 27401361]
10. Goossens LMA, Rutten-van Mölken MPMH, Boland MRS, Donkers B, Jonker MF, Slok AHM, Research team that developed the ABC tool. ABC Index: quantifying experienced burden of COPD in a discrete choice experiment and predicting costs. BMJ Open 2017 Dec 26;7(12):e017831 [FREE Full text] [doi: 10.1136/bmjopen-2017-017831] [Medline: 29282261]
11. Slok AHM, Bemelmans TCH, Kotz D, van der Molen T, Kerstjens HAM, In 't Veen JCCM, et al. The Assessment of Burden of COPD (ABC) Scale: a reliable and valid questionnaire. COPD 2016 Aug;13(4):431-438. [doi: 10.3109/15412555.2015.1118025] [Medline: 26788838]
12. COPD. NHG-Richtlijnen. URL: https://richtlijnen.nhg.org/standaarden/copd [accessed 2020-11-29]
13. Boudewijns EA, Claessens D, van Schayck OCP, Keijsers LCEM, Salomé PL, In 't Veen JCCM, et al. ABC-tool reinvented: development of a disease-specific 'Assessment of Burden of Chronic Conditions (ABCC)-tool' for multiple chronic conditions. BMC Fam Pract 2020 Jan 13;21(1):11 [FREE Full text] [doi: 10.1186/s12875-019-1075-8] [Medline: 31931729]
14. COVID-19. LCI richtlijnen. 2020. URL: https://lci.rivm.nl/richtlijnen/covid-19 [accessed 2020-07-04]
15. Disease background of COVID-19. European Centre for Disease Prevention and Control. URL: https://www.ecdc.europa.eu/en/2019-ncov-background-disease [accessed 2020-07-04]
16. Coronavirus disease 2019 (COVID-19) pandemic: increased transmission in the EU/EEA and the UK – seventh update. European Centre for Disease Prevention and Control. 2020 Mar 25. URL: https://www.ecdc.europa.eu/sites/default/files/documents/RRA-seventh-update-Outbreak-of-coronavirus-disease-COVID-19.pdf [accessed 2020-07-04]
17. Ngai JC, Ko FW, Ng SS, To KW, Tong M, Hui DS. The long-term impact of severe acute respiratory syndrome on pulmonary function, exercise capacity and health status. Respirology 2010 Apr;15(3):543-550 [FREE Full text] [doi: 10.1111/j.1440-1843.2010.01720.x] [Medline: 20337995]
18. Lam MH, Wing Y, Yu MW, Leung C, Ma RCW, Kong APS, et al. Mental morbidities and chronic fatigue in severe acute respiratory syndrome survivors: long-term follow-up. Arch Intern Med 2009 Dec 14;169(22):2142-2147. [doi: 10.1001/archinternmed.2009.384] [Medline: 20008700]
19. Longfibrose. Longfonds. URL: https://www.longfonds.nl/longfibrose/symptomen-longfibrose [accessed 2020-07-04]
20. Hui DS, Azhar EI, Kim Y, Memish ZA, Oh M, Zumla A. Middle East respiratory syndrome coronavirus: risk factors and determinants of primary, household, and nosocomial transmission. Lancet Infect Dis 2018 Aug;18(8):e217-e227 [FREE Full text] [doi: 10.1016/S1473-3099(18)30127-0] [Medline: 29680581]
21. Mo X, Jian W, Su Z, Chen M, Peng H, Peng P, et al. Abnormal pulmonary function in COVID-19 patients at time of hospital discharge. Eur Respir J 2020 Jun;55(6). [doi: 10.1183/13993003.01217-2020] [Medline: 32381497]
22. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020 Feb 15;395(10223):497-506 [FREE Full text] [doi: 10.1016/S0140-6736(20)30183-5] [Medline: 31986264]
23. 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]
24. Bein T, Weber-Carstens S, Apfelbacher C. Long-term outcome after the acute respiratory distress syndrome: different from general critical illness? Curr Opin Crit Care 2018 Feb;24(1):35-40 [FREE Full text] [doi: 10.1097/MCC.0000000000000476] [Medline: 29189296]
25. Matthay MA, Zemans RL, Zimmerman GA, Arabi YM, Beitler JR, Mercat A, et al. Acute respiratory distress syndrome. Nat Rev Dis Primers 2019 Mar 14;5(1):18 [FREE Full text] [doi: 10.1038/s41572-019-0069-0] [Medline: 30872586]
26. GPS: Global Psychotrauma Screen. Global Collaboration on Traumatic Stress. URL: https://www.global-psychotrauma.net/gps [accessed 2020-01-01]
27. Hamer M, Kivimäki M, Gale CR, Batty GD. Lifestyle risk factors, inflammatory mechanisms, and COVID-19 hospitalization: a community-based cohort study of 387,109 adults in UK. Brain Behav Immun 2020 Jul;87:184-187 [FREE Full text] [doi: 10.1016/j.bbi.2020.05.059] [Medline: 32454138]
28. Ahmed A, Haque T, Rahman MM. Lifestyle acquired immunity, decentralized intelligent infrastructures, and revised healthcare expenditures may limit pandemic catastrophe: a lesson from COVID-19. Front Public Health 2020;8:566114. [doi: 10.3389/fpubh.2020.566114] [Medline: 33224915]
29. Coronalongplein. URL: https://www.coronalongplein.nl [accessed 2021-03-19]

Abbreviations

ABC: assessment of burden of chronic obstructive pulmonary disease
ABCoV: assessment of burden of COVID-19
ARDS: acute respiratory distress syndrome
CAPHRI: Care and Public Health Research Institute
CEO: chief executive officer
COPD: chronic obstructive pulmonary disease
GP: general practitioner
GPS: Global Psychotrauma Screen
MERS: Middle East respiratory syndrome
PROM: patient-reported outcome measurement
SARS: severe acute respiratory syndrome

Edited by G Eysenbach; submitted 17.07.20; peer-reviewed by J Delos, J Bricker; comments to author 17.11.20; revised version received 02.12.20; accepted 17.03.21; published 31.03.21.

Please cite as:
van Noort EMJ, Claessens D, Moor CC, Berg CALVD, Kasteleyn MJ, in 't Veen JCCM, Van Schayck OCP, Chavannes NH
Online Tool for the Assessment of the Burden of COVID-19 in Patients: Development Study
JMIR Form Res 2021;5(3):e22603
URL: https://formative.jmir.org/2021/3/e22603
doi:10.2196/22603
PMID:33729982

Original Paper

Correlations Between Facial Expressivity and Apathy in Elderly People With Neurocognitive Disorders: Exploratory Study

Radia Zeghari[1,2], PhD; Alexandra König[1,3], PhD; Rachid Guerchouche[3], PhD; Garima Sharma[4], PhD; Jyoti Joshi[4], PhD; Roxane Fabre[1,5], MSc; Philippe Robert[1,2], MD, PhD; Valeria Manera[1,2], PhD

[1]Cognition Behaviour Technology Research Unit, Memory Center, Université Côte d'Azur, Nice, France

[2]Association Innovation Alzheimer, Nice, France

[3]French Institute for Research in Computer Science and Automation, Valbonne, France

[4]Monash University, Monash, Australia

[5]Centre Hospitalier Universitaire de Nice, Département de Santé Publique, Nice, France

Corresponding Author:
Radia Zeghari, PhD
Cognition Behaviour Technology Research Unit
Memory Center
Université Côte d'Azur
10 Rue Molière
Nice
France
Phone: 33 4 92 03 47 66
Email: radia.zeghari@univ-cotedazur.fr

Abstract

Background: Neurocognitive disorders are often accompanied by behavioral symptoms such as anxiety, depression, and/or apathy. These symptoms can occur very early in the disease progression and are often difficult to detect and quantify in nonspecialized clinical settings.

Objective: We focus in this study on apathy, one of the most common and debilitating neuropsychiatric symptoms in neurocognitive disorders. Specifically, we investigated whether facial expressivity extracted through computer vision software correlates with the severity of apathy symptoms in elderly subjects with neurocognitive disorders.

Methods: A total of 63 subjects (38 females and 25 males) with neurocognitive disorder participated in the study. Apathy was assessed using the Apathy Inventory (AI), a scale comprising 3 domains of apathy: loss of interest, loss of initiation, and emotional blunting. The higher the scale score, the more severe the apathy symptoms. Participants were asked to recall a positive and a negative event of their life, while their voice and face were recorded using a tablet device. Action units (AUs), which are basic facial movements, were extracted using OpenFace 2.0. A total of 17 AUs (intensity and presence) for each frame of the video were extracted in both positive and negative storytelling. Average intensity and frequency of AU activation were calculated for each participant in each video. Partial correlations (controlling for the level of depression and cognitive impairment) were performed between these indexes and AI subscales.

Results: Results showed that AU intensity and frequency were negatively correlated with apathy scale scores, in particular with the emotional blunting component. The more severe the apathy symptoms, the less expressivity in specific emotional and nonemotional AUs was displayed from participants while recalling an emotional event. Different AUs showed significant correlations depending on the sex of the participant and the task's valence (positive vs negative story), suggesting the importance of assessing male and female participants independently.

Conclusions: Our study suggests the interest of employing computer vision-based facial analysis to quantify facial expressivity and assess the severity of apathy symptoms in subjects with neurocognitive disorders. This may represent a useful tool for a preliminary apathy assessment in nonspecialized settings and could be used to complement classical clinical scales. Future studies including larger samples should confirm the clinical relevance of this kind of instrument.

(JMIR Form Res 2021;5(3):e24727) doi:10.2196/24727

KEYWORDS

apathy; action units; assessment; ICT; facial video analysis; neurocognitive disorders; neurocognitive; facial analysis

Introduction

Apathy is one of the most common neuropsychiatric symptoms in neurocognitive disorders (NCDs) such as Alzheimer disease, Parkinson disease, Huntington disease, and vascular dementia and is prevalent in several psychiatric pathologies such as schizophrenia and major depression [1]. It can be defined as a reduction in goal-directed behavior that persists over time causing impairment in global functioning. Three dimensions of apathy have been identified, including loss/reduction in goal-directed behavior and goal-directed cognitive activity (eg, reduced interests and reduced indoor and outdoor activities), emotions (eg, emotional blunting), and social interactions (eg, reduced interactions with family members and friends) [2]. Apathy is a debilitating symptom: it significantly decreases the quality of life of patients with NCD and their caregivers and increases the risk for institutionalization in outpatients [3] and mortality for nursing home residents (even after controlling for depression) [4]. Apathy has also been associated with faster cognitive and functional decline [5,6]. A recent study suggested that apathy significantly increases the risk of developing dementia, resulting in a loss of autonomy in activities of daily living [7]. As apathetic persons interact less with their surroundings, show less interests in their family and in leisure activities, they are significantly less stimulated, which may increase the progression of NCD. Critically, preliminary evidence suggests that interventions targeting apathy in people with mild cognitive impairment (through repetitive transcranial magnetic stimulation) may be effective in improving global cognitive functioning [8], thus suggesting that identifying apathy early in disease progression and putting in place early treatment options could offer new opportunities for dementia prevention [9].

Currently, apathy is assessed by using various clinical scales or questionnaires (including self-reports and scales completed by caregivers and/or the clinician; see Radakovic et al [10] for a review), which allow for quantification of apathy symptoms over continuous scales. Furthermore, apathy can be assessed using the Apathy Diagnostic Criteria (ADC), which allow for classification of patients as apathetic versus nonapathetic based on the observed symptomatology [2]. All these instruments suffer from the risk of bias resulting from the assessor's subjectivity [11]. For instance, self-report clinical scales rely on the patient's ability to recall a change in their activities, interests, or emotional reactivity, which might be difficult considering that this requires preserved memory and insight to some extent. Similarly, the clinician may have access only to limited information concerning the patient's changes in everyday activities, thus resulting in underestimation of apathy [11].

Another obstacle that can emerge is that apathy can often be misdiagnosed as depression because of frequent comorbidities such as fatigue and anhedonia and a considerable overlap in key symptoms such as diminished interest, psychomotor retardation, or social withdrawal [12-14]. However, apathy and depression can be dissociated based on emotional deficits. Indeed, the emotion dimension in apathy can be described as a blunted affect whereas depression is characterized by the presence of negative emotions and sadness. Early differential diagnosis is of great importance since treating apathy with antidepressants can lead to aggravation of symptoms [2,11]. Therefore, there is an urgent need to detect apathy more objectively at early stages using objective, noninvasive methods to provide timely treatment and prevent it from aggravating cognitive symptoms. There is today a growing interest in finding new sensitive measures to detect behavioral aspects of dementia as they may represent a heavier burden for caregivers than the cognitive dysfunctions themselves [15,16].

New information and communication technologies can provide objective and more sensitive measures of human behaviors and have been recommended for neurocognitive disorders and apathy [17,18]. The behavioral correlates of apathy have been investigated through oculomotor movement using eye tracking [19], global activity or sleep disturbances using accelerometers [20,21], voice features [22,23], and facial expressivity [24-26]. The additional value of combining multimodal measures has been demonstrated before in depression by merging audio features with facial activity [27-29]. Associating these new markers with other recent methods such as ecological momentary assessment [30,31], which allows patients to report on symptoms remotely, could represent the future of more naturalistic psychiatric evaluations. These techniques allow the tracking of changes in mood and cognition continuously leading to timely prevention of further decline.

Today, facial expressivity, which may be altered in apathetic subjects, can be measured automatically by means of computer vision-based facial analysis methods. The Facial Action Coding System allows detection of facial behavior by identifying specific facial muscle movements called action units (AUs) [32]. Some AUs are specifically linked to emotions such as AU 4 (cheek raiser) and 12 (lips puller, smile) being linked to joy. Some AUs are mainly associated with positive emotions, while others are more associated with negative emotions (see Table 1 for a description of AUs). Girard and colleagues [33] found that automatic facial expression analysis was consistent with manual coding, validating its use in clinical research. In their study, they analyzed AUs one by one to extract a pattern linked to depression. Seidl and colleagues [26] studied the association of cognitive decline and facial expressivity. The results showed apathy moderated the effect of cognitive decline on facial expression in Alzheimer disease and was significantly correlated to decreased general and specific facial expressivity.

Table 1. Action unit descriptions [32] and emotion valence [34].

AU[a] number	FACS[b] name	Emotion valence related
1	Inner brow raiser	–[c]
2	Outer brow raiser	+[d]
4	Brow lowerer	–
5	Upper lid raiser	–
6	Cheek raiser	+
7	Lid tightener	+/–[e]
9	Nose wrinkler	–
10	Upper lip raiser	–
11	Nasolabial deepener	=[f]
12	Lip corner puller	+
14	Dimpler	+
15	Lip corner depressor	–
17	Chin raiser	=
20	Lip stretcher	–
23	Lip tightener	+/–
25	Lips part	+
26	Jaw drop	+
45	Blink	=

[a]AU: action unit.

[b]FACS: Facial Action Coding System.

[c]Negative valence.

[d]Positive valence.

[e]Either positive or negative valence.

[f]Neutral.

In a previous pilot study, we aimed to identify facial behavior associated with apathy through automated video analysis and investigate how facial variations (expression and movement) could help characterize this symptom [24]. Our algorithm was able to classify subjects as apathetic versus nonapathetic with 84% accuracy. To validate the deep learning algorithm, we relied on the leave-one-out cross-validation technique, which consists of training a model using n–1 available subjects and validate using the remaining subject. In another research work, we employed a different approach with the same dataset [25]. Here again we used deep learning techniques to build models but added audio features to the video ones. The final model was 76% accurate in discriminating apathetic versus nonapathetic subjects.

Despite these promising findings, the problem of using deep learning techniques for such studies is that the generally low number of included subjects makes it difficult to allow for a generalization of the built models. In addition, while correct classification of apathetic versus nonapathetic subjects is an important challenge, from a clinical point of view it is crucial to precisely understand which features are more sensitive for apathy assessment and investigate links between these features and the degree of apathy severity as well as its subdomains, as measured by continuous scales.

Therefore, we aim with this exploratory study to better understand the quantitative relationship between facial expressivity and apathy and its subdomains (emotional blunting, loss of initiation, loss of interest) independently of depression and level of cognitive decline, which have been previously shown to affect emotional expressivity. For this, we included a larger sample of participants from which we correlated basic facial AU activation and intensity with apathy scale scores.

We hypothesize that the higher the apathy score of a participant, the lower and less intense would be his facial expressivity especially for AUs involved in positive or negative emotions [34].

Methods

Participants

A total of 63 subjects were included in the study; 7 participants had subjective memory complaints and the rest were diagnosed with NCDs [35] (39 mild and 17 major), including 11 subjects with Alzheimer disease, 22 subjects with vascular dementia,

and 11 subjects with affective disorders. Participants were recruited from the Memory Center of Nice University Hospitals, France, and from the Cognition Behavioral Technology research lab of the Université Cote d'Azur in the context of motivation activation research protocol. Participants with mild NCD were previously followed at the Memory Center. Participants were not included if they had sensory or motor impairments interfering with the protocol completion. For each participant, clinicians assessed the global level of cognitive impairment using the Mini-Mental State Examination (MMSE) [36], the presence of depression using the Neuropsychiatric Inventory (NPI) [37], the presence of apathy using the ADC [2], and the severity of apathy symptoms using the Apathy Inventory (AI) [38]. The AI is divided into three subscales: loss of initiation (AI-initiation), loss of interest (AI-interest), and emotional blunting (AI-affect). The NPI depression score is calculated by reporting the intensity (1=mild, 2=moderate, 3=severe) and frequency (1=rarely to 4=very often). The demographic and clinical profiles of participants are presented in Table 2. Additional demographic features (including comparisons between apathetic versus nonapathetic subjects based on the ADC) are reported in Multimedia Appendix 1.

Table 2. Sociodemographic, cognitive, and behavioral variables of study participants.

Characteristic	Total (n=63), mean (SD)	Females (n=38), mean (SD)	Males (n=25), mean (SD)	Cohen *d*
Age in years	72.95 (8.40)	72.66 (8.89)	73.40 (7.76)	0.09
MMSE[a]	24.06 (3.84)	23.68 (4.09)	24.64 (3.43)	0.25
NPI[b] depression	1.17 (1.95)	1.13 (2.03)	1.24 (1.85)	0.06
AI[c] affect	0.33 (0.74)	0.24 (0.68)	0.48 (0.82)	0.33
AI initiation	1.32 (1.47)	1.00 (1.38)	1.80 (1.50)	0.56
AI interest	1.17 (1.33)	0.92 (1.30)	1.56 (1.29)	0.49
AI total	2.81 (3.03)	2.11 (2.83)	3.88 (3.07)	0.61

[a]MMSE: Mini-Mental State Examination.

[b]NPI: Neuropsychiatric Inventory.

[c]AI: Apathy Inventory.

Ethics

The study was performed as defined in the Declaration of Helsinki. The protocol was approved by the ethics committee (Comité de Protection de Personnes—CPP Est III, France; MoTap: RCB ID No. 2017-A01366-4). Informed written consent was obtained from all participants before the study.

Free Emotional Speech Task

Participants were asked to recall a positive and negative event of their life in maximum 1 minute. This free speech task requires only a low cognitive load and is supposed to trigger an emotional response. Participants were audio and videorecorded using a tablet device in a quiet room of the Resources and Research Memory Center at the Claude Pompidou Institut in Nice. The tablet was facing the subjects but not hiding the investigator, allowing a more natural interaction. In a previous study, we have analyzed the simple audio files and found that certain voice features were associated with apathy presence [23].

Video Features Extraction

For the features extraction we used OpenFace 2.0 software, which is an open-source facial behavior analysis toolkit [39]. OpenFace is an implementation of face recognition with deep neural networks. OpenFace allows detection of a single or multiple faces in an image using pretrained models. Many basic features can be extracted such as facial landmarks, eyes, and head positions. Higher level features are also provided by OpenFace in a video: head direction and movements, eye gaze estimations, and expressed emotions. Emotions are mainly computed by combining several AUs. Since the accuracy of detected emotions is highly related to the datasets used for training the neural networks and since we are working with a small dataset of elderly people that doesn't allow us to retrain the model, the obtained accuracy on detected emotions is quite unsatisfying. We then investigated the use of AUs as middle-level features to detect apathy. We use here the 17 main computed AUs by OpenFace: 1, 2, 4, 5, 6, 7, 9, 10, 12, 14, 15, 17, 20, 23, 25, 26, and 45 (see Table 1). For each frame of the recorded videos, OpenFace provides AU intensity and presence. We then calculated 2 measures for each AU in both videos: the average intensity and activation of each AU in each video. Intensity ranged from 0 (absent) to 1 (present at minimum intensity) to 5 (present at maximum intensity), with continuous values in between. The mean intensity for each AU was calculated as the average score across all the video frames. To obtain a more general measure of AU activation, we also computed for each subject the mean activation and intensity across all AUs.

Statistical Analysis

Statistical analyses were performed using SPSS Statistics version 23.0.0 for Mac software (IBM Corporation) and R version 3.6.3 (R Foundation for Statistical Computing). To investigate the linear relationships between apathy and facial expressivity, we performed partial linear correlations between AU activation and intensity and the AI subscales using depression (NPI depression) and level of cognitive impairment (MMSE score) as covariates for both the positive and negative story. As most of the clinical scales were not normally distributed (as indexed by Shapiro-Wilks tests), Spearman rho correlations were employed. We wanted to control for the effect of cognitive impairment as

it might have an impact on the task itself to recall an emotional event. Similarly, we wanted to control for the effects of depressive symptoms. In the analyses run on each specific AU, we separated males and females as studies have found that women and men can express their emotions differently [40-42].

Results

Descriptive Analysis

A total of 63 subjects (38 females and 25 males) were included in the study. The demographic and clinical profiles of participants are presented in Table 2. Sociodemographic features of apathetic and nonapathetic subjects and correlations among the different clinical scales are reported in Multimedia Appendix 1.

The average activation frequency and intensity for each AU are reported in Multimedia Appendix 1. In the positive story, the average global intensity of all AUs was 0.52 (SD 0.15) and the average global frequency was 0.29 (SD 0.07). In the negative story, the average global intensity was 0.51 (SD 0.28) and global frequency 0.28 (SD 0.08). Average presence and intensity for each AU for males and females are presented in Multimedia Appendix 1. In the negative story, AU 7 was the most intense facial expression in the whole sample (mean 1.60 [SD 0.84] on a 0 to 5 scale) followed by AU 10 (mean 1.02 [SD 0.47]) and AU 4 (mean 0.45 [SD 0.57]). AU 5 was the most frequently seen in the video (mean 0.68 [SD 0.30]) followed by AU 4 (mean 0.37 [SD 0.34) and AU 7 (mean 0.37 [SD 0.31]). In the positive story, AU 7 was also the most intense (mean 1.68 [SD 0.86]) followed by AU 10 (mean 1.1 [SD 0.50]) and AU 6 (mean 0.85 [SD 0.53]). AU 5 was the most frequently seen (mean 0.63 [SD 0.33]) followed by AU 7 (mean 0.40 [SD 0.32]) and AU 23 (mean 0.39 [SD 0.29]).

AU Correlations to Scales

Global Scales

A general analysis considering the global activation and intensity across AUs revealed a small but significant negative partial correlation between global AU intensity to AI-affect score in the negative story ($r_{s(61)}$=–0.26, *P*=.04), suggesting that more apathetic participants showed lower AU intensity in the negative story. No other significant linear correlation was found between the apathy scales and global activation.

Single AUs

Spearman partial correlations between each AU and apathy scale score controlled by the NPI depression and MMSE scores and divided by sex are presented in Table 3. Only significant correlations are reported. All correlations are presented in Multimedia Appendix 1.

For females, we found significant negative partial correlations between AUs and apathy, predominantly with the AI-affect subscale score, in both negative and positive stories (all of them of medium effect size, rho ranging from –0.33 to –0.47). All correlations were negative. The more severe the AI-affect score, the less frequent or less intense the AU. Specifically, significant correlations were found for AU 1, 2, 9, 10, 12, 25, and 45. In the positive story, the mean activations of AU 10 and 25 were correlated to AI-affect score. In the negative story, the mean activations of AU 1, 2, 9, 10 12, and 45 were correlated to AI-affect scores. The mean activations of AU 10 and 12 were correlated to AI-total score, and the mean activation of AU 10 was correlated to AI-initiation score.

For males, significant correlations were found for AU 1, 12, 14, 15, 17, 20, and 45. All of these correlations were negative, with effect sizes ranging from medium to large (rho ranging from –0.42 to –0.63). In the positive story, the mean intensity of AU 1 was correlated to AI-initiation score, 12 to AI-interest score, 14 to AI-total score, and 17 was correlated to AI-affect score. The mean activation of AU 14 was correlated to AI-initiation, AI-interest, and AI-total scores. The mean activation of AU 45 was correlated to AI-interest and AI-total scores. In the negative story, the mean intensity of AU 1 was correlated to AI-initiation and AI-total scores. The mean intensities of AU 15, 17, and 20 were correlated to AI-affect score. The mean activation of AU 28 was correlated to AI-initiation and NPI-apathy score.

Table 3. Spearman rho correlations between action units and apathy scales for males and females.

Characteristics, AU[a]	AI[b]-affect	AI-initiation	AI-interest	AI-total
Female				
Negative story				
Mean activation				
Inner brow raiser	–0.340	__[c]	—	—
Outer brow raiser	–0.422	—	—	—
Nose wrinkler	–0.382	—	—	—
Upper lip raiser	–0.350	–0.426	—	–0.337
Lip corner puller		—	—	–0.350
Blink	–0.349	—	—	—
Mean intensity				
Inner brow raiser	–0.361	—	—	—
Nose wrinkler	–0.470	—	—	—
Positive story				
Mean activation				
Upper lip raiser	–0.461	—	—	—
Lips part	–0.333	—	—	—
Male				
Negative story				
Mean activation				
Lip stretcher	–0.474	—	—	—
Mean intensity				
Inner brow raiser	—	–0.432	—	–0.426
Lip corner depressor	–0.628	—	—	—
Chin raiser	–0.510	—	—	—
Lip stretcher	–0.472	—	—	—
Positive story				
Mean activation				
Dimpler	—	–0.501	–0.625	–0.503
Blink	—	—	–0.428	–0.466
Mean intensity				
Inner brow raiser	—	–0.472	—	—
Lip corner puller	—	—	–0.418	—
Dimpler	—	—	—	–0.416
Chin raiser	–0.437	—	—	—

[a]AU: action unit.

[b]AI: Apathy Inventory.

[c]Not applicable.

Discussion

Principal Findings

In this exploratory study we aimed to verify whether facial expressivity, assessed using automatic video analysis, correlates with apathy severity in elderly subjects with NCDs. Specifically, we aimed to explore which particular AUs, intensity and activation frequency, are associated with apathy and its different subdomains in male and female subjects, employing 2 different emotion-related tasks (telling a positive and a negative story).

To assess apathy, we used the AI scale (clinician version) and its subdomains reduced affect, loss of interest, and loss of initiation. We extracted AUs from videorecordings of participants while they performed an emotional task consisting of recalling a positive and a negative event from their life. Studies have found that women and men can express their emotions differently [40-42]. Thus, we separated male and female participants for the analyses. We hypothesized that the AUs implicated in positive and negative emotions (joy, sadness, disgust, anger, surprise) would correlate with apathy levels. Specifically, the more important the apathy symptomatology, the lower the emotional activation/intensity. We classified AUs with emotional valence (positive, negative, neutral) according to Haines et al [34].

Partial correlations (controlling for depressive symptoms and level of cognitive impairment) showed that, overall, apathetic participants showed lower average AU intensity in the negative story. This result suggests that more apathetic subjects had less intense facial expressivity compared with less apathetic subjects, at least while telling a negative event. At the level of single AU analysis, several AUs were significantly less expressed with more pronounced apathy symptoms in the upper region (inner and outer brow raiser) and lower region (nose and lip movement) of the face. Moreover, the relevant AUs were different depending on sex and the task's emotional valence.

For instance, for men, AU 12 (smile) and 14 (dimpler), both considered positive AUs, were significantly less expressed the more important the behavioral and cognitive dimensions of apathy in the positive story. AU 15 (lip corner depressor) and 20 (lip stretcher), both considered negative AUs, correlated significantly with affect dimension in the negative story. AU 1 (inner brow raiser), involved in sadness, anger and fear, was less expressed in both stories the more severe the initiation symptoms. For women, in the positive story, AU 1 (inner brow raiser) intensity and 4 (brow lowerer) frequency (both involved in sadness) were significantly correlated to AI-affect in the negative story. Thus, some of the expected facial expressions are significantly less expressed in more apathetic subjects. For men, this was observed in both positive and negative storytelling. For women, this was mainly observed in the negative story. For women, AU 25 (lip part), involved in talking, was significantly less frequent in the positive story. This is in line with the symptomatology of apathetic subjects being less talkative [23]. Different AUs were correlated to apathy scales depending on the task's emotional valence. In the negative story, mainly negative valence AUs were correlated to apathy symptoms for men. Both positive and negative AUs showed significant correlations for women. In the positive story, positive AUs (12 and 14) and neutral AUs (17 and 45) were significantly less expressed for men. For women, only two AUs (10 and 25) showed significant correlations, one of negative and one of positive valence. Our findings suggest that the more severe the apathy symptoms, the less expressed are the AUs expected for the task's valence.

Correlations With the Affect Apathy Dimension

For both men and women in the negative story, the majority of AUs were associated with the affect subdomain. AI-affect (blunted affect) is mainly assessed with facial emotional responses in clinical interviews, which could explain this finding. Depending on the patient's emotional responsiveness in the interaction, the clinician will infer the presence of a potential emotional blunting. This is one of the main biases in the evaluation of apathy due to a lack of objectivity (clinician can miss cues) and context (patients are in a particular environment that does not translate the way they experience emotions in their daily lives). Similar studies found that flattened affect in schizophrenic or depressed patients is reflected in decreased spontaneous facial expression compared with control groups [43]. Furthermore, they found that depressed patients showed even less facial expressivity compared with schizophrenic subjects.

Interestingly, blunted affect was the subdomain of apathy most linked to lack of facial expressivity for women. This could be explained by the hypothesis that women tend to express their emotions more [40,41,44]. Studies have shown that emotional expressivity and emotional memory retrieval are different depending on sex [45,46]. Men experience emotions more strongly whereas women tend to express them more. Another explanation would be that women are expected to express more happiness and sadness, and a lack of facial expressivity due to apathy would be identified more quickly than for men who are less expected to display these emotions [47]. Future studies should investigate the link between subdomains of apathy and sex. We found that for women, AU 9 (nose wrinkler, involved in disgust) intensity was significantly correlated with AI-affect in the negative story. In one study, the authors predicted the sex of a person simply from their emotional expressions of happiness and disgust but not from expressions like sadness and surprise [41]. This AU is expected to be more pronounced and intense for women but seems significantly less expressed in people with apathy.

Correlations With the Loss of Interest and Lack of Initiation Dimensions

Only a few AUs were linked to loss of interest and lack of initiation. For women, only AU 10 was linked to lack of initiation, and none to loss of interest. However, for men, in the positive story, all the significant correlations to AUs (such as 1 and 14) were linked to either lack of initiation or loss of interest. One explanation could be that the positive condition triggered memories that might have involved activities such as gatherings with family or friends, eliciting less intense positive emotions the more severe the apathy symptoms. Emotional valence (positive, negative, and neutral) and arousal of an event will have an impact on the way emotions are stored and retrieved in autobiographical memory [48]. Positive memories are retrieved with richer details supposedly due to a bonding social role, while negative memories would have more of a survival role, so the events are not repeated. In the positive story, women often recalled the birth of a child or a wedding day while men recalled more diverse events.

Our results, despite being preliminary, are in line with previous research showing that besides cognitive deficits, apathy was significantly correlated with decreased overall and specific facial expression [26]. In depression, similar results have been found

[33]. Girard et al [33] compared specific AUs to depression severity and found that depressed people would smile less and express the emotion of contempt more. In apathy, the AUs involved in these emotions were all decreased for the male sample. However, different AUs were involved for the female sample, which underlines how important it is to consider sex differences in the study of emotional expressivity. Overall, apathy seems to have an impact on global facial expressivity and thus, can be detected by automated video analysis. Nevertheless, at this stage it remains a rather sensitive but not specific tool for detecting variations in emotional facial expression (which in turn can be an indicator for apathy or depression or negative symptoms, etc). Peham et al [49] studied facial emotional behavior of female subjects suffering from mental disorders (personality disorder, depression, anxiety, eating disorder, etc) during a clinical interview and found no distinctive patterns that were disease specific.

In depression detection, increasingly research efforts have been placed on merging audio and video features with encouraging results, demonstrating that when both modalities are combined more precise assessments can be obtained [27]. We previously demonstrated that, simply by extracting and analyzing voice features from the free emotional tasks, apathy prediction can be quite accurate [22,23]. Therefore, it can be assumed that in a next step, the fusion of audio and video features will improve on our current results, supporting the long-term goal of validating such technology for use in daily clinical practice.

Limitations

The main study limitations are the relatively small sample size, coupled with the rather high number of statistical tests performed. As this study was exploratory, we did not employ statistical corrections for multiple comparisons, which could increase the probability of making type I errors. This is one of the first studies that analyzes correlations between facial expressivity and apathy, and these results could be used to formulate more precise hypotheses on specific AU involvement, to be tested with more robust statistical methods in future studies. Another issue is that this study relied on clinicians to rate levels of apathy and other neuropsychiatric scales. Future research should employ multiple types of scales such as self-administered scales in order to understand if there is a gap between how the patient feels in his daily life and what the clinician is observing. Combining ecological momentary assessment with these novel measures (facial behavior, voice parameters, activity monitoring, etc) and biomarkers (magnetic resonance imaging) should also be included to ensure covering all types of assessments available since there are no existing gold standard for apathy scales [11,50]. More variables should be considered such as fatigue and anhedonia, as they both are related to apathy [14,51].

Conclusion

Overall, it can be concluded that computer vision-based facial analysis showed promising results in detecting blunted affect and global apathy in neurocognitive disorders. These preliminary results should be corroborated by further studies including a larger sample size, allowing researchers to test a reduced number of relevant hypotheses and apply corrections for multiple comparisons. As the effect size found in our study was medium to large, reduced facial expressivity may represent a promising proxy for emotional blunting. Specifically, as hypothesized, our results suggested that the presence of AUs relating to positive emotions is particularly relevant to assess apathy during the positive storytelling, while the presence of AUs relating to negative emotions may be relevant during negative storytelling, especially for men. Women may show a wider range of emotions in both positive and negative storytelling, thus suggesting the interest of assessing AUs related to both positive and negative emotions in both stories.

More research is needed to identify specific facial expressions associated with apathy and combine this method with other technologies such as automatic speech analysis and eye tracking to provide additional information for differential diagnosis. Identifying multimodal digital biomarkers (eg, voice features, activity patterns, eye paths, facial expression) and combining them with ecological momentary assessment could be the future of neuropsychiatric and cognitive assessment, allowing early detection of changes and therefore better adapted treatment [30]. These technologies could facilitate continuous monitoring to prevent relapses in depression, psychotic crisis in schizophrenia, or the detection of early signs of cognitive deficits and behavioral changes in neurocognitive disorders.

Acknowledgments

This work was supported by the Association Innovation Alzheimer, the JL Noisiez Foundation, and by the French government, through the UCA-JEDI (Cote d'Azur University: Joint, Excellent, and Dynamic Initiative) Investments in the Future project managed by the National Research Agency (reference number ANR-15-IDEX-01). This work was done in the context of the Digital Medicine–Brain, Cognition, Behavior program of the University Côte d'Azur. Thanks to all patients who participated in our study.

Authors' Contributions

AK and PR designed the studies. RZ and AK ran the experiments. RZ, VM, RF, and RG analyzed the results. GS, JJ, and RG extracted the facial features. All authors participated in writing the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Auxiliary tables.
[DOCX File , 64 KB - formative_v5i3e24727_app1.docx]

References

1. Husain M, Roiser JP. Neuroscience of apathy and anhedonia: a transdiagnostic approach. Nat Rev Neurosci 2018 Aug;19(8):470-484. [doi: 10.1038/s41583-018-0029-9] [Medline: 29946157]
2. Robert P, Lanctôt KL, Agüera-Ortiz L, Aalten P, Bremond F, Defrancesco M, et al. Is it time to revise the diagnostic criteria for apathy in brain disorders? The 2018 international consensus group. Eur Psychiatry 2018 Oct;54:71-76 [FREE Full text] [doi: 10.1016/j.eurpsy.2018.07.008] [Medline: 30125783]
3. Bakker C, de Vugt ME, van Vliet D, Verhey FRJ, Pijnenburg YA, Vernooij-Dassen MJFJ, et al. Predictors of the time to institutionalization in young- versus late-onset dementia: results from the Needs in Young Onset Dementia (NeedYD) study. J Am Med Dir Assoc 2013 Apr;14(4):248-253. [doi: 10.1016/j.jamda.2012.09.011] [Medline: 23123009]
4. Nijsten JMH, Leontjevas R, Pat-El R, Smalbrugge M, Koopmans RTCM, Gerritsen DL. Apathy: risk factor for mortality in nursing home patients. J Am Geriatr Soc 2017 Oct;65(10):2182-2189. [doi: 10.1111/jgs.15007] [Medline: 28791690]
5. Starkstein SE, Jorge R, Mizrahi R, Robinson RG. A prospective longitudinal study of apathy in Alzheimer's disease. J Neurol Neurosurg Psychiatry 2006 Jan;77(1):8-11 [FREE Full text] [doi: 10.1136/jnnp.2005.069575] [Medline: 16361584]
6. Robert PH, Berr C, Volteau M, Bertogliati C, Benoit M, Sarazin M, PréAL study. Apathy in patients with mild cognitive impairment and the risk of developing dementia of Alzheimer's disease: a one-year follow-up study. Clin Neurol Neurosurg 2006 Dec;108(8):733-736. [doi: 10.1016/j.clineuro.2006.02.003] [Medline: 16567037]
7. Ruthirakuhan M, Herrmann N, Vieira D, Gallagher D, Lanctôt KL. The roles of apathy and depression in predicting Alzheimer disease: a longitudinal analysis in older adults with mild cognitive impairment. Am J Geriatr Psychiatry 2019 Aug;27(8):873-882 [FREE Full text] [doi: 10.1016/j.jagp.2019.02.003] [Medline: 30910421]
8. Padala PR, Padala KP, Lensing SY, Jackson AN, Hunter CR, Parkes CM, et al. Repetitive transcranial magnetic stimulation for apathy in mild cognitive impairment: a double-blind, randomized, sham-controlled, cross-over pilot study. Psychiatry Res 2018 Mar;261:312-318 [FREE Full text] [doi: 10.1016/j.psychres.2017.12.063] [Medline: 29331848]
9. Manera V, Abrahams S, Agüera-Ortiz L, Bremond F, David R, Fairchild K, et al. Recommendations for the nonpharmacological treatment of apathy in brain disorders. Am J Geriatr Psychiatry 2020 Apr;28(4):410-420. [doi: 10.1016/j.jagp.2019.07.014] [Medline: 31495772]
10. Radakovic R, Harley C, Abrahams S, Starr JM. A systematic review of the validity and reliability of apathy scales in neurodegenerative conditions. Int Psychogeriatr 2015 Jun;27(6):903-923. [doi: 10.1017/S1041610214002221] [Medline: 25355282]
11. Clarke DE, Ko JY, Kuhl EA, van Reekum R, Salvador R, Marin RS. Are the available apathy measures reliable and valid? A review of the psychometric evidence. J Psychosom Res 2011 Jan;70(1):73-97 [FREE Full text] [doi: 10.1016/j.jpsychores.2010.01.012] [Medline: 21193104]
12. Yeager CA, Hyer L. Apathy in dementia: relations with depression, functional competence, and quality of life. Psychol Rep 2008 Jun;102(3):718-722. [doi: 10.2466/pr0.102.3.718-722] [Medline: 18763441]
13. Tagariello P, Girardi P, Amore M. Depression and apathy in dementia: same syndrome or different constructs? A critical review. Arch Gerontol Geriatr 2009;49(2):246-249. [doi: 10.1016/j.archger.2008.09.002] [Medline: 19022508]
14. Ang Y, Lockwood P, Apps MAJ, Muhammed K, Husain M. Distinct subtypes of apathy revealed by the Apathy Motivation Index. PLoS One 2017;12(1):e0169938 [FREE Full text] [doi: 10.1371/journal.pone.0169938] [Medline: 28076387]
15. Papastavrou E, Kalokerinou A, Papacostas SS, Tsangari H, Sourtzi P. Caring for a relative with dementia: family caregiver burden. J Adv Nurs 2007 Jun;58(5):446-457. [doi: 10.1111/j.1365-2648.2007.04250.x] [Medline: 17442030]
16. Matsumoto N, Ikeda M, Fukuhara R, Shinagawa S, Ishikawa T, Mori T, et al. Caregiver burden associated with behavioral and psychological symptoms of dementia in elderly people in the local community. Dement Geriatr Cogn Disord 2007;23(4):219-224. [doi: 10.1159/000099472] [Medline: 17299264]
17. Gros A, Bensamoun D, Manera V, Fabre R, Zacconi-Cauvin A, Thummler S, et al. Recommendations for the use of ICT in elderly populations with affective disorders. Front Aging Neurosci 2016;8:269 [FREE Full text] [doi: 10.3389/fnagi.2016.00269] [Medline: 27877126]
18. König A, Aalten P, Verhey F, Bensadoun G, Petit P, Robert P, et al. A review of current information and communication technologies: can they be used to assess apathy? Int J Geriatr Psychiatry 2014 Apr;29(4):345-358. [doi: 10.1002/gps.4017] [Medline: 24006215]
19. Chau SA, Chung J, Herrmann N, Eizenman M, Lanctôt KL. Apathy and attentional biases in Alzheimer's disease. J Alzheimers Dis 2016;51(3):837-846. [doi: 10.3233/JAD-151026] [Medline: 26890774]
20. Mulin E, Zeitzer JM, Friedman L, Le Duff F, Yesavage J. Relationship between apathy and sleep disturbance in mild and moderate Alzheimer's disease: an actigraphic study. J Alzheimers Dis 2011;25(1):85-91. [doi: 10.3233/JAD-2011-101701] [Medline: 21335662]

21. Valembois L, Oasi C, Pariel S, Jarzebowski W, Lafuente-Lafuente C, Belmin J. Wrist actigraphy: a simple way to record motor activity in elderly patients with dementia and apathy or aberrant motor behavior. J Nutr Health Aging 2015 Aug;19(7):759-764. [doi: 10.1007/s12603-015-0530-z] [Medline: 26193860]
22. Linz N, Klinge X, Tröger J. Automatic detection of apathy using acoustic markers extracted from free emotional speech. 2018 Presented at: 2nd Workshop on AI for Aging, Rehabilitation and Independent Assisted Living (ARIAL) IJCAI'18; 2018; Stockholm.
23. König A, Linz N, Zeghari R, Klinge X, Tröger J, Alexandersson J, et al. Detecting apathy in older adults with cognitive disorders using automatic speech analysis. J Alzheimers Dis 2019;69(4):1183-1193. [doi: 10.3233/JAD-181033] [Medline: 31127764]
24. Happy S, Dantcheva A, Das A, Zeghari R, Robert P, Bremond F. Characterizing the state of apathy with facial expression and motion analysis. 2019 Presented at: 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG); 2019; Lille p. 1-8. [doi: 10.1109/fg.2019.8756545]
25. Sharma G, Joshi J, Zeghari R, Guerchouche R. Audio-visual weakly supervised approach for apathy detection in the elderly. 2020 Presented at: 2020 International Joint Conference on Neural Networks (IJCNN); 2020; Glasgow p. 1-7. [doi: 10.1109/ijcnn48605.2020.9206829]
26. Seidl U, Lueken U, Thomann PA, Kruse A, Schröder J. Facial expression in Alzheimer's disease: impact of cognitive deficits and neuropsychiatric symptoms. Am J Alzheimers Dis Other Demen 2012 Mar;27(2):100-106 [FREE Full text] [doi: 10.1177/1533317512440495] [Medline: 22495337]
27. Joshi J, Goecke R, Alghowinem S, Dhall A, Wagner M, Epps J, et al. Multimodal assistive technologies for depression diagnosis and monitoring. J Multimodal User Interfaces 2013 Sep 7;7(3):217-228. [doi: 10.1007/s12193-013-0123-2]
28. Cohn JF, Kruez TS, Matthews I, Yang Y, Nguyen MH, Padilla MT, et al. Detecting depression from facial actions and vocal prosody. Proc 3rd Int Conf Affective Comput Intelligent Interaction Workshops 2009. [doi: 10.1109/ACII.2009.5349358]
29. Cummins N, Scherer S, Krajewski J, Schnieder S, Epps J, Quatieri TF. A review of depression and suicide risk assessment using speech analysis. Speech Commun 2015 Jul;71:10-49. [doi: 10.1016/j.specom.2015.03.004]
30. Kubiak T, Smyth J. Connecting domains—ecological momentary assessment in a mobile sensing framework. In: Baumeister H, Montag C, editors. Digital Phenotyping and Mobile Sensing. Studies in Neuroscience, Psychology and Behavioral Economics. Cham: Springer; 2019.
31. Ramsey AT, Wetherell JL, Depp C, Dixon D, Lenze E. Feasibility and acceptability of smartphone assessment in older adults with cognitive and emotional difficulties. J Technol Hum Serv 2016;34(2):209-223 [FREE Full text] [doi: 10.1080/15228835.2016.1170649] [Medline: 27683018]
32. Ekman P, Friesen WV. Manual for the Facial Action Coding System. Palo Alto: Consulting Psychologists Press; 1978.
33. Girard JM, Cohn JF, Mahoor MH, Mavadati SM, Hammal Z, Rosenwald DP. Nonverbal social withdrawal in depression: evidence from manual and automatic analysis. Image Vis Comput 2014 Oct;32(10):641-647 [FREE Full text] [doi: 10.1016/j.imavis.2013.12.007] [Medline: 25378765]
34. Haines N, Southward MW, Cheavens JS, Beauchaine T, Ahn W. Using computer-vision and machine learning to automate facial coding of positive and negative affect intensity. PLoS One 2019;14(2):e0211735 [FREE Full text] [doi: 10.1371/journal.pone.0211735] [Medline: 30721270]
35. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington: American Psychiatric Association; 2013.
36. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
37. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994 Dec;44(12):2308-2314. [doi: 10.1212/wnl.44.12.2308] [Medline: 7991117]
38. Robert PH, Clairet S, Benoit M, Koutaich J, Bertogliati C, Tible O, et al. The apathy inventory: assessment of apathy and awareness in Alzheimer's disease, Parkinson's disease and mild cognitive impairment. Int J Geriatr Psychiatry 2002 Dec;17(12):1099-1105. [doi: 10.1002/gps.755] [Medline: 12461757]
39. Baltrusaitis T, Robinson P, Morency L. OpenFace: an open source facial behavior analysis toolkit. 2016 Presented at: 2016 IEEE Winter Conference on Applications of Computer Vision (WACV); 2016; Lake Placid p. 1-10. [doi: 10.1109/wacv.2016.7477553]
40. Vail A, Grafsgaard J, Boyer K, Wiebe E, Lester J. Gender differences in facial expressions of affect during learning. Proc 2016 Conf User Modeling Adaptation Personalization 2016:65-73. [doi: 10.1145/2930238.2930257]
41. Xia B. Which facial expressions can reveal your gender? A study with 3D faces. ArXiv. Preprint posted online May 1, 2018 2018 [FREE Full text] [doi: 10.5860/choice.189890]
42. McDuff D, Kodra E, Kaliouby RE, LaFrance M. A large-scale analysis of sex differences in facial expressions. PLoS One 2017;12(4):e0173942 [FREE Full text] [doi: 10.1371/journal.pone.0173942] [Medline: 28422963]
43. Trémeau F, Malaspina D, Duval F, Corrêa H, Hager-Budny M, Coin-Bariou L, et al. Facial expressiveness in patients with schizophrenia compared to depressed patients and nonpatient comparison subjects. Am J Psychiatry 2005 Jan;162(1):92-101. [doi: 10.1176/appi.ajp.162.1.92] [Medline: 15625206]

44. Dimberg U, Lundquist LO. Gender differences in facial reactions to facial expressions. Biol Psychol 1990 Apr;30(2):151-159. [doi: 10.1016/0301-0511(90)90024-q] [Medline: 2285765]
45. Piefke M, Weiss PH, Markowitsch HJ, Fink GR. Gender differences in the functional neuroanatomy of emotional episodic autobiographical memory. Hum Brain Mapp 2005 Apr;24(4):313-324 [FREE Full text] [doi: 10.1002/hbm.20092] [Medline: 15704151]
46. Deng Y, Chang L, Yang M, Huo M, Zhou R. Gender differences in emotional response: inconsistency between experience and expressivity. PLoS One 2016;11(6):e0158666 [FREE Full text] [doi: 10.1371/journal.pone.0158666] [Medline: 27362361]
47. Hess U, Adams R, Kleck R. When two do the same, it might not mean the same: the perception of emotional expressions shown by men and women. In: Group Dynamics and Emotional Expression. Cambridge: Cambridge University Press; 2009:33-50.
48. Ford JH, Addis DR, Giovanello KS. Differential effects of arousal in positive and negative autobiographical memories. Memory 2012;20(7):771-778 [FREE Full text] [doi: 10.1080/09658211.2012.704049] [Medline: 22873402]
49. Peham D, Bock A, Schiestl C, Huber E, Zimmermann J, Kratzer D, et al. Facial affective behavior in mental disorder. J Nonverbal Behav 2015 Jul 11;39(4):371-396. [doi: 10.1007/s10919-015-0216-6]
50. Cummings J, Friedman JH, Garibaldi G, Jones M, Macfadden W, Marsh L, et al. Apathy in neurodegenerative diseases: recommendations on the design of clinical trials. J Geriatr Psychiatry Neurol 2015 Sep;28(3):159-173. [doi: 10.1177/0891988715573534] [Medline: 25809634]
51. Skorvanek M, Gdovinova Z, Rosenberger J, Saeedian RG, Nagyova I, Groothoff JW, et al. The associations between fatigue, apathy, and depression in Parkinson's disease. Acta Neurol Scand 2015 Feb;131(2):80-87. [doi: 10.1111/ane.12282] [Medline: 25288130]

Abbreviations

ADC: Apathy Diagnostic Criteria
AI: Apathy Inventory
AU: action unit
MMSE: Mini-Mental State Examination
NCD: neurocognitive disorder
NPI: Neuropsychiatric Inventory
UCA-JEDI: Cote d'Azur University: Joint, Excellent, and Dynamic Initiative

Edited by J Torous; submitted 02.10.20; peer-reviewed by W Eikelboom, N Nazari; comments to author 30.10.20; revised version received 23.11.20; accepted 18.01.21; published 31.03.21.

Please cite as:
Zeghari R, König A, Guerchouche R, Sharma G, Joshi J, Fabre R, Robert P, Manera V
Correlations Between Facial Expressivity and Apathy in Elderly People With Neurocognitive Disorders: Exploratory Study
JMIR Form Res 2021;5(3):e24727
URL: https://formative.jmir.org/2021/3/e24727
doi:10.2196/24727
PMID:33787499

Original Paper

Educational Needs and Preferences for Patient-Centered Outcomes Research in the Cystic Fibrosis Community: Mixed Methods Study

Emily M Godfrey[1*], MPH, MD; Traci M Kazmerski[2*], MSc, MD; Georgia Brown[3*], MLA; Erin K Thayer[1*], MPH; Laura Mentch[3*], EdM; Molly Pam[3*], BSc; Morhaf Al Achkar[1*], PhD, MD

[1]Department of Family Medicine, University of Washington School of Medicine, Seattle, WA, United States

[2]Department of Pediatrics, School of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

[3]Cystic Fibrosis Reproductive and Sexual Health Collaborative, Seattle, WA, United States

[*]all authors contributed equally

Corresponding Author:
Emily M Godfrey, MPH, MD
Department of Family Medicine
University of Washington School of Medicine
4311 11th Ave NE
Suite 210, Box 354982
Seattle, WA, 98105
United States
Phone: 1 206 685 4895
Fax: 1 206 616 4768
Email: godfreye@uw.edu

Abstract

Background: Cystic fibrosis (CF) is a rare, life-shortening, multiorgan disease, the treatment of which has seen significant increases in the life expectancy of those with CF. Many advances in CF care are thanks to the dedicated and active participation of people with CF as research participants. Unfortunately, most CF research teams still do not fully partner with people with CF or their caregivers.

Objective: The aim of this study was to determine the interest, knowledge gaps, and desired format for patient-centered outcomes research (PCOR) training in the CF community.

Methods: We surveyed patients, caregivers, researchers, research staff, and diverse health care providers via list servers and social media outreach about their knowledge of, experience with, and preferences for PCOR training components. We followed the survey with 3 small-group discussion sessions with 22 participants who completed the survey to establish consensus and prioritize key learning components of a PCOR training program. We summarized results using descriptive statistics.

Results: A total of 170 participants completed the survey (patients/caregivers: 96/170, 56.5%; researchers/health care providers: 74/170, 43.5%). Among providers, 26% (19/74) were physicians/advanced practice providers, 20% (15/74) were nurses, and 54% (40/74) were from other disciplines. Among all participants, 86.5% (147/170) expressed interest in learning about PCOR, although training topics and training format differed between the patient/caregiver and researcher/health care provider groups. Before participating in PCOR, patients/caregivers wanted to understand more about expectations of them as partners on PCOR research teams (82/96, 85%). Meanwhile, researchers/health care providers desired information on how to include outcomes important to patients/caregivers (55/74, 74%) and the quality and impact of PCOR research (52/74, 70% and 51/74, 69%, respectively). Patients/caregivers were most interested in learning about the time commitment as a PCOR team member (75/96, 78%). Researchers/health care providers wanted to receive training about how to establish trust (47/74, 64%) and maintain confidentiality (47/74, 64%) when including patient or caregiver partners on the PCOR team. During follow-up discussions, participants emphasized the importance of addressing the traditional patient/caregiver and researchers/health care provider hierarchy by teaching about transparency, appreciation, creating a common language between the groups, and providing specific training on “how” to do PCOR.

Conclusions: Our findings suggest CF community members are interested in PCOR. A high-quality training program would fill a current deficit in methodological research. This assessment identified the topics and formats desired and can be used to develop targeted training to enhance meaningful PCOR in CF.

(JMIR Form Res 2021;5(3):e24302) doi:10.2196/24302

KEYWORDS

cystic fibrosis; needs assessment; patient-centered outcomes research; training; education; team building; patient engagement

Introduction

Cystic fibrosis (CF) is a rare, life-shortening, multi-organ disease that affects approximately 30,000 patients in the United States [1]. CF impedes cell chloride protein channels leading to a cycle of impaired mucociliary clearance, inflammation, and infection in the respiratory tract, with related effects on the digestive, endocrine, immune, and reproductive systems. It can lead to severe respiratory and digestive problems as well as other complications such as infections and diabetes. Although CF traditionally affected children, today more than 50% of individuals with CF are adults with a median survival of almost 45 years [2].

People with CF have a long history of actively being involved as participants in research and thus have played a critical role in the medical advances seen in the CF community. The Cystic Fibrosis Foundation (CFF) is a large stakeholder organization that has led the effort to fund clinical research studies that advance the care and treatment of CF patients in the United States [3]. The latest breakthrough has been the development and recent Food and Drug Administration approval of a highly effective modulator therapy medication that corrects defective protein channels for 90% of those with CF 12 years and older, which is approximately 27,000 people in the United States [4].

Despite these achievements, patient participation in CF-related research has been mostly limited to involvement as participants enrolled in clinical trials. In the past two decades, the CF community and CFF have worked to deepen partnerships with patients through skill-building opportunities in quality improvement methods and the formation of patient and family advisory boards [5,6]. Another CF stakeholder organization, Cystic Fibrosis Research, Inc (CFRI), hosts an annual conference for researchers, clinicians, caregivers, and patients with CF and provides monetary support for CF research driven by stakeholders in the CF community [7]. Yet, despite these opportunities, partnership with people with CF and their families as equal players in research design and performance remains limited.

The Patient-Centered Outcomes Research Institute (PCORI), established in 2010 in the United States, now requires research engagement in all of its funded studies [8,9]. Patient-centered outcomes research fosters coproduction by engaging patients, caregivers, and other stakeholders as equal members on research teams [10]. Essentially, PCOR shifts power that typically rests with researchers over to service users (ie, patients) [11]. Stakeholder engagement has been shown to improve the relevance of research, increase stakeholder trust in research and researchers, enhance mutual learning between stakeholders and researchers, and improve research adoption [12,13].

Even with the apparent benefits of PCOR, the PCORI recognizes that before patient–researcher partnerships can work effectively and successfully, some level of initial training for both the researchers and the patient partners is necessary [10]. In 2013, the PCORI stated that patients and stakeholders need training "to have productive conversations with research partners" while researchers need training to "adopt a vernacular that is familiar to patients and stakeholders and facilitates best communication" [10]. We found several studies that evaluated PCOR programs and interactions, findings from which we identified priorities for PCOR training [14-16]. One PCOR training program for patient partners of the National Organization of Rare Disorders (NORD) in conjunction with the University of Maryland included PCOR funding opportunities, use of data sources to help support PCOR partnerships, different levels of patient engagement, and techniques for communication and collaboration [17]. An evaluation of a training program for new PCOR researchers conducted at the University of Maryland suggested the need for qualified and skilled mentors in PCOR methodology [15]. A separate study found that training priorities should include helping team members identify appropriate patient partners, devising an engagement strategy that clarifies roles and expectations, and building skills for positive team dynamics [14]. PCOR training for researchers and patients can support productive relationships that advance patient-centered outcomes research.

To better inform PCOR training, we performed an educational needs assessment to determine the interests and concerns of the CF community. Such assessments are considered fundamental to the success of training programs and to identifying potential gaps and discrepancies between learner types, in this case those between patients/caregivers and researchers/health care providers [18,19]. The aim of this study was to inform the development of a future PCOR training program for the CF community.

Methods

Study Design

We developed our needs assessment using a mixed-methods approach, which helps to strengthen a study's conclusions and provide greater validity to the findings [20]. We designed our needs assessment using both a quantitative survey and qualitative in-depth discussion groups to inquire about the PCOR training needs of the CF community. The survey sought to assess the overall interest of PCOR in the community and to understand respondents' perceived barriers and concerns about PCOR. We

followed up with these findings with 3 in-depth discussion groups, called World Café, to further explore these concepts. In both the survey and discussion sessions, we assessed PCOR knowledge and experience, possible PCOR training program topics, and potential training session formats. This study was provided exempt status by the University of Washington Human Subjects Division (Institutional Review Board no. 6146).

Patient Involvement

Within the CF community is a well-established organization built upon PCOR principles, the Cystic Fibrosis Reproductive and Sexual Health Collaborative (CFReSHC). Through PCORI pilot funding, CFReSHC was established in 2016 with a team of CF researchers, patients with CF, and reproductive health–trained family physicians, obstetricians, and gynecologists. The aim of this collaborative is to create and maintain an authentic coproduction partnership [21-23]. To accomplish this, the CFReSHC meaningfully engages patients and other stakeholders (such as clinicians, payers, and policy makers) throughout the research process, including the planning, conduct, and dissemination of the research [24-26]. CFReSHC members, including patients with CF, were involved in this and other studies' design, execution, interpretation of findings, dissemination, and authoring.

Participants

Our primary objective was to make the survey accessible to as many people as possible in the CF community. We invited a convenience sample of diverse members of the CF community to participate in our needs assessment, including patients with CF, caregivers, all members of the CF clinical care team, researchers, and research staff. Because CF is a complex multi-organ disease, the CF clinical care team is comprehensive and interdisciplinary in nature and consists of clinicians, nurses, respiratory therapists/physical therapists, social workers, and nutritionists. Many CF clinics have additional specialists in gastroenterology, endocrinology, complex pharmacy, advanced practice, occupational therapy, mental health, and care coordination.

Recruitment

Our initial phase of recruitment involved reaching out to organizations that serve the CF community and spreading the word about the survey through multiple communication platforms and social media channels. We contacted 4 prominent organizations in the CF community, including the CFF, CFRI, CFReSHC, and CF Roundtable. We also used CF community list servers through the CFF and thus reached each of the 130 existing care teams in the United States. Once the responses slowed to less than 1 response per week, we stopped recruitment. We invited survey respondents who provided their contact information to participate in 1 of 3 follow-up 1.5-hour World Café discussion sessions.

Survey

We developed an anonymous 35-question online survey based on questions from a prior needs assessment conducted through the University of Washington Institute of Translational Health Sciences (ITHS) Community Voices Program, which is a 5-state program that connects community organizations with academic researchers [27]. The questionnaire (Multimedia Appendix 1) included definitions from the PCORI about specific terms, including patient engagement, PCOR, and comparative effectiveness research in order to increase baseline knowledge for the respondents about the questions being asked. Questions included respondent identification to specific self-identified participant groups (patient/caregiver vs researcher/health care provider). We then tailored a few questions to the specified groups based on topic areas deemed as priority areas in other PCOR training sessions [14,15,17]. We asked about participant knowledge of and experience with PCOR, preferred PCOR topics, and formats for additional training and asked participants to rank the importance of their selection using a 4-point Likert scale: "extremely important," "important," "not important," or "not at all important." We used a 4-point Likert scale to ensure participants selected a nonneutral position and to reduce the complexity of responses, given the breadth of participants from whom we sought input [28,29]. Two questions were mandatory, including the type of participant and interest in participating in a follow-up discussion session, while the other questions did not require responses to move forward. Survey features included a "back" button allowing participants to review and change responses. Participants also had the option to save and return to the survey later using their specific survey link. Our educational specialist and 4 CF community team members reviewed and modified the survey to ensure relatedness and understandability of questions in our target population. Surveys were administered for 7 weeks in November and December of 2018. We collected data through Research Electronic Data Capture (REDCap; Vanderbilt University), an encrypted and secure, HIPAA (Health Insurance Portability and Accountability Act)-compliant, survey database hosted by the University of Washington ITHS [30]. The survey was in English, took approximately 15 minutes to complete, and was voluntary. At the end of the first survey, participants had the option to enter personal information in a second survey if they were interested in participating in a World Café discussion session. The second survey with personal information was not linked to the responses of the first survey.

Follow-up World Café Sessions

In January 2019, we conducted 3 separate in-depth discussion sessions to further explore survey respondents' preferences for PCOR training in the CF community. We structured our discussion sessions using the World Café methodology [31]. World Café is a consensus-building community participatory tool designed to allow several small-group conversations to take place at separate tables, with participants systematically rotating to different tables approximately every 20 minutes. World Café provides a setting in which community participants discuss diverse perspectives in order to generate new collective intelligence [32]. World Café was selected as the method with the understanding that the ideas gathered remain in the domain of the participants, not with the researchers. Because of the need for strict infection control in CF and the inability of affected individuals to convene in a single room [33], we conducted our World Café sessions online via videoconferencing using Zoom (Zoom Video Communications, Inc). The online video conferencing software allowed geographically dispersed

discussant participants to convene as a single large group and in concurrent, separate, smaller group discussions.

In keeping with World Café guidelines, we started each discussion session with a large-group 20-minute introduction to set the context and create a hospitable space. This included providing a project overview and reviewing the definition of engagement and benefits of PCOR. We then broke the larger group into 2 small-group discussions facilitated by a patient partner–researcher/clinician dyad experienced in PCOR methodology. One team member facilitated the discussion with questions, while the other took detailed notes. Patients/caregiver discussant participants met separately from researchers/health care providers to encourage sincere feedback and to minimize any power dynamics [34]. Each breakout room discussed their questions for 15 to 20 minutes, with facilitators encouraging everyone's contribution before switching to the other breakout room. We based topic discussion questions on the survey findings from the 170 participants from the CF community. Both breakout rooms discussed participants' interests in PCOR and motivations to encourage others to participate in PCOR. In one breakout room, participants discussed their concerns about participating in PCOR and likelihood of their attending trainings with other patients or researchers. In the other breakout room, participants discussed important topics and skills to incorporate into PCOR training for the CF community. At the start of each small breakout room discussion, participants were shown the responses to the questions the prior small group discussed, so that the new group was quickly brought up to speed on the issues raised in the prior discussion. At the end of each discussion session, the facilitators simultaneously reviewed the detailed breakout sessions notes and compiled and repeated to the group the ideas that emerged for the participants to review to ensure all ideas discussed during their session were captured. After each breakout room session was complete, both the patient/caregiver and researchers/health care provider groups were convened into a large group and asked to vote for their top 3 responses to each theme to create community consensus and define priorities. Voting was chosen because it allowed the final decision-making to remain in the hands of the participants and created a venue for consensus building after a series of discussions [35]. Incentives were offered to participants who participated in a World Café session.

Analysis

Because nonresearch stakeholders have not been traditionally part of CF research teams, we assumed the PCOR learning needs of patients and caregivers would be similar, and thus combined our findings for these 2 stakeholder groups. Similarly, we combined the findings of the researchers and health care provider groups as well.

Survey

We analyzed complete survey data using descriptive statistics to compare and contrast frequencies of responses between participant groups.

Follow-up World Café Sessions

We took the top 4 ideas within each theme as indicated by simple majority vote as priorities for a future PCOR training program. Quotations from the World Café discussion notes provided examples to illustrate and clarify priority areas.

Results

At the time of survey closure, we had a total of 170 responses. Among the respondents were 96 CF patients and caregivers, and 74 researchers, research staff, and health care providers, all of whom completed the online survey, with 22 participating in the follow-up World Café discussion sessions. We included all participants in the analysis.

Survey

More than half (52/96, 54%) of patients/caregivers and 86% (64/74) of researchers/health care providers had heard of PCOR, but only about one-third of both groups had ever previously participated in PCOR (Table 1).

Table 1. Participant roles and survey responses (N=170).

Participant role and responses	Value, n (%)
Survey participant role	
Patients/caregivers	96 (56)
Providers/researchers (n=74)	74 (44)
Physician/advanced practice provider	19 (26)
Nurse	15 (20)
Other[a]	40 (54)
Participated in 3 discussion sessions (n=22)	
Patients/caregivers	12 (55)
Providers/researchers	10 (46)
Previous experiences with PCOR[b]	
Participants who had heard of PCOR	116 (68)
Patients/caregivers	52 (54)
Providers/researchers	64 (86)
Participants who had participated in PCOR	57 (34)
Patients/caregivers	34 (35)
Providers/researchers	23 (31)

[a]Other provider/researcher roles included dietitians, respiratory therapists, psychologists, mental health coordinators, social workers, researchers, and research coordinators.

[b]PCOR: patient-centered outcomes research.

Almost 86.5% (147/170) of all respondents were interested in participating in PCOR. Before participating in PCOR, patients/caregivers wanted to understand more about expectations of them as partners on PCOR research teams (82/96, 85%; Table 2). Researchers/health care providers, in contrast, wanted to hear more about how to include outcomes important to patients/caregivers (55/74, 74%) and the quality and impact of PCOR research (52/74, 70% and 51/74, 69%, respectively; Table 2).

Table 2. Participant ratings of what to learn before engaging in patient-centered outcomes research.

Participant responses by role	Rating (N=170), n (%)			
	Extremely important	Important	Not important	Not at all important
Patients/caregivers[a]				
What the expectations are[b]	82 (85)	11 (11)	0 (0)	0 (0)
The benefits to you or your child	60 (62)	28 (29)	3 (3)	3 (3)
How to share your expertise	59 (61)	29 (30)	6 (6)	0 (0)
How to partner with researchers	52 (54)	40 (42)	2 (2)	0 (0)
How to establish open communication	61 (64)	30 (31)	2 (2)	1 (1)
Researchers/providers				
How to include PCOR[c] outcomes	55 (74)	18 (24)	1 (1)	0 (0)
The quality of PCOR	52 (70)	19 (26)	2 (3)	0 (0)
The impact of PCOR	51 (69)	21 (28)	1 (1)	0 (0)
How to navigate IRB[d]	33 (45)	39 (53)	2 (3)	0 (0)
How to identify PCOR research topics	48 (65)	23 (31)	3 (4)	0 (0)
How to design a PCOR study	32 (43)	34 (46)	7 (9)	0 (0)

[a]Two missing values are from respondents misclassified as researchers/providers who did not receive these questions.

[b]One missing value is from a participant who did not respond.

[c]PCOR: patient-centered outcomes research.

[d]IRB: institutional review board.

Desired PCOR training topics and training format differed between the patient/caregiver and researcher/health care provider groups (Figure 1). Patients/caregivers were most interested in learning about the time commitment as a PCOR team member (75/96, 78%). Researchers/health care providers wanted to equally receive training about how to establish trust (47/74, 64%) and maintain confidentiality (47/74, 64%) when patient or caregiver partners are on the PCOR team. The majority of patients/caregivers wanted to learn about PCOR using online interactive sessions (80/96, 83%), whereas researchers/health care providers preferred to have training at their CF center (54/74, 73%) or as a webinar (51/74, 69%; Table 3).

Figure 1. Desired topics to address in patient-centered outcomes research (PCOR) training by participant type.

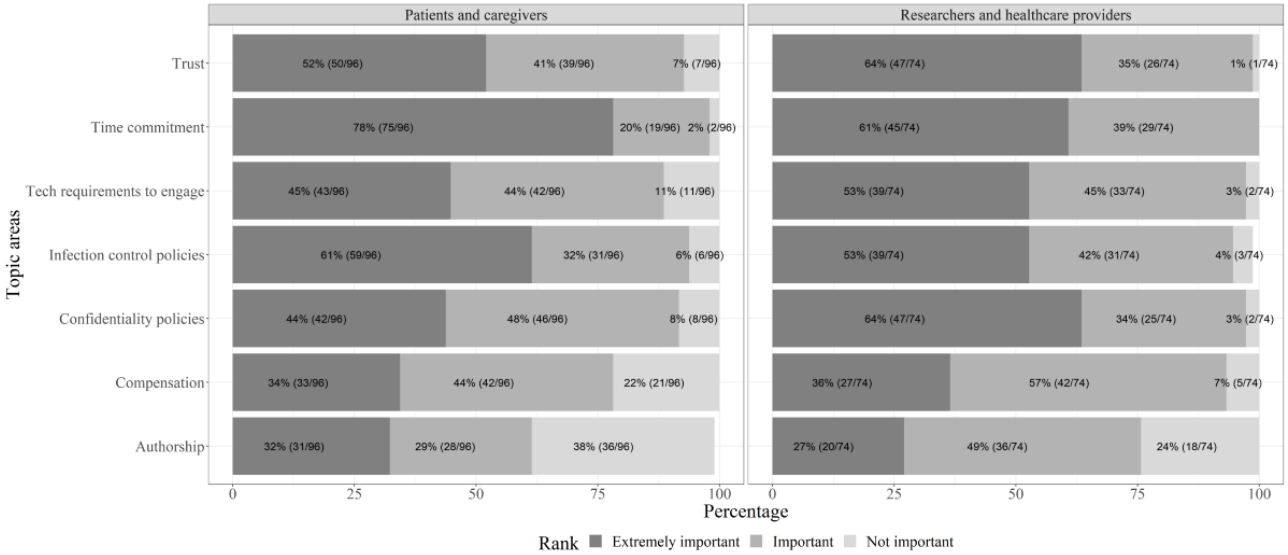


Table 3. Session format preferences[a].

Format	Patients/caregivers (n=96), n (%)	Researchers/providers (n=74), n (%)	Total (N=170), n (%)
Webinar	69 (72.0)	51 (68.9)	120 (70.6)
Onsite training at your CF[b] center	44 (45.8)	54 (73.0)	98 (57.6)
Online training	80 (83.3)	46 (62.2)	126 (74.1)
Self-directed learning	73 (76.0)	43 (58.1)	116 (68.2)

[a]Formats that participants indicated they were "likely" to participate in.
[b]CF: cystic fibrosis.

Follow-up World Café Sessions

We categorized the top 4 concerns about participating in PCOR by patients/caregivers and researchers/health care providers into 2 separate categories (Table 4).

Concerns About Participating in PCOR

Participants were concerned about ensuring PCOR groups have a diverse representation of patient partners and that everyone on the PCOR team would have a clear idea of their role. Other priorities to address included overcoming the existing power dynamic that exists between patients/caregivers and researchers/health care providers, defining roles, and keeping patient partners engaged throughout the research process.

Topics and Skills to Include in a PCOR Training Program

The top 4 topics and skills to include in a PCOR training program for the CF community included the illustration of good PCOR team dynamics (Table 4). For example, participants discussed solutions to this recommendation to include tips such as icebreaker games and team building exercises to help level the hierarchy and allow team members to get to know each other. Other topics participants indicated for a PCOR program included teaching members of the research team how to genuinely appreciate contributions of patient partners and other stakeholders as well as how to construct a transparent process and create a shared language. Participants thought that providing these tips in a reference manual would be important.

Table 4. Top 4 priorities related to concerns about engaging in PCOR and PCOR training topics and skills from World Café discussion sessions.

Issue	Quotation to illustrate issue (source)
Top concerns about engaging in PCOR[a]	
How to engage a representative group of patients	"The families who are able to partner with us may represent a more resourced family. Participants that need the most support may be unable to participate." (researcher/healthcare provider, discussion session 3)
How to create defined roles for patients/caregivers	"[It's important for patients to have] a defined role. Sometimes it is not clear to patients why [they] are there and what [they're] supposed to be doing." (patient/caregiver, discussion session 1)
Level the team hierarchy	"Clinicians often can't really figure out where the patient fits in. Patients often defer to the clinicians because they don't have the confidence to speak up." (researcher/health care provider, discussion session 1)
How to retain patients as team members throughout the research project	"Making sure [patients'] role is meaningful, integral and acknowledged. Make sure it isn't tokenism. Inviting patients into initial study planning but not including them in later study design or data analysis or acknowledging them in research products [is tokenism]." (patient/caregiver, discussion session 2)
Top priorities for a PCOR training program	
Provide examples to explain PCOR	"[A PCOR 101 training program should include] a combination of personal testimony, to get investigators engaged, and a nuts and bolts manual to refer to after the training is over." (researcher/provider, discussion session 1)
Teach how to appreciate all PCOR team members	"Help providers understand that patients are the experts in their own story." (patient/caregiver, discussion session 3)
Demonstrate how to construct transparency	"[There is] suspicion in the community because there is no feedback loop and no sustained benefit from coming to the table." (patient/caregiver, discussion session 3)
Create a shared language	"[Training] should be done jointly. [I worry] about medical jargon used by medical personnel, but patients know much of that too. [A joint training session] will open up a dialog—doctors would actually hear patient concerns." (patient/caregiver, discussion session 2)

[a]PCOR: patient-centered outcomes research.

Training Logistics

Across all 3 discussion sessions, most participants (4/5, 80%; 3/5, 60%; 7/9, 77%) thought the 2 separate learner groups (patients/caregivers and researchers/providers) should be trained together to mimic PCOR, with patients and researchers/health care providers working as equal members of the research team. They also thought that these 2 learner groups may have specific learning needs for which training could take place separately, such as patients/caregivers needing to learn basic research concepts and researchers/providers needing to learn skills on group facilitation and creating an inclusive working environment. Only 5% (1/19) thought each learner group should be trained entirely as separate groups.

Discussion

Principal Findings

We performed this educational needs assessment as a first step in designing a PCOR training program for the CF community. The majority of respondents reported being interested in participating in PCOR. Most participants desired joint patient/caregiver and research/provider learning sessions, except in cases where each group had unique learning needs such as training patients in research fundamentals. Respondents acknowledged that there is currently no formal PCOR training and that such a training program would fill a methodological need for the CF community.

Training priorities expressed by our participants were similar to prior research [14,17]. In our study, both patient/caregiver and researchers/providers wanted to know how to partner with one another, indicating the importance of creating an engagement strategy. In the qualitative study of diverse PCOR team members, Lavallee et al [14] notes that depending on the level at which patient and other stakeholders participate on the team (ie, collaborator vs consultant), researcher training should include tips about taking additional time to build trust, clarifying roles, and ensuring that patient input is not limited once the patient agrees to participate. PCOR training for patient partners conducted by NORD similarly devoted 2 sessions to outlining patient engagement throughout the research project and emphasizing the importance of defining roles by using case examples from rare diseases, including CF [17]. These training sessions included team dynamics such as transparency, bidirectional learning, and developing a structure for collaboration [14,17], which were also highlighted by our World Café participants. Our follow-up World Café discussions emphasized the importance of providing training that shows

"how" to partner with one another, similar to how the University of Maryland has used experienced PCOR mentors for new PCOR teams [15]. PCOR guidelines developed by the American Thoracic Society (ATS), a professional organization to which many CF specialists belong, also touch on supporting new PCOR teams with "how" to do PCOR by recommending that researchers create a mechanism in which to share lessons with one another [36].

Building upon prior research, our finding highlighted concerns about apparent hierarchical issues that exist between clinicians/researchers and patients in the CF community. The ATS guidelines suggest the importance of teams developing processes in which perspectives are balanced to help reach consensus and fostering a collaborative spirit from the start, thus mitigating hierarchy [36]. Through our literature review, we found that PCOR teams that were deemed to be successful included members with excellent facilitation and communication skills and incorporated evidence-based strategies to achieve the teams' aims and outcomes [11]. Similarly, a mixed methods study of PCORI pilot project awardees found that successful PCOR teams require strong relationships between members, engagement expertise, and institutional leadership that supports PCOR [37]. A separate qualitative study consisting of a large hospital research collaborative also stressed the importance of building supportive environments between patients, families, and researchers [38]. Despite not being specifically mentioned by our participants, a number of articles underscore the importance of including financial support for researchers specifically to build and maintain PCOR teams [37,39].

The CF community is well positioned to build capacity in PCOR, and even prior to the era of COVID-19, people with CF were well versed in engaging online. This study successfully demonstrates the ability to incorporate community engagement and mixed methods research virtually. Additionally, previously mentioned, many people with CF and their caregivers are already active in stakeholder-sponsored research events through patient advocacy organizations such as the CFF and CFRI. The CFF currently trains and onboards new patient partners and family members in clinical quality improvement work. In organizations like the CFReSHC, patients with CF are self-stewards of research proposal ideas, and they partner with academicians, stakeholder organizations, and health care providers as equals on research teams. The members of the CFReSHC are both willing and well situated to assist research teams new to PCOR with coaching and guidance as needed.

Limitations

There were several limitations to this study. Due to our community-wide recruitment, we are unable to report our actual response rate. Rather, our findings are limited to a convenience sample and may not be representative of the learning needs or desires of everyone in the CF community. Additionally, our World Café discussion group participants were limited to those who responded to the survey. We do not believe this was a critical flaw because the survey opportunity likely helped build rapport with the researchers and allowed us to select participants who were eager to reflect on various issues in PCOR [40]. Follow-up questions to further assess learning gaps are needed to further confirm our findings.

Conclusions

The majority of respondents in the CF community are interested in PCOR. A PCOR training program would fill a current methodological research gap in the CF community. The results of this needs assessment were used to create a pilot PCOR training program for the CF community.

Acknowledgments

This work was funded in part through the PCORI Eugene Washington Program Award (no. 10569-UWASH) and the National Center for Advancing Translational Sciences of the National Institutes of Health (no. UL1 TR002319). The content is solely the responsibility of the authors and does not necessarily represent the official views of the PCORI or the National Institutes of Health. The funders had no role in the design and conduct of the study, collection, management, analysis, interpretation of the data, preparation, review, or decision to submit the findings for publication. The sponsor reviewed the manuscript only for proper acknowledgement of the organization. The authors would also like to thank our stakeholders, the CFF, CFRI, CF Roundtable, and the CFReSHC. We also thank members of the CF community who participated in the survey and follow-up World Café discussion groups.

Authors' Contributions

EMG was the principal investigator for the study and conceived and designed the study. TMK, LM, GB, EKT, MAA, and MP contributed to study design. EKT performed the analysis. All authors contributed to manuscript preparation and critically reviewed each draft. All authors approved the final manuscript for publication.

Conflicts of Interest

TMK receives grant funding from and is a consultant for the Cystic Fibrosis Foundation.

Multimedia Appendix 1
Online needs assessment survey for the cystic fibrosis community.
[PDF File (Adobe PDF File), 1651 KB - formative_v5i3e24302_app1.pdf]

References

1. Rowe SM, Miller S, Sorscher EJ. Cystic fibrosis. N Engl J Med 2005 May 12;352(19):1992-2001. [doi: 10.1056/NEJMra043184] [Medline: 15888700]
2. Cystic Fibrosis Foundation Patient Registry. 2019 annual data report. Cystic Fibrosis Foundation. Bethesda MD; 2019. URL: https://www.cff.org/Research/Researcher-Resources/Patient-Registry/ [accessed 2021-02-19]
3. Ramsey BW, Welsh MJ. AJRCCM: 100-Year Anniversary. Progress along the pathway of discovery leading to treatment and cure of cystic fibrosis. Am J Respir Crit Care Med 2017 May;195(9):1092-1099. [doi: 10.1164/rccm.201702-0266ed] [Medline: 28459323]
4. Aschenbrenner DS. New treatment for cystic fibrosis. Am J Nurs 2020 Feb;120(2):20-21. [doi: 10.1097/01.NAJ.0000654292.06762.bd]
5. Cystic Fibrosis Foundation. Community voice brings your voice into everything that we do. Cystic Fibrosis Foundation. Bethesda, MD URL: https://www.cff.org/CF-Community-Blog/Posts/2017/Community-Voice-Brings-Your-Voice-Into-Everything-That-We-Do/ [accessed 2018-04-10]
6. Help shape CF research. Cystic Fibrosis Foundation. Bethesda, MD URL: https://www.cff.org/Get-Involved/Connect/Help-Shape-CF-Research/ [accessed 2020-08-01]
7. About Cystic Fibrosis Research, Inc. Cystic Fibrosis Research Inc. Palo Alto, CA URL: http://www.cfri.org/home.shtml [accessed 2018-04-09]
8. The value of engagement. Patient-Centered Outcomes Research Institute. Washington, DC; 2021. URL: https://www.pcori.org/engagement/value-engagement [accessed 2021-02-19]
9. Frank L, Basch E, Selby JV. The PCORI perspective on patient-centered outcomes research. JAMA 2014 Oct 15;312(15):1513. [doi: 10.1001/jama.2014.11100] [Medline: 25167382]
10. Fleurence R, Selby JV, Odom-Walker K, Hunt G, Meltzer D, Slutsky JR, et al. How the Patient-Centered Outcomes Research Institute is engaging patients and others in shaping its research agenda. Health Aff (Millwood) 2013 Feb;32(2):393-400 [FREE Full text] [doi: 10.1377/hlthaff.2012.1176] [Medline: 23381533]
11. Oliver K, Kothari A, Mays N. The dark side of coproduction: do the costs outweigh the benefits for health research? Health Res Policy Syst 2019 Mar 28;17(1):33 [FREE Full text] [doi: 10.1186/s12961-019-0432-3] [Medline: 30922339]
12. Concannon TW, Fuster M, Saunders T, Patel K, Wong JB, Leslie LK, et al. A systematic review of stakeholder engagement in comparative effectiveness and patient-centered outcomes research. J Gen Intern Med 2014 Jun 4;29(12):1692-1701. [doi: 10.1007/s11606-014-2878-x] [Medline: 24893581]
13. Forsythe LP, Carman KL, Szydlowski V, Fayish L, Davidson L, Hickam DH, et al. Patient engagement in research: early findings from the Patient-Centered Outcomes Research Institute. Health Affairs 2019 Mar;38(3):359-367. [doi: 10.1377/hlthaff.2018.05067] [Medline: 30830822]
14. Lavallee D, Gore J, Lawrence S. Initiative to support patient involvement in research (INSPIRE). Patient-Centered Outcomes Research Institute. Washington, DC; 2016 Jul 18. URL: https://www.pcori.org/research-results/2014/initiative-support-patient-involvement-research-inspire [accessed 2020-08-01]
15. Huang J, Lipman PD, Daniel Mullins C. Bridging the divide: building infrastructure to support community-academic partnerships and improve capacity to conduct patient-centered outcomes research. Behav. Med. Pract. Policy Res 2017 Apr 3;7(4):773-782. [doi: 10.1007/s13142-017-0487-z] [Medline: 28374212]
16. Oehrlein EM, Love TR, Anyanwu C, Hanna ML, Kraska J, Perfetto EM. Multi-method patient-engagement approach: a case example from a PCORI-funded training project. Patient 2019 Apr;12(2):277-280. [doi: 10.1007/s40271-018-0341-5] [Medline: 30390238]
17. PCOR training: a program for rare disease patient advocates. Patient-Centered Outcomes Research Institute. Washington, DC; 2019. URL: https://www.pcori.org/research-results/2015/pcor-training-program-rare-disease-patient-advocates [accessed 2020-10-20]
18. Ambrose S, Bridges M, DiPietro M, Lovett M, Norman M, Mayer R. How Learning Works: Seven Research-Based Principles for Smart Teaching. Hoboken, NJ: John Wiley & Sons, Inc; 2010.
19. Brown J. Training needs assessment: a must for developing an effective training program. Public Personnel Management 2002 Dec 01;31(4):569-578. [doi: 10.1177/009102600203100412]
20. Schoonenboom J, Johnson RB. How to construct a mixed methods research design. Köln Z Soziol Sozpsychol 2017 Jul 5;69(S2):107-131. [doi: 10.1007/s11577-017-0454-1] [Medline: 28989188]
21. Sufian S, Mentch L, Godfrey E. Women with CF are helping guide reproductive and sexual health research. CF Roundtable. Beaverton, OR: United States Adult Cystic Fibrosis Association; 2017 Jul. URL: https://www.cfroundtable.com/newsletter [accessed 2020-08-10]
22. Godfrey E, Kazmerski T, Thayer E, Brown G, Mentch L, Pam M. Determining the needs of a patient-centered outcomes research training program for the cystic fibrosis community. In: Pediatric pulmonology.: Cystic Fibrosis Foundation; 2019 Presented at: North American Cystic Fibrosis Conference; October 31-November 2, 2019; Nashville, TN p. S462-S462.
23. Godfrey E, Thayer E, Mentch L, Al-Achkar M, Kazmerski T, Brown G. Developing an effective training program to build patient-centered outcomes research capacity within the cystic fibrosis community. In: Pediatric pulmonology.: Cystic

Fibrosis Foundation; 2019 Presented at: North American Cystic Fibrosis Conference; October 31-November 2, 2019; Nashville, TN p. S209-S209.
24. Largent EA, Weissman JS, Gupta A, Abraham M, Rozenblum R, Lynch HF, et al. Patient-centered outcomes research: stakeholder perspectives and ethical and regulatory oversight issues. IRB: Ethics & Human Research 2018 Jan 01;40(1):7-17. [doi: 10.1002/eahr.401002] [Medline: 30631218]
25. Kazmerski T, Ladores SL, Jain R, Mody S, Pam M, Brown G, et al. The Cystic Fibrosis Reproductive and Sexual Health Collaborative: Next steps in building a sustainable partnership. In: Pediatric pulmonology.: Cystic Fibrosis Foundation; 2017 Presented at: North American Cystic Fibrosis Conference; Nov 2-4, 2017; Indianapolis, IN p. S481-S481.
26. Kazmerski T, Josephy T, Sufian S, Gilmore KC, Jain R, Ladores SL, et al. Advancing comprehensive care through the interdisciplinary cystic fibrosis reproductive and sexual health collaborative. In: Pediatric pulmonology.: Cystic Fibrosis Foundation; 2018 Presented at: North American Cystic Fibrosis Conference; Oct 18-20, 2018; Denver, CO p. 412-412.
27. The Community Voices Program. Institute of Translational Health Sciences. Seattle, WA; 2018. URL: https://www.iths.org/community/partners/community-voices-program/ [accessed 2020-08-10]
28. Lietz P. Research into questionnaire design: a summary of the literature. International Journal of Market Research 2010 Mar 01;52(2):249-272. [doi: 10.2501/S147078530920120X]
29. Losby J, Wetmore A. Using Likert scales in evaluation survey work. Centers for Disease Control and Prevention. Atlanta, GA; 2012 Feb 14. URL: https://www.cdc.gov/dhdsp/pubs/docs/cb_february_14_2012.pdf [accessed 2020-10-10]
30. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
31. Brown J, Isaacs D, Cafe CW. Shaping our futures through conversations that matter. In: The World Café. Oakland, US: Berrett-Koehler Publishers; 2005.
32. MacFarlane A, Galvin R, O'Sullivan M, McInerney C, Meagher E, Burke D, et al. Participatory methods for research prioritization in primary care: an analysis of the World Café approach in Ireland and the USA. Fam Pract 2016 Sep 27;34(3):278-284. [doi: 10.1093/fampra/cmw104] [Medline: 27677298]
33. Gifford A, Stoudemire W, Saiman L. Infection prevention and control clinical care guidelines. Cystic Fibrosis Foundation. Bethesda, MD; 2019. URL: https://www.cff.org/Care/Clinical-Care-Guidelines/Infection-Prevention-and-Control-Clinical-Care-Guidelines/Infection-Prevention-and-Control-Clinical-Care-Guidelines/ [accessed 2020-08-17]
34. Harrison JD, Anderson WG, Fagan M, Robinson E, Schnipper J, Symczak G, et al. Patient and family advisory councils (PFACs): identifying challenges and solutions to support engagement in research. Patient 2018 Aug;11(4):413-423. [doi: 10.1007/s40271-018-0298-4] [Medline: 29392529]
35. Making decisions. The Community Toolbox. Center for Community Health and Development: University of Kansas; 2020. URL: https://ctb.ku.edu/en/table-of-contents/leadership/leadership-functions/make-decisions/main [accessed 2020-10-30]
36. Feemster LC, Saft HL, Bartlett SJ, Parthasarathy S, Barnes T, Calverley P, American Thoracic Society Behavioral Sciences Health Services Research Nursing Assembly. Patient-centered outcomes research in pulmonary, critical care, and sleep medicine. An official American Thoracic Society workshop report. Ann Am Thorac Soc 2018 Sep;15(9):1005-1015. [doi: 10.1513/AnnalsATS.201806-406WS] [Medline: 30168741]
37. Nease DE, Burton D, Cutrona SL, Edmundson L, Krist AH, Laws MB, et al. "Our lab is the community": defining essential supporting infrastructure in engagement research. J Clin Transl Sci 2018 Aug;2(4):228-233. [doi: 10.1017/cts.2018.325] [Medline: 30800479]
38. Harrison JD, Anderson WG, Fagan M, Robinson E, Schnipper J, Symczak G, et al. Patient and family advisory councils for research: recruiting and supporting members from diverse and hard-to-reach communities. J Nurs Adm 2019 Oct;49(10):473-479. [doi: 10.1097/NNA.0000000000000790] [Medline: 31490796]
39. LeClair AM, Kotzias V, Garlick J, Cole AM, Kwon SC, Lightfoot A, et al. Facilitating stakeholder engagement in early stage translational research. PLoS One 2020;15(7):e0235400 [FREE Full text] [doi: 10.1371/journal.pone.0235400] [Medline: 32614885]
40. Conducting focus group interviews. USAID Center for Development Information and Evaluation. Washington, DC; 1996. URL: https://www.ndi.org/sites/default/files/USAID%20Guide_Conducting%20Focus%20Groups.pdf [accessed 2020-10-18]

Abbreviations

ATS: American Thoracic Society
CF: cystic fibrosis
CFF: Cystic Fibrosis Foundation
CFReSHC: Cystic Fibrosis Reproductive and Sexual Health Collaborative
CFRI: Cystic Fibrosis Research, Incorporated
HIPAA: Health Insurance Portability and Accountability Act
ITHS: Institute of Translational Health Sciences

NORD: National Organization of Rare Disorders
PCOR: patient-centered outcomes research
PCORI: Patient-Centered Outcomes Research Institute
REDCap: Research Electronic Data Capture

Edited by G Eysenbach; submitted 14.09.20; peer-reviewed by E Schwind, K Mrklas; comments to author 07.10.20; revised version received 04.11.20; accepted 17.01.21; published 04.03.21.

Please cite as:
Godfrey EM, Kazmerski TM, Brown G, Thayer EK, Mentch L, Pam M, Al Achkar M
Educational Needs and Preferences for Patient-Centered Outcomes Research in the Cystic Fibrosis Community: Mixed Methods Study
JMIR Form Res 2021;5(3):e24302
URL: https://formative.jmir.org/2021/3/e24302
doi:10.2196/24302
PMID:33661127

Original Paper

An Acute Stress Scale for Health Care Professionals Caring for Patients With COVID-19: Validation Study

Jose Joaquin Mira[1,2,3*], PhD; Angel Cobos[4], MSc; Olga Martínez García[5], MSc; María José Bueno Domínguez[6], RNC; María Pilar Astier-Peña[7], PhD; Pastora Pérez Pérez[8], PhD; Irene Carrillo[1,2,3], PhD; Mercedes Guilabert[2], PhD; Virtudes Perez-Jover[2], PhD; Cesar Fernandez[9], PhD; María Asuncion Vicente[9], PhD; Matilde Lahera-Martin[10], BSc; Carmen Silvestre Busto[11], MSc; Susana Lorenzo Martínez[12], PhD; Ascension Sanchez Martinez[13], PhD; Jimmy Martin-Delgado[1,3], MSc; Aurora Mula[1], BSc; Barbara Marco-Gomez[14], MSc; Cristina Abad Bouzan[14], MSc; Carlos Aibar-Remon[15], PhD; Jesus Aranaz-Andres[16], PhD; SARS-CoV-2 Second Victims Working Group[17*]

[1]Atenea Research Group, Foundation for the Promotion of Health and Biomedical Research, Alicante, Spain
[2]Department of Health Psychology, Miguel Hernandez University, Elche, Spain
[3]European Research Network on Second Victims (COST Action CA-19113), Alicante, Spain
[4]Intensive Care Unit, San Cecilio Clinical University Hospital, Granada, Spain
[5]Psychiatry Department, San Cecilio Clinical University Hospital, Granada, Spain
[6]Department of Patient Safety, Sagessa Group, Reus, Spain
[7]Centro de Salud la Jota, Zaragoza, Spain
[8]Spanish Society for Quality Care, Oviedo, Spain
[9]Telematics Engineering Area, Miguel Hernandez University, Elche, Spain
[10]Department of Occupational Health, Osasunbidea, Pamplona, Spain
[11]Patient Safety Department, Osasunbidea, Pamplona, Spain
[12]Department of Quality and Patient Management, Alcorcon Foundation University Hospital, Madrid, Spain
[13]Department of Quality of Care, General University Hospital Reina Sofía, Murcia, Spain
[14]Mental Health Department, Calatayud Health Area, Zaragoza, Spain
[15]Department of Preventive Medicine, Lozano Blesa Univerisity Clinical Hospital, Zaragoza, Spain
[16]Department of Preventive Medicine, Ramón y Cajal University Hospital, Madrid, Spain
[17]Be+ Against COVID, Alicante, Spain
[*]these authors contributed equally

Corresponding Author:
Jose Joaquin Mira, PhD
Atenea Research Group
Foundation for the Promotion of Health and Biomedical Research
Hermanos López de Osaba
Alicante
Spain
Phone: 34 606433599
Email: jose.mira@umh.es

Abstract

Background: The COVID-19 pandemic has affected the response capacity of the health care workforce, and health care professionals have been experiencing acute stress reactions since the beginning of the pandemic. In Spain, the first wave was particularly severe among the population and health care professionals, many of whom were infected. These professionals required initial psychological supports that were gradual and in line with their conditions.

Objective: In the early days of the pandemic in Spain (March 2020), this study aimed to design and validate a scale to measure acute stress experienced by the health care workforce during the care of patients with COVID-19: the Self-applied Acute Stress Scale (EASE).

Methods: Item development, scale development, and scale evaluation were considered. Qualitative research was conducted to produce the initial pool of items, assure their legibility, and assess the validity of the content. Internal consistency was calculated

using Cronbach α and McDonald ω. Confirmatory factor analysis and the Mann-Whitney-Wilcoxon test were used to assess construct validity. Linear regression was applied to assess criterion validity. Back-translation methodology was used to translate the scale into Portuguese and English.

Results: A total of 228 health professionals from the Spanish public health system responded to the 10 items of the EASE scale. Internal consistency was .87 (McDonald ω). Goodness-of-fit indices confirmed a two-factor structure, explaining 55% of the variance. As expected, the highest level of stress was found among professionals working in health services where a higher number of deaths from COVID-19 occurred (P<.05).

Conclusions: The EASE scale was shown to have adequate metric properties regarding consistency and construct validity. The EASE scale could be used to determine the levels of acute stress among the health care workforce in order to give them proportional support according to their needs during emergency conditions, such as the COVID-19 pandemic.

(JMIR Form Res 2021;5(3):e27107) doi:10.2196/27107

KEYWORDS

SARS-CoV-2 virus; COVID-19 outbreak; medical staff; acute stress; moral injury; posttraumatic stress; COVID-19

Introduction

From the beginning of the COVID-19 pandemic, the rate of reproduction, lethality, and social alarm that accompanies SARS-CoV-2 [1,2] poses a challenge to health systems in all countries. The impact of the COVID-19 pandemic has been devastating for infected and uninfected patients and societies, as well as for the health care workforces who have been considered second victims of SARS-CoV-2.

Although we have more information about the new coronavirus and health institutions are now more capable of responding to patients' needs, this was not the case at first. At that time, in addition to the initial clinical uncertainty, the shortage of equipment, and the difficulties in maintaining the supply chain, there were constant changes in instructions, the interruption of all nondelayable care, and the experience of isolation of admitted patients, who died alone in more cases than could be admitted. At that time, there was a perception that the uncertainty about COVID-19 patients' evolution, care pressure, and lack of means affected the emotional balance of health professionals. The most observed responses were related to acute stress, moral injury, and compassion fatigue [3-6].

The Spanish health system, in particular, has been overwhelmed by the number of patients with COVID-19; during the first wave of the outbreak, as of May 26, 2020, there had been 236,259 cases and 27,117 deaths. In addition, 20.4% of COVID-19 patients were health professionals [7]. At that time, the physical and mental effort involved in caring for patients with COVID-19 has caused acute stress, compassion fatigue, and other affective pathologies, which, together with psychosomatic reactions, have affected work morale [4,8,9]. Integrated care has been jeopardized. Health care professionals have been reassigned to areas where they have no expertise or preparation, protocols have been made overnight and are continuously changing, and continuity of care was interrupted at all levels. Under these conditions, health care has been compromised [10]. Moreover, a significant proportion of these professionals were at risk, since most were reluctant to seek help to face affective and anxiety disorders [11].

In these circumstances, the priority was to offer supporting tools to frontline health care providers, including resources they could access to support them in their feelings of being overwhelmed and down. These resources had to be a quick response to the needs identified in informal conversations and in the proposals and experience of leading professionals in organizations who demanded support to deal with everyday stressful situations. Within this framework, the project *Be+ against COVID* emerged, given a reluctance by health care providers to request support; the idea was to develop a self-assessment tool to measure acute stress levels adjusted to the experiences that the professionals described having lived.

The aim of this study was to show the design and validation process of a self-applied acute stress scale for people who worked in the direct care of patients with COVID-19 during the early period of the outbreak; this scale is called the Self-applied Acute Stress Scale (EASE). This instrument focused on facilitating awareness of the level of stress endured by health care professionals; the instrument would then be used to assess the impact of organizational changes in health care systems on coping with the acute stress caused by (1) the limited resources available to treat patients during the early period of this pandemic, (2) uncertainty about the appropriate treatment of COVID-19 patients, (3) the risk of becoming infected in the course of care, and (4) the interrupted care for non-COVID-19 patients. As a result, this could contribute to secondary prevention of emotional and anxiety responses once the critical phase of the pandemic is over.

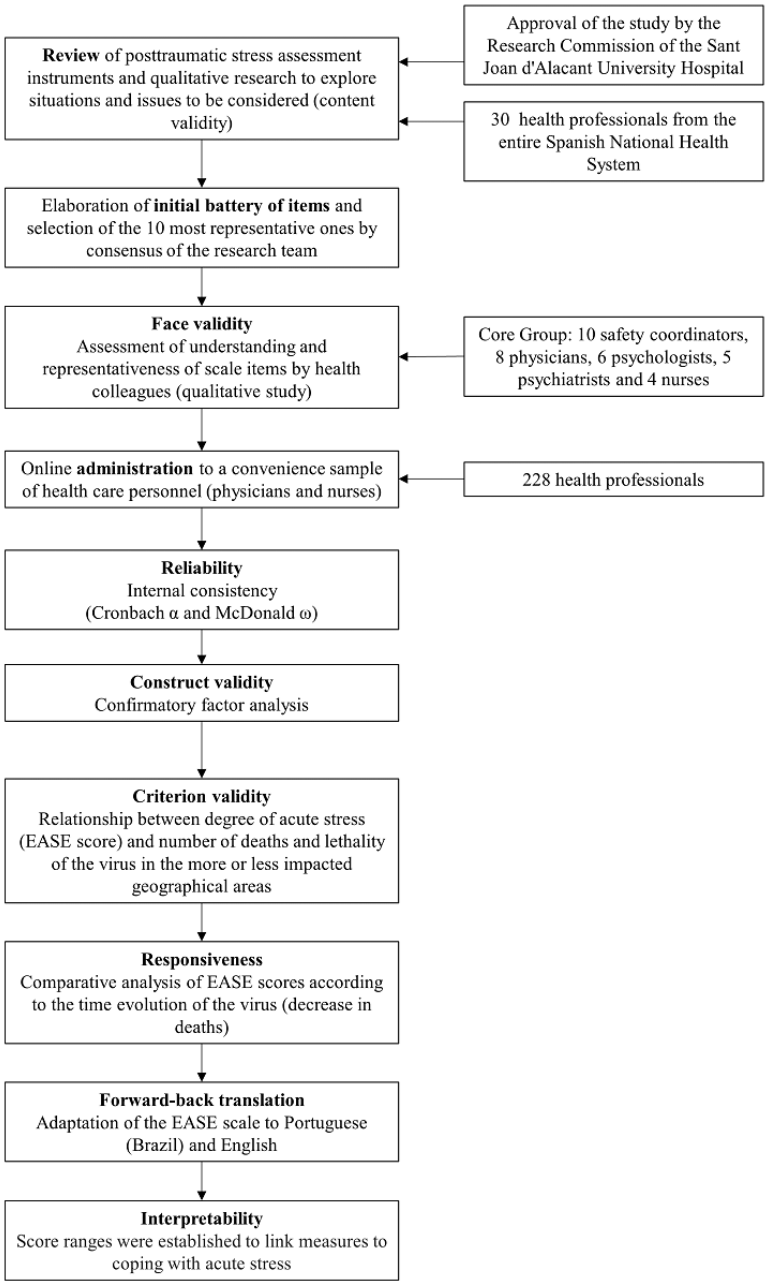
Methods

Overview

This study was conducted in Spain between March 20 and April 19, 2020, and coincided with the worst moments of the first wave of the COVID-19 outbreak. Health care providers could complete the EASE scale on a website or via a mobile app (see Figure 1). This study's protocol considered the three phases to creating a scale that were described by Boateng et al (ie, item development, scale development, and scale evaluation) [12], the standards and guidelines for validation practices summarized by Chan [13], and the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) recommendations [14]. This study was approved by the Research Commission of the Sant Joan d'Alacant University Hospital.

The target population of this instrument was the health care workforce, which includes physicians, nurses, and other health care personnel. The rationale for developing a new instrument was based on the demanding circumstances that were being experienced during the pandemic. Another reason was to develop a short tool, to be answered in less than 5 minutes, that identified the problematic situations being experienced during the early period of the pandemic by health care professionals to motivate them to look for support in facing the affective- and anxiety-related situations described. Other instruments, such as general scales for posttraumatic stress disorder, anxiety, and depression, were discarded due to their focus on mental health and because they did not include items reflecting the experiences that were observed to be more frequent during the treatment of COVID-19 patients [12,13].

Figure 1. Flow diagram of the Self-applied Acute Stress Scale (EASE) validation process.



Definitions

Acute stress was defined as an intense, unpleasant, and dysfunctional reaction beginning shortly after an overwhelming traumatic event and lasting less than a month due to a particularly stressful event. Acute stress responses may be adaptive but, in other cases, may impact well-being over time [15].

Participants

A convenience sample of physicians, nurses, and other health care personnel were recruited. The sample size was adjusted according to the number of items (ie, maximum of 10), applying the criterion of a minimum of 20 subjects with valid responses per item. To preserve complete anonymity, no information on age or experience in the workplace was requested.

Content Validity

A pragmatic literature review of empirical articles, letters, and reviews describing the experiences of the health care workforce was conducted. Additionally, eight physicians, four nurses, five psychiatrists, six psychologists, and 10 safety personnel from 10 hospitals and four health centers (Core Group) participated in collecting information during informal interviews with their colleagues. All of this information was grouped into categories according to the similarity of problematic situations. In this way, the sets of problematic situations and experiences were mapped and grouped into two theoretical factors: affective responses and anxiety responses (ie, content validity).

Item Generation

Two to three items were developed for each problem situation by the research team considering the results of the literature review—regarding previous SARS, Middle East respiratory syndrome (MERS), and Ebola outbreaks and studies in progress on the impact of the COVID-19 outbreak—and the emotional experiences and signs of acute stress reported by professionals in health care centers where the team members were working. Items measuring acute stress in health care professionals were assessed for possible inclusion. The successive approximations method was applied. In successive consensus rounds, the items considered most appropriate were either discarded or selected.

Face Validity and Legibility

The resulting scale was assessed by the Core Group considering its representativeness (ie, face validity). The Core Group also assessed the understanding of each item (ie, legibility) and the response options on a 4-point scale.

Reliability

Internal consistency was calculated using Cronbach α and McDonald ω. A value greater than .70 was considered acceptable for this analysis.

Construct Validity

Confirmatory factor analysis (CFA) was used to confirm the underlying two-factor structure, estimating several fit indices to test which CFA model best represented the data set: the comparative fix index, Jöreskog and Sörbom's adjusted goodness-of-fit index, the standardized root mean square residual, Jöreskog and Sörbom's goodness-of-fit index, and the normed fit index. Additionally, we tested the hypothesis stating that scores on the EASE scale would be higher for those professionals working in centers located in territories with higher mortality from COVID-19. Using the Mann-Whitney-Wilcoxon test and linear regression analysis (ie, the *Enter* method), the capacity to classify the professionals' responses in the EASE scale was assessed according to the number of deaths registered in the geographical area where the health center was located during the period from March 10 to April 19, 2020.

Translatability

The translation-back-translation method was used to ensure language and cultural equivalence between the Portuguese and English versions of the scale.

Criterion Validity

Criterion validity was determined by the scale's ability to discriminate between levels of acute stress over time, hypothesizing that it would be greater in those cases with higher care burden. We compared the EASE scale scores during two different periods of the pandemic's evolution—March 25 to April 1, 2020, and April 14 to 19, 2020—which were marked by different daily numbers of cases and deaths by COVID-19 (ie, between 800 and 900 deaths/day versus <500 deaths/day, respectively). The data published by the Ministry of Health on April 22, 2020, were taken as a reference.

Interpretability

To determine the impact on well-being, a consensus among researchers was established to understand scores using four segments: scores up to the 50th percentile, scores between the 50th and 80th percentile, scores between the 80th and 90th percentile, and scores above the 90th percentile.

Results

Overview

The first 228 responses from health care professionals were included in the database and used to evaluate the scale. A total of 42.1% (96/228) of respondents were physicians, 28.1% (64/228) were nurses, and 29.8% (68/228) were health support staff (ie, nursing assistants and care attendants). They mostly worked in Madrid (66/228, 28.9%), Andalusia (50/228, 21.9%), Valencia (40/228, 17.5%), and Catalonia (15/228, 6.6%). All key groups were adequately represented. The subsample of physicians was slightly overrepresented considering the proportion of physicians on staff to nurses. The responses originated from both territories—one with the highest and one with the lowest incidence of COVID-19—proportional to where the incidence of new cases of COVID-19 was high and moderate.

Content Validity, Face Validity, and Legibility

The most relevant sources of acute stress identified from the professionals' experiences were constant changes in instructions, shortage of material to avoid contagion, reduction in the number of staff due to the risk exposure or contagion, bitter feelings when seeing patients die lonely, fear of infecting their families, making decisions reserved for situations of major catastrophes with a high component of ethical conflict, and the passing away of colleagues. These issues were represented by an initial set of 17 possible items on the EASE scale. This number was finally reduced to 10 items (ie, version 0 of the scale) once participants considered their representativeness and comprehension (see Multimedia Appendix 1).

Reliability

Internal consistency values were adequate—Cronbach α=.85 and McDonald ω=.87—when all items of the EASE scale were considered. For the affective response factor, Cronbach α=.81 and McDonald ω=.81; for the fear and anxiety factor, Cronbach α=.73 and McDonald ω=.74.

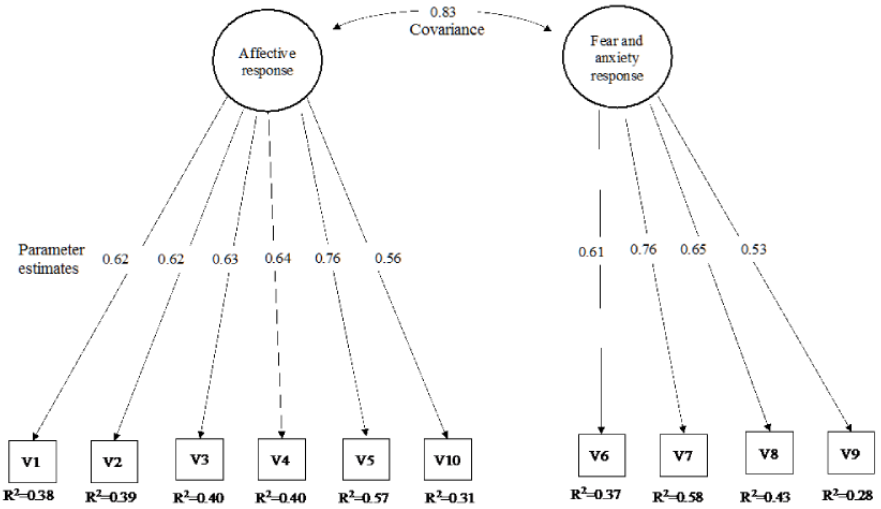
Construct Validity

The CFA indicated an acceptable fit to the data (see Table 1 and Figure 2). Two factors explaining 55% of the variance were confirmed by CFA. The first factor measured affective responses and the second factor measured fear and anxiety responses. Global scores ranged from 3 to 30 (see Table 2). They were higher among professionals working in health services that accumulated a higher number of deaths (mean 10.6, 95% CI 9.5-11.7 vs mean 8.2, 95% CI 6.5-9.9; *P*<.05). This tendency was also observed in the scores for the two-solution factors (affective responses: mean 6.3, 95% CI 5.6-7.9 vs mean 4.9, 95% CI 3.7-6.1; fear and anxiety responses: mean 4.3, 95% CI 2.8-4.8 vs mean 3.2, 95% CI 2.4-4.0). Lethality rate was positively related to the EASE scale scores (β=1.07, 95% CI 1.00-1.15; *P*<.05) (see Table 3).

Table 1. Confirmatory factor analysis to determine fit to the data.

Goodness-of-fit index	Value
Absolute adjustment rates	
Standardized root mean square residual	0.06
Jöreskog and Sörbom's goodness-of-fit index	0.92
Jöreskog and Sörbom's adjusted goodness-of-fit index	0.90
Relative adjustment rates	
Normed fit index	0.90
Comparative fit index	0.93

Figure 2. Path diagram of the confirmatory factor analysis. Standardized weights and measurement errors of each item of the Self-applied Acute Stress Scale (EASE). CFI: comparative fix index; GFI: goodness-of-fit index; RMSEA: root mean square error of approximation; V1: I can't help but think of recent critical situations. I can't get out of work; V2: I have completely lost the taste for things that gave me peace of mind; V3: I keep my distance, I resent dealing with people, and I'm irascible even at home; V4: I feel that I am neglecting many people who need my help; V5: I have difficulty thinking and making decisions, I have many doubts, and I have entered a kind of emotional blockage; V6: I feel intense physiological reactions (shock, sweating, dizziness, shortness of breath, insomnia, etc) related to the current crisis; V7: I feel on permanent alert. I believe that my reactions now put other patients, my colleagues, or myself at risk; V8: Worrying about not getting sick causes me a strain that is hard to bear; V9: I'm afraid I'm going to infect my family; V10: I have difficulty empathizing with patients' suffering or connecting with their situation (emotional distancing and emotional anesthesia).



χ^2_{34}=9.04, *P*<.001; RMSEA=0.085; CFI=0.92; GFI=0.93

Table 2. Construct validity of the acute stress scale used among professionals caring for patients with COVID-19.

Scale item	Scale score (N=228), mean (SD), 95% CI[a]	Loading
Factor 1: affective responses		
I keep my distance, I resent dealing with people, and I am irascible even at home.	1.1 (0.9), 1.0-1.2	0.83
I have completely lost the taste for things that used to bring me peace of mind or well-being.	1.1 (0.9), 1.0-1.2	0.72
I feel that I am neglecting many people who need my help.	0.9 (0.9), 0.8-1.0	0.71
I cannot help but think of recent critical situations. I can't get out of work.	1.3 (0.1), 1.2-1.4	0.60
I have difficulty thinking and making decisions, I have many doubts, and I have entered a kind of emotional blockage.	0.9 (0.9), 0.8-1.0	0.58
All factor 1 items	6.0 (3.9), 5.5-6.5	N/A[b]
Factor 2: fear and anxiety responses		
I have difficulty empathizing with patients' suffering or connecting with their situation (emotional distancing and emotional anesthesia).	0.7 (0.9), 0.6-0.8	0.48
Worrying about not getting sick causes me a strain that is hard to bear.	0.8 (0.9), 0.7-0.9	0.86
I'm afraid I'm going to infect my family.	1.3 (1.0), 1.2-1.4	0.70
I feel on permanent alert. I believe that my reactions now put other patients, my colleagues, or myself at risk.	0.9 (0.9), 0.8-1.0	0.66
I feel intense physiological reactions (shock, sweating, dizziness, shortness of breath, insomnia, etc) related to the current crisis.	1.0 (0.9), 0.9-1.1	0.55
All factor 2 items	4.0 (2.8), 3.6-4.4	N/A
Total (score ranges from 3 to 30)	10.0 (6.1), 9.2-10.8	N/A

[a]Individual scores range from 0 to 3.

[b]N/A: not applicable; this value was calculated for individual items only.

Table 3. Acute stress scale used among professionals caring for patients with COVID-19 over two periods and two geographical areas.

Scale item	Scale score (N=228), mean (SD), 95% CI[a]			
	First period: 800-900 deaths per day[b]	Second period: <500 deaths per day[b]	Geographical area: more impact[c]	Geographical area: less impact[d]
I can't help but think of recent critical situations. I can't get out of work.	1.5 (0.5), 1.2-1.8	0.8 (1.0), 0.4-1.2	1.4 (1.0), 1.2-1.6	1.2 (0.9), 0.9-1.5
I have completely lost the taste for things that gave me peace of mind.	1.5 (0.5), 1.2-1.8	1.2 (1.0), 0.8-1.6	*1.2 (0.9), 1.1-1.3*	*0.8 (0.8), 0.5-1.1*
I keep my distance, I resent dealing with people, and I'm irascible even at home.	*1.6 (0.9), 1.0-2.2*	*0.8 (0.8), 0.4-1.2*	1.2 (1.0), 1.0-1.4	0.9 (0.8), 0.6-1.2
I feel that I am neglecting many people who need my help.	1.3 (0.9), 0.7-1.9	0.8 (1.0), 0.4-1.2	0.9 (0.9), 0.8-1.0	0.8 (0.9), 0.5-1.1
I have difficulty thinking and making decisions, I have many doubts, and I have entered a kind of emotional blockage.	*1.5 (0.9), 0.9-2.1*	*0.6 (0.8), 0.2-1.0*	1.0 (1.0), 0.8-1.2	0.7 (0.9), 0.4-1.0
I feel intense physiological reactions (shock, sweating, dizziness, shortness of breath, insomnia, etc) related to the current crisis.	1.5 (0.8), 0.9-2.1	1.1 (0.9), 0.7-1.5	1.1 (1.0), 0.9-1.3	0.8 (0.9), 0.5-1.1
I feel on permanent alert. I believe that my reactions now put other patients, my colleagues, or myself at risk.	1.6 (1.1), 0.8-2.4	0.8 (0.8), 0.4-1.2	0.9 (1.0), 0.7-1.1	0.6 (0.7), 0.4-0.8
Worrying about not getting sick causes me a strain that is hard to bear.	1.1 (1.0), 0.4-1.8	0.7 (0.9), 0.3-1.1	0.9 (0.9), 0.8-1.0	0.6 (0.8), 0.3-0.9
I'm afraid I'm going to infect my family.	1.3 (0.7), 0.8-1.8	0.9 (1.0), 0.5-1.3	1.4 (1.0), 1.2-1.6	1.3 (1.0), 1.0-1.6
I have difficulty empathizing with patients' suffering or connecting with their situation (emotional distancing and emotional anesthesia).	1.4 (1.1), 0.6-2.2	0.6 (1.1), 0.1-1.1	0.7 (1.9), 0.6-0.8	0.6 (0.9), 0.3-0.9
Total (scores range from 3 to 30)	*14.3 (5.5), 10.5-18.1*	*8.3 (6.6), 5.3-11.3*	*10.6 (6.4), 9.5-11.7*	*8.2 (5.0), 6.5-9.9*

[a]Individual scores range from 0 to 3 and italicized values represent statistical significance: $P<.05$.

[b]The first period was from March 25 to April 1, 2020; the second period was from April 1 to 19, 2020.

[c]This geographical area was most affected by COVID-19, with a higher mortality per 1000 inhabitants, and included Madrid, Catalonia, Basque Country, Aragon, Castile and Leon, and Valencia.

[d]This geographical area was least affected by COVID-19, with a lower mortality per 1000 inhabitants, and included Asturias, Canary Islands, La Rioja, Murcia, and Navarre.

Criterion Validity

The EASE scale score was higher in the first period (March 25 to April 1, 2020) compared to the second period (April 1 to 19, 2020), in which the number of deaths per day decreased by half (mean 14.3, SD 5.5, 95% CI 10.5-18.1 vs mean 8.3, SD 6.6, 95% CI 5.3-11.3; $P<.05$) (see Table 3).

Interpretability

The following categories based on score ranges were established: 0-9 points, good emotional adjustment; 10-14 points, emotional distress; 15-24 points, medium-high emotional overload; and ≥25 points, extreme acute stress. Most of the respondents were in the first range (115/228, 50.4%) and 28.9% (66/228) were in the second. Only 2.6% (6/228) were in the fourth bracket.

Translation-Back-Translation

The Portuguese and English versions of the scale are shown in Multimedia Appendix 1.

Availability of Data and Materials

Data and materials are available upon reasonable request.

Discussion

Principal Findings

The instrument designed and validated in this study has adequate metric properties and seems useful for professionals to become aware of their levels of stress while caring for COVID-19 patients during the pandemic. The average level of self-reported stress was between 9 and 11 points up to 30. This level of stress was measured during a period of unprecedented pressure while caring for COVID-19 patients. During this period, there were relevant organizational changes, uncertainty about the evolution of patients, a lack of personal protective equipment, and an increase in the number of professionals infected. It is expected that by decreasing the pressure caused by these situations, the emotional response of these health care professionals would increase as they become fully aware of their experience [3,16]. This scale has been used to identify and prevent such progression of stressful experiences among health professionals, the second victims of SARS-CoV-2 during the pandemic. The EASE scale has been linked to intervention measures to support the staff who are in direct contact with COVID-19 patients, using digital materials designed to improve psychological

well-being [17]. This approach has been developed in several countries in addition to Spain [18,19].

This is particularly relevant because some current data in Spain using the EASE scale are suggesting that during the third wave, the level of distress reported by health care professionals is increasing up to 3 times, with 8.8% of respondents being classified into the fourth category of scores, probably as a result of accumulated fatigue and a feeling of starting over [5]. Therefore, caring for those who care has become a priority for public health strategies [20,21].

Stress, hypervigilance, fatigue, difficulty sleeping, inability to relax, and fear were common symptoms among health care professionals at the start of the pandemic [5]. The content validity analysis conducted in this study identified behaviors and responses related to these symptoms; the EASE scale items seek to assess the scale's effect on those who provide care to COVID-19 patients. However, although mental symptoms have been identified, physical symptoms must also be considered when interventions to support them are being designed [20].

The EASE scale has sufficient sensitivity; the staff in contact with COVID-19 patients had higher levels of acute stress than other staff, as suggested in other studies [22]. When the EASE scale has been used in other studies, it was revealed that the levels of acute stress were higher in the absence of personal protective equipment, when the care pressure was greater, among professionals in critical or emergency units, and, as in this case, in those territories with a greater number of cases [23,24]. Since the current situation (ie, the third wave) seems different from the previous one, the scale's utility in the next 3 to 6 months, after the worst of the current health crisis, should be checked.

In previous pandemics and in cases of natural disasters (earthquakes, tsunamis, etc), terrorist attacks with numerous victims, air or train crashes, and war conflicts, professionals have also experienced distress with consequences that have lasted a while [25]. COVID-19 is not the first pandemic in recent times. SARS-CoV in 2003 in China and Canada, MERS-CoV in 2012 in Saudi Arabia, and the Ebola outbreak in 2014 in several African countries, which reached Spain and other Organisation for Economic Co-operation and Development countries, impacted the well-being of health care professionals [26,27]. However, the magnitude of the health crisis caused by the COVID-19 pandemic has been more global and temporarily more widespread than in recent previous cases. Some signs of acute stress measured by the EASE scale are similar to all of these situations. However, some items are specific to an outbreak where the biological risk for professionals is present. In this sense, the EASE scale could be used in future outbreaks to check the level of acute stress of the health care workforce and to give them proportional support according to their needs in these conditions.

The rapid spread of SARS-CoV-2, coupled with the breakdown of the supply chain, lack of equipment, and uncertainty about how to deal with the treatment of COVID-19 patients, prompted the need for rapid responses to the pressing problems of the time. The EASE scale emerged at a time when the response capacity of professionals was under threat, as they were physically and mentally overwhelmed and many colleagues were becoming infected; in Spain, the number of professionals who were COVID-19 patients in the first wave were among the highest in the world.

The EASE scale includes four score brackets; each of them is linked to specific recommendations to address the psychological burden due to caring for COVID-19 patients [17]. Having a specific instrument such as the EASE scale to monitor acute stress during a pandemic can be beneficial for two reasons. First, it helps professionals become aware of their situation and could contribute to initiating self-help behaviors early on. Second, health care organizations have a "thermometer" with which to advise health care providers of when they should take a mandatory rest in order to benefit patients in their care, their colleagues, and themselves, before succumbing to the overload they are enduring.

Now, in this crisis, these same professionals should continue to care for patients with COVID-19, patients with sequelae from COVID-19, and all patients whose care processes have been interrupted during the acute phase of the pandemic. In this scenario, if rapid action is not taken, professionals' capacity and, therefore, the quality and safety of patient care may be compromised.

This scenario may be a valuable opportunity to consolidate integrated care; now is the time to consider appropriate strategies to introduce measures that will increase health care professionals' well-being at work and strengthen clinical leadership. In addition, it is time to commit to a model such as the Quadruple Aim of health care, which considers patient outcomes to be dependent on how caregivers are supported. The Quadruple Aim recognizes this focus within the context of the broader transformation required in our health care system toward high-value care. While the first three aims provide a rationale for a health system [28], the fourth aim becomes a foundational element for the other goals to be realized [29]. The key is the fourth aim: creating the conditions for the health care workforce to find joy and meaning in their work and improve the experience of providing care [30]. For this reason, it seems that "caring for the caregiver" [31] is necessary in the transformation of the health system, and having instruments to be able to monitor the effects of measures implemented can be very useful; at this point in time, this aim becomes "caring for the caregiver in times of pandemic" as a way to achieve optimal care for patients.

A limitation of this study is that it does not discriminate between professional categories, nor does it consider critical services separately during this crisis, such as critical care and resuscitation, internal medicine, pneumology, and infectious diseases. As well, these data are limited to Spain. Although a cultural and linguistic equivalence analysis of the scale was made, a measurement of the invariance among languages could be conducted in future studies. In the absence of a gold standard to assess criterion validity, in this study the ability of the scale to correctly classify respondents' answers was considered.

In the forthcoming months, we can expect professionals to be affected during the outbreak as a consequence of stress overload [5], having seen their professional codes violated [3] due to

insufficient resources to care for COVID-19 patients, contradictory instructions, or the interruption in the continuity of care of non-COVID-19 patients. Also, we can expect an increase of affective and anxiety reactions and symptoms, including, in some cases, posttraumatic stress among the professionals who saw their health and that of their loved ones threatened [32,33]. Continuing to monitor acute stress levels may be advisable in order to check on the effective recovery of health care professionals, which implies that the health of the population will continue to be adequately cared for.

Most probably, the initial impact due to the conditions in which the treatment of patients was carried out (ie, lack of equipment, lack of guidelines, overloading , and fear of contagion) is not related to the current experience of the health care providers. The recovery of professionals' health, teamwork, and workers' morale in health care organizations, particularly where the deceased include professionals from the team itself, will probably require specific and wide-ranging actions. Actions that promote a positive dynamic within the health care teams are more likely to be successful [5], for example, by reflecting on how they have acted, what has worked, and what could have been done differently [34-36].

Conclusions

The EASE scale has been shown to exhibit adequate metric properties, such that it may be considered a reliable and valid scale. Its usefulness is two-fold: firstly, to help professionals become aware of their emotional overload and that it can be supported and, secondly, to measure the effect of this overload to avoid the progression toward more severe psychopathological conditions.

Acknowledgments

We would like to thank all the health professionals who, faced with an unprecedented health care challenge, have taken a step forward to face this health care crisis. During this study, JMD, IC, and JJM benefited from a grant from the Regional Ministry of Education, Research, Culture and Sports of the Valencian Region for Excellence Research Groups (reference No. PROMETEO/2017/173).

Authors' Contributions

JJM, IC, MG, VPJ, AC, and PPP designed the study. VPJ, OMG, MJBD, BMG, and CAR reviewed the literature and all authors described the experiences of frontline professionals. All authors contributed to redefining and improving the scale and collected responses. CF and MAV designed the platforms to collect responses. JMD and AM participated in the analysis of data. JJM, IC, MG, and JMD prepared a first draft of this manuscript. The Core Group was composed of Adriana Lopez-Pineda, Julian Vitaller Burillo, Juan Francisco Herrera Cuenca, Maria Luisa Torrijano Casalengua, Antonio Guilabert Gimenez, Carmen Muñoz Ruiperez, Isabel María Galán Meléndez, Carolina Varela Rodríguez, Auxi Javaloyes Sanchís, Inmaculada Palazón, Daniela Campos de Andrade Lourenção, and Jorge de Vicente Guijarro. All authors contributed to and approved the final version of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The Self-applied Acute Stress Scale (EASE).
[DOCX File , 30 KB - formative_v5i3e27107_app1.docx]

References

1. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020 Feb 15;395(10223):497-506 [FREE Full text] [doi: 10.1016/S0140-6736(20)30183-5] [Medline: 31986264]
2. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, China Novel Coronavirus Investigating and Research Team. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 2020 Feb 20;382(8):727-733 [FREE Full text] [doi: 10.1056/NEJMoa2001017] [Medline: 31978945]
3. Greenberg N, Docherty M, Gnanapragasam S, Wessely S. Managing mental health challenges faced by healthcare workers during COVID-19 pandemic. BMJ 2020 Mar 26;368:m1211. [doi: 10.1136/bmj.m1211] [Medline: 32217624]
4. Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, et al. Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. JAMA Netw Open 2020 Mar 02;3(3):e203976 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.3976] [Medline: 32202646]
5. Vanhaecht K, Seys D, Bruyneel L, Cox B, Kaesemans G, Cloet M, et al. COVID-19 is having a destructive impact on health-care workers' mental well-being. Int J Qual Health Care 2021 Feb 20;33(1):mzaa15 [FREE Full text] [doi: 10.1093/intqhc/mzaa158] [Medline: 33270881]

6. Kisely S, Warren N, McMahon L, Dalais C, Henry I, Siskind D. Occurrence, prevention, and management of the psychological effects of emerging virus outbreaks on healthcare workers: Rapid review and meta-analysis. BMJ 2020 May 05;369:m1642 [FREE Full text] [doi: 10.1136/bmj.m1642] [Medline: 32371466]
7. Coronavirus Disease 2019 (COVID-19) in the EU/EEA and the UK– ninth update. Stockholm, Sweden: European Centre for Disease Prevention and Control; 2020 Apr 23. URL: https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-rapid-risk-assessment-coronavirus-disease-2019-ninth-update-23-april-2020.pdf [accessed 2020-04-29]
8. Liu C, Yang Y, Zhang X, Xu X, Dou Q, Zhang W, et al. The prevalence and influencing factors in anxiety in medical workers fighting COVID-19 in China: A cross-sectional survey. Epidemiol Infect 2020 May 20;148:e98 [FREE Full text] [doi: 10.1017/S0950268820001107] [Medline: 32430088]
9. Cai H, Tu B, Ma J, Chen L, Fu L, Jiang Y, et al. Psychological impact and coping strategies of frontline medical staff in Hunan between January and March 2020 during the outbreak of coronavirus disease 2019 (COVID-19) in Hubei, China. Med Sci Monit 2020 Apr 15;26:e924171 [FREE Full text] [doi: 10.12659/MSM.924171] [Medline: 32291383]
10. Ibáñez-Vizoso JE, Alberdi-Páramo Í, Díaz-Marsá M. International mental health perspectives on the novel coronavirus SARS-CoV-2 pandemic. Rev Psiquiatr Salud Ment 2020;13(2):111-113 [FREE Full text] [doi: 10.1016/j.rpsm.2020.04.002] [Medline: 32389514]
11. Galbraith N, Boyda D, McFeeters D, Hassan T. The mental health of doctors during the COVID-19 pandemic. BJPsych Bull 2020 Apr 28:1-4 [FREE Full text] [doi: 10.1192/bjb.2020.44] [Medline: 32340645]
12. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: A primer. Front Public Health 2018;6:149 [FREE Full text] [doi: 10.3389/fpubh.2018.00149] [Medline: 29942800]
13. Chan EKH. Standards and guidelines for validation practices: Development and evaluation of measurement instruments. In: Zumbo BD, Chan EKH, editors. Validity and Validation in Social, Behavioral, and Health Sciences. Social Indicators Research Series, vol 54. Cham, Switzerland: Springer International Publishing; 2014:9-24.
14. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. Qual Life Res 2010 May;19(4):539-549 [FREE Full text] [doi: 10.1007/s11136-010-9606-8] [Medline: 20169472]
15. Garfin DR, Thompson RR, Holman EA. Acute stress and subsequent health outcomes: A systematic review. J Psychosom Res 2018 Sep;112:107-113. [doi: 10.1016/j.jpsychores.2018.05.017] [Medline: 30097129]
16. Borges LM, Bahraini NH, Holliman BD, Gissen MR, Lawson WC, Barnes SM. Veterans' perspectives on discussing moral injury in the context of evidence-based psychotherapies for PTSD and other VA treatment. J Clin Psychol 2020 Mar;76(3):377-391. [doi: 10.1002/jclp.22887] [Medline: 31714610]
17. Mira JJ, Vicente MA, Lopez-Pineda A, Carrillo I, Guilabert M, Fernández C, et al. Preventing and addressing the stress reactions of health care workers caring for patients with COVID-19: Development of a digital platform (Be + Against COVID). JMIR Mhealth Uhealth 2020 Oct 05;8(10):e21692 [FREE Full text] [doi: 10.2196/21692] [Medline: 32936769]
18. Strametz R, Raspe M, Ettl B, Huf W, Pitz A. Recommended actions: Reinforcing clinicians' resilience and supporting second victims during the COVID-19 pandemic to maintain capacity in the healthcare system [Article in German]. Zentralbl Arbeitsmed Arbeitsschutz Ergon 2020 Sep 02:1-5 [FREE Full text] [doi: 10.1007/s40664-020-00405-7] [Medline: 32905028]
19. Blake H, Bermingham F, Johnson G, Tabner A. Mitigating the psychological impact of COVID-19 on healthcare workers: A digital learning package. Int J Environ Res Public Health 2020 Apr 26;17(9):2997 [FREE Full text] [doi: 10.3390/ijerph17092997] [Medline: 32357424]
20. Salazar de Pablo G, Vaquerizo-Serrano J, Catalan A, Arango C, Moreno C, Ferre F, et al. Impact of coronavirus syndromes on physical and mental health of health care workers: Systematic review and meta-analysis. J Affect Disord 2020 Oct 01;275:48-57 [FREE Full text] [doi: 10.1016/j.jad.2020.06.022] [Medline: 32658823]
21. Morgantini L, Naha U, Wang H, Francavilla S, Acar Ö, Flores J, et al. Factors contributing to healthcare professional burnout during the COVID-19 pandemic: A rapid turnaround global survey. PLoS One 2020;15(9):e0238217 [FREE Full text] [doi: 10.1371/journal.pone.0238217] [Medline: 32881887]
22. Kisely S, Warren N, McMahon L, Dalais C, Henry I, Siskind D. Occurrence, prevention, and management of the psychological effects of emerging virus outbreaks on healthcare workers: Rapid review and meta-analysis. BMJ 2020 May 05;369:m1642 [FREE Full text] [doi: 10.1136/bmj.m1642] [Medline: 32371466]
23. Mira JJ, Carrillo I, Guilabert M, Mula A, Martin-Delgado J, Pérez-Jover MV, SARS-CoV-2 Second Victim Study Group. Acute stress of the healthcare workforce during the COVID-19 pandemic evolution: A cross-sectional study in Spain. BMJ Open 2020 Nov 06;10(11):e042555 [FREE Full text] [doi: 10.1136/bmjopen-2020-042555] [Medline: 33158839]
24. Martin-Delgado J, Viteri E, Mula A, Serpa P, Pacheco G, Prada D, et al. Availability of personal protective equipment and diagnostic and treatment facilities for healthcare workers involved in COVID-19 care: A cross-sectional study in Brazil, Colombia, and Ecuador. PLoS One 2020;15(11):e0242185 [FREE Full text] [doi: 10.1371/journal.pone.0242185] [Medline: 33175877]

25. Mattei A, Fiasca F, Mazzei M, Necozione S, Bianchini V. Stress and burnout in health-care workers after the 2009 L'Aquila earthquake: A cross-sectional observational study. Front Psychiatry 2017;8:98 [FREE Full text] [doi: 10.3389/fpsyt.2017.00098] [Medline: 28659831]
26. Khalid I, Khalid TJ, Qabajah MR, Barnard AG, Qushmaq IA. Healthcare workers emotions, perceived stressors and coping strategies during a MERS-CoV outbreak. Clin Med Res 2016 Mar;14(1):7-14 [FREE Full text] [doi: 10.3121/cmr.2016.1303] [Medline: 26847480]
27. Lehmann M, Bruenahl CA, Löwe B, Addo MM, Schmiedel S, Lohse AW, et al. Ebola and psychological stress of health care professionals. Emerg Infect Dis 2015 May;21(5):913-914 [FREE Full text] [doi: 10.3201/eid2105.141988] [Medline: 25897490]
28. Berwick DM, Nolan TW, Whittington J. The Triple Aim: Care, health, and cost. Health Aff (Millwood) 2008;27(3):759-769. [doi: 10.1377/hlthaff.27.3.759] [Medline: 18474969]
29. Sikka R, Morath JM, Leape L. The Quadruple Aim: Care, health, cost and meaning in work. BMJ Qual Saf 2015 Oct;24(10):608-610. [doi: 10.1136/bmjqs-2015-004160] [Medline: 26038586]
30. Bodenheimer T, Sinsky C. From Triple to Quadruple Aim: Care of the patient requires care of the provider. Ann Fam Med 2014;12(6):573-576 [FREE Full text] [doi: 10.1370/afm.1713] [Medline: 25384822]
31. Karb R, Natsui S. Reflections: Addressing the Quadruple Aim in social emergency medicine. Ann Emerg Med 2019 Nov;74(5S):S31-S32. [doi: 10.1016/j.annemergmed.2019.08.439] [Medline: 31655671]
32. Chen Q, Liang M, Li Y, Guo J, Fei D, Wang L, et al. Mental health care for medical staff in China during the COVID-19 outbreak. Lancet Psychiatry 2020 Apr;7(4):e15-e16 [FREE Full text] [doi: 10.1016/S2215-0366(20)30078-X] [Medline: 32085839]
33. Kang L, Ma S, Chen M, Yang J, Wang Y, Li R, et al. Impact on mental health and perceptions of psychological care among medical and nursing staff in Wuhan during the 2019 novel coronavirus disease outbreak: A cross-sectional study. Brain Behav Immun 2020 Jul;87:11-17 [FREE Full text] [doi: 10.1016/j.bbi.2020.03.028] [Medline: 32240764]
34. Traylor AM, Tannenbaum SI, Thomas EJ, Salas E. Helping healthcare teams save lives during COVID-19: Insights and countermeasures from team science. Am Psychol 2021 Jan;76(1):1-13. [doi: 10.1037/amp0000750] [Medline: 33119329]
35. Janssen M, van der Voort H. Agile and adaptive governance in crisis response: Lessons from the COVID-19 pandemic. Int J Inf Manage 2020 Dec;55:102180 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102180] [Medline: 32836637]
36. Schippers M. Optimizing decision-making processes in times of COVID-19: Using reflexivity to counteract information processing failures. ERIM Report Series reference forthcoming. SSRN 2020 May 13:1-17 [FREE Full text] [doi: 10.2139/ssrn.3599939]

Abbreviations

CFA: confirmatory factor analysis
COSMIN: COnsensus-based Standards for the selection of health status Measurement INstruments
EASE: Self-applied Acute Stress Scale
MERS: Middle East respiratory syndrome

Edited by J Torous; submitted 12.01.21; peer-reviewed by S Rostam Niakan Kalhori , L Guan; comments to author 06.02.21; revised version received 08.02.21; accepted 12.02.21; published 09.03.21.

Please cite as:
Mira JJ, Cobos A, Martínez García O, Bueno Domínguez MJ, Astier-Peña MP, Pérez Pérez P, Carrillo I, Guilabert M, Perez-Jover V, Fernandez C, Vicente MA, Lahera-Martin M, Silvestre Busto C, Lorenzo Martínez S, Sanchez Martinez A, Martin-Delgado J, Mula A, Marco-Gomez B, Abad Bouzan C, Aibar-Remon C, Aranaz-Andres J, SARS-CoV-2 Second Victims Working Group
An Acute Stress Scale for Health Care Professionals Caring for Patients With COVID-19: Validation Study
JMIR Form Res 2021;5(3):e27107
URL: https://formative.jmir.org/2021/3/e27107
doi:10.2196/27107
PMID:33687343

Viewpoint

Recruitment and Retention Strategies for Community-Based Longitudinal Studies in Diverse Urban Neighborhoods

Emily B Ferris[1], MPH; Katarzyna Wyka[1], MA, PhD; Kelly R Evenson[2], MS, PhD; Joan M Dorn[3], PhD; Lorna Thorpe[4], PhD; Diane Catellier[5], DrPH, MS; Terry T-K Huang[1], MBA, PhD, MPH

[1]Center for Systems and Community Design, Graduate School of Public Health and Health Policy, City University of New York, New York, NY, United States

[2]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, United States

[3]Sophie Davis Biomedical Education Program, City University of New York School of Medicine, New York, NY, United States

[4]Department of Population Health, School of Medicine, New York University, New York, NY, United States

[5]Research Triangle Institute, Research Triangle Park, Durham, NC, United States

Corresponding Author:
Terry T-K Huang, MBA, PhD, MPH
Center for Systems and Community Design
Graduate School of Public Health and Health Policy
City University of New York
55 W 125th St Room 80
New York, NY, 10027
United States
Phone: 1 646 364 0247
Email: terry.huang@sph.cuny.edu

Abstract

Longitudinal, natural experiments provide an ideal evaluation approach to better understand the impact of built environment interventions on community health outcomes, particularly health disparities. As there are many participant engagement challenges inherent in the design of large-scale community-based studies, adaptive and iterative participant engagement strategies are critical. This paper shares practical lessons learned from the Physical Activity and Redesigned Community Spaces (PARCS) study, which is an evaluation of the impact of a citywide park renovation initiative on physical activity, psychosocial health, and community well-being. The PARCS study, although ongoing, has developed several approaches to improve participant engagement: building trust with communities, adapting the study protocol to meet participants' needs and to reflect their capacity for participation, operational flexibility, and developing tracking systems. These strategies may help researchers anticipate and respond to participant engagement challenges in community-based studies, particularly in low-income communities of color.

(JMIR Form Res 2021;5(3):e18591) doi:10.2196/18591

KEYWORDS

community-based; participant engagement; natural experiment; built environment intervention; health disparities; study adaptations

Introduction

Given the projection that obesity prevalence among US adults will rise to 49% by 2030 [1] and the many health problems associated with obesity [2], obesity and physical inactivity continue to be major public health issues [3-5]. Black and Hispanic communities in the United States have higher rates of obesity and physical inactivity and are disproportionately at risk of associated, myriad health issues [6-9]. Compared with non-Hispanic White adults (37.9%), both Hispanic adults of all races (47.0%) and non-Hispanic Black adults (46.8%) have a higher prevalence of obesity [10]. Among women, increased income and educational attainment are associated with decreased obesity prevalence [11]. Fewer Hispanic adults of all races (21.3%) and non-Hispanic Black adults (20.1%) met the 2008 Federal Physical Activity Guidelines compared with non-Hispanic White adults (25.6%) [12]. Given the complex drivers fueling the obesity epidemic and the entrenched social and environmental causes of health disparities, a strategic range of interventions tailored to diverse communities is critical to effectively address obesity, physical inactivity, and the associated health disparities in the United States [13].

Existing research has found associations between obesity prevalence and physical activity behaviors and many aspects

of the built environment, including land use mix, connectivity [14], access to parks, and food retail options [15]. Inequities in the built environment may contribute to socioeconomic and racial disparities in obesity and physical activity. A nationally representative cohort study found that census blocks of lower socioeconomic status and a large proportion of minoritized residents had less access to physical activity facilities, which in turn was associated with increased overweight and decreased levels of physical activity [16]. As a result of institutional racism, longstanding racial residential segregation is a driving force of this type of inequitable distribution of and access to resources, with negative impacts on socioeconomic status (SES) and health outcomes [17].

With the potential to reach a large number of people and promote sustainable behavior change, built environment interventions offer a promising approach to prevent and reduce obesity [15,18]. For example, some park renovation interventions have positively impacted residents' physical activity behaviors [19,20]. As obesogenic built environment features are more prevalent in lower-income neighborhoods and neighborhoods with large communities of color, built environment interventions can contribute to addressing racial, ethnic, and socioeconomic disparities in prevalence of obesity and levels of physical activity [16,21-23]. However, much of the limited, existing research on built environment interventions has used cross-sectional study designs and lacks a focus on communities of color or lower SES communities, suggesting a need for more rigorous study designs to provide higher quality evidence [24] and to better understand the potential impact of built environment interventions over time, especially in lower-income neighborhoods or communities of color [15,25,26].

The limited evidence base is, in part, due to challenges in evaluating built environment interventions. The potential of built environment interventions lies in their ability to shift daily habits and behaviors [27]; however, it can take time to measure and detect the impact of these everyday behavior changes on the health of the residents. Although longitudinal studies, especially randomized controlled trials, are ideal to examine the impact of built environment changes over time, this is not always possible in the real-world context [15]. Natural experiments are an appropriate method for evaluating policy or large-scale changes such as built environment interventions [24,28] but they are often difficult to implement rigorously. Systematic reviews of naturally occurring experiments evaluating the impact of policy and built environment changes on obesity highlighted the need to provide sufficient follow-up time and collect data at multiple time points to provide a more valid measurement of potential changes [15,24].

As two of the major challenges to rigorous, longitudinal natural experiments are participant recruitment and retention [29], effective participant engagement strategies and well-run study operations are key elements to ensure the study's success. There are many potential impediments to enrolling participants in a large-scale longitudinal study and successfully following up with them. Potential participants must be willing to commit to participating over a long period and anticipate continuing to meet the study eligibility criteria. Throughout the study, some participants may lose interest or experience study burn out, whereas others may move or change their contact information without informing the research team [30]. Additional barriers emerge in communities of color and low-income communities. Mistrust of medical professionals and fear of exploitation have been identified as challenges in recruiting participants of color in the United States [31,32]. For low-income individuals of diverse racial and ethnic identities, competing demands are a major barrier and, in general, the burden of participation is much higher for low-income individuals than for higher-income study participants [31,32]. Owing to these challenges, iterative participant engagement strategies and operational flexibility to keep participants involved become critical to the success of the study, as high attrition can increase the risk of bias and impact study validity [30,33]. The growing interest and increasing body of literature exploring effective participant engagement and retention strategies reflect this importance [31,34,35].

Given the dynamic nature of longitudinal studies, research teams' participant engagement strategies may change frequently to adapt to participant and study needs. Research teams' documented engagement protocols, which are created either for internal use or for institutional board review, do not always reflect these adaptations and the full range of strategies actually used by a research team [33]. A review of studies with high participant retention found that research teams iteratively adapted and tailored engagement strategies throughout the life of the study based on the specific needs of participants and that this process was rarely documented [33]. Successful participant engagement strategies are often labor-intensive, including substantial *in-the-field tracking*, and iteratively translate insights from researchers and participants into protocol adaptations [31,36]. This paper, presented as a practical viewpoint from the field, seeks to document insights learned from the Physical Activity in Redesigned Community Spaces (PARCS) study regarding adaptive participant engagement strategies in a natural experiment evaluating changes to the built environment. These insights demonstrate the diverse range of strategies and the related operational flexibility that research teams can deploy to effectively engage participants, particularly from low-income communities and/or communities of color, in longitudinal studies.

PARCS Study Background

The PARCS study is a longitudinal, natural experiment evaluating the impact of the Community Parks Initiative, an equity-based renovation initiative by the New York City Department of Parks and Recreation (NYC Parks), on residents' physical activity, mental health, and community well-being. The Community Parks Initiative renovations focus on neighborhood parks that had not received significant capital investment in the past two decades and met two of the following three criteria: above average population density, above average percentage of residents living below the federal poverty line, and recent population growth. Control parks were matched based on neighborhood demographic characteristics and met the Community Parks Initiative inclusion criteria but were not slated for renovation during the study timeline. The full research protocol has been published elsewhere [37]. This study was

approved by the City University of New York Institutional Review Board. All participants provided informed consent before their inclusion in the study.

Baseline data were collected from June 2016 to August 2018, with the initial goal of recruiting approximately 1700 participants. For most study sites, the second wave of data collection occurred 2 years after baseline data collection and wave 3 occurred 3 years after baseline. However, due to changes in the park renovation timelines, the second wave of data collection took place 3-4 years after baseline in the selected sites. Although many studies experience more difficulty recruiting control participants compared with intervention participants, this has not been the case in the PARCS study. Although information about parks included in the Community Parks Initiative is publicly available, PARCS study participants were not explicitly told the study arm of their neighborhood site to minimize potentially influencing participants' park use behaviors, and the preliminary average number of participants per study site at baseline was similar in the intervention and control sites (approximately 30 participants per site). In a preliminary baseline sample, 54.2% of participants reported an annual income of ≤US $20,000, and the majority were identified as Hispanic of any race (43.1%) or non-Hispanic Black (49.5%) [38]. As follow-up data collection is ongoing, retention rates are not yet known.

All participants lived within 0.3 miles from one of the 54 study parks (33 intervention parks and 21 control parks). The study included two cohorts: an adult-only cohort and a parent-child dyad cohort in which a primary caregiver and a child aged between 3 years and 8 years were enrolled together. Participants agreed to wear an accelerometer for at least 10 hours per day over a 7-day period to measure both physical activity and sedentary behavior. Over the same 7-day period, adult participants used the PARCS study app to respond to a survey with questions on a range of psychosocial and community well-being measures. Through the study app, participants also responded to brief real-time ecological momentary assessment surveys regarding park use. Using mobile geographic information system-enabled technology, the PARCS study app geofenced each study park and recorded participants' usage of study parks. Field staff referred to as project coordinators were responsible for participant recruitment, implementing engagement strategies, distributing study materials, following up with participants, and providing a community-level interface for the study. Three project managers were responsible for overseeing field staff, coordinating site scheduling and accelerometer distribution, monitoring participant engagement strategies, and tracking adherence metrics to monitor progress toward study goals.

The PARCS study methodology is a response to the call for more rigorous study designs to generate stronger evidence regarding the relationship between built environment interventions and health behaviors and outcomes. As the included park renovations mainly occur in lower-income communities of color throughout New York City, the study further aims to generate evidence about the relationship between built environment interventions and health equity. Due to the long timeframe and a focus on low-income communities and neighborhoods where a majority of residents are people of color, the PARCS study team has encountered many of the previously reported challenges of successfully engaging participants [33]. Throughout the process, the research team attempted to respond by adapting diverse participant engagement and operational management strategies.

Study Adaptations in Participant Engagement and Operational Flexibility

It is critical to employ both external (ie, participant and community focused) and internal (ie, within the research team) strategies to optimize participant engagement. Looking outward, it is key to have iterative, responsive strategies that directly engage with participants and support participants' continued involvement in the project. It is equally important to have internal, project, and staff management strategies that are adaptable, clear, and synergistic. On the basis of assessing protocol changes to date and tracking the adaptations' efficacy when feasible, the PARCS study team's approach to participant engagement has centered on the following four dimensions: (1) building trust with communities, (2) adapting the protocol to meet participant capacity, (3) establishing operational flexibility, and (4) developing tracking systems. The PARCS study team identified these priority dimensions based on the key themes that emerged repeatedly throughout the study in discussions with the investigator team and study staff and that encapsulated the most pressing participant engagement and operations challenges. These priority dimensions also reflect previous research on the importance of developing trust with communities [31,32], having flexible operational structures to solicit feedback from staff and implement rapid protocol adjustments [31], and developing a consistent study identity [33,34]. Table 1 provides a summary of all strategies and the hypothesized impact.

Table 1. Summary of participant engagement strategies.

Methods and strategies	Hypothesized impact
Dimension 1: building trust with communities	
Creating added value	
• Regular social media content (eg, Instagram, Facebook) • Monthly touchpoints (birthday and holiday cards, pulse surveys, raffles, and photo contests)	• Participants feel part of a community with shared values and mutual interest in contributing to their neighborhoods • Participants have continued opportunities to engage and feel connected with the study community
Maintaining a professional and legitimate presence	
• Maximize project coordinator's ability to quickly signal their association with known organizations • Address verification	• Participants recognize affiliated organizations and are more likely to trust staff and believe in the project's legitimacy and mission • Further demonstrated validity of study and level of commitment necessary to participate
Branding	
• Consistent branding of all materials	• The PARCS[a] study becomes increasingly familiar and trustworthy within study neighborhoods
Dimension 2. Adapting the protocol to meet participant capacity	
Scheduled appointments versus rapid deployment	
• Participants were screened and enrolled at the same initial meeting	• Participants were more likely to successfully enroll
Additional sites	
• Four additional sites were added	• Helped address recruitment challenges related to variance in density and zoning among neighborhoods
Supplemental sample cohort	
• A supplemental cohort was recruited at each follow-up wave of data collection	• Helped address attrition
Dimension 3. Establishing operational flexibility	
Switching from teams to one operating unit	
• Operational structure shifted from three distinct units to one more centralized team	• Increased operational flexibility allowed for scheduling to be more efficiently managed by one-point person
Case management approach	
• Field staff were responsible for checking in with the participants they enrolled	• Developed rapport and a deeper connection between participants and the study
Reporting mechanisms	
• Establish communication channels so field staff can efficiently report issues from the field and managers can communicate protocol changes	• Research can adapt quickly and efficiently as issues come up in the field
Dimension 4. Developing tracking systems	
Tracking enrollment and retention	
• Weekly tracking of the number of participants enrolled per hour worked by each project coordinator and of contact attempts to connect with returning participants for follow-up waves of data collection	• Allowed staff to address any training or site-specific enrollment issues and provided field staff a sense of ownership and investment with the broader project goals and identify sites which needed additional support
Protocol adherence	

Methods and strategies	Hypothesized impact
• Weekly tracking of survey completion, accelerometer return, and accelerometer wear adherence rates	• Allowed staff to gauge level of participants' engagement with protocol, provide individualized feedback to participants, and quickly identify any protocol adherence issues

[a]PARCS: Physical Activity and Redesigned Community Spaces.

Dimension 1: Building Trust With Communities

Starting with the initial approach and throughout follow-up interactions, it is critical to build trust and a respectful rapport with participants. The PARCS study team has employed the following methods to build and maintain trust in communities: (1) creating added value for participants, (2) maintaining a professional and legitimate presence, and (3) building a consistent brand.

Creating Added Value

Given the long-term nature of longitudinal research and changes to the built environment, there may not be an immediate benefit to study participants. To offset this, research teams can create an additional *value add* for participants. In a qualitative substudy using key informant interviews with 20 PARCS participants to understand their motivations for participating, many participants identified the importance of helping their communities and contributing to the society [39]. On the basis of these findings, the PARCS team recognized that the most compelling *valueaddition* for many participants may be the opportunity to contribute to their neighborhoods and to be part of a community of people with similar values. In this vein, the study team attempted to foster a sense of community among participants, aligned with participants' values and motivations for participating.

The PARCS team primarily used social media strategies to develop this sense of community. The team posted content weekly on the PARCS study Instagram and Facebook accounts to build the study identity and develop an online community specifically for PARCS study participants. Through this content, we aimed to reflect the participants' interest in meaningful connections and in contributing to their communities and to further convey the message that they found a community of people with shared values within the PARCS study.

The research project provided additional value through monthly touchpoints with study participants, including quarterly newsletters, holiday and birthday cards, pulse surveys (eg, short text-based surveys based on timely topics), raffles, and photo contests. The monthly touchpoints were designed to help participants feel connected to the project and to the larger PARCS community. Just as importantly, they were designed to involve minimal effort from the participant.

Maintaining a Professional and Legitimate Presence

With 54 sites across all five New York City boroughs and limited recruitment time windows (2-4 weeks per site), field staff needed recruitment materials that were easy to transport on public transit and communication strategies, which were effective for diverse audiences. Field staff needed to quickly establish a reputable professional presence in the study neighborhood. Given that residents must live within a 0.3-mile radius of a park to be eligible, this created hyperlocalized recruitment zones with different physical and social features. In some study sites, field staff were able to collaborate with local organizations within a 0.3-mile radius and recruited from these established local community fixtures. However, most study sites did not have this type of existing infrastructure within the study zone. At most sites, field staff relied on street-intercept recruitment strategies where they approached residents in key locations to describe the study. Due to the informal nature of street-intercept approaches, it is difficult to establish a professional rapport and to communicate the project's legitimacy. Field staff reported receiving the following feedback from potential participants: wariness to share personal information with strangers, distrust that the study was affiliated with a credible institution, and skepticism regarding the accelerometers and fear of being *tracked* by the device.

To address these concerns, the study team developed strategies to maximize the field staff's ability to quickly signal their affiliations with a known organization to legitimize the project. The research team purchased folding tables, banners, postcards, branded pens, and wristbands. When appropriate, the field staff set up the table and decorated it with promotional materials. Being able to approach community residents from an established (if temporary) space helped community residents feel comfortable that the study was a legitimate undertaking. Field staff were supplied with branded PARCS study t-shirts, tote bags, and lanyards to demonstrate their affiliations with a well-known official organization and project. Where it was not appropriate to use a table, field staff approached potential participants with study flyers or postcards visible so that potential participants could immediately see supporting documentation.

Upon the recommendation of a staff member with ties to some of the study communities, an optional address verification question was added to the eligibility screener. Potential participants were asked to verify their addresses with an ID or mail. This was optional, and upon meeting the rest of the eligibility criteria, participants could enroll without verifying their address. The idea was to further demonstrate the validity of the project by confirming this information. Many services that people in New York are familiar with, such as Citi Bike rentals or library services, require proof of residency. After adding this question, 714 out of 1590 screened potential participants and 551 out of 996 enrolled participants verified their addresses. The enrollment rate increased from 48% to 54% after the inclusion of the address verification question. However, address verification did not appear to have increased survey completion rates, as survey completion was similar between participants who did and did not verify their address.

Branding

Consistent branding is another key component of building trust and establishing the project's legitimacy. NYC Parks created a study logo that was used on all printed materials and throughout the study's social media presence. Many of the printed materials incorporated the City University of New York School of Public Health and NYC Parks logos to communicate the study's affiliations to these two reputable New York City institutions. Branded promotional materials were widely distributed in the hope that the study brand would also become familiar and trusted within the community. Field staff hung large banners with study branding on study parks' fences and distributed branded, informational materials in community centers, New York City Housing Authority (NYCHA) lobbies and tenant association offices, bodegas, laundromats, schools, day care, and libraries near study sites.

Dimension 2: Adapting the Protocol to Meet Participant Capacity

The second critical dimension to optimize participant engagement was the study's commitment to adapt the proposed protocol to meet participants' capacity to participate. The PARCS study made protocol adaptations as needed in the following areas: (1) scheduled appointments versus rapid deployments, (2) additional sites, and 3) supplemental sample cohort.

Scheduled Appointments Versus Rapid Deployment

The initial study protocol required field staff to screen potential participants to determine eligibility and schedule follow-up appointments. At the follow-up appointment, participants would enroll in the study, receive the accelerometer, and complete the baseline survey in person with the project coordinator. The in-person appointment took approximately 90 minutes. The field staff attempted to maximize participant attendance at the scheduled follow-up appointment by meeting at the time and location of the participants' choice. Field staff also reached out to potential participants before their appointment for confirmation. However, the rate of follow-up meeting attendance was low and would have made it impossible to meet the study recruitment targets on time. As increasing incentives and appointment reminders did not increase appointment attendance, a more substantial protocol change was required. Instead of scheduling follow-up appointments, field staff implemented a *rapid deployment* strategy where residents who met the eligibility criteria were enrolled and received all the study materials at the initial meeting. This strategy boosted recruitment numbers but, as expected, also negatively impacted survey completion adherence, as participants had the option to complete the survey at home without a project coordinator. To mitigate this negative impact, additional outreach, such as phone calls, text messages, and home visits, was added to the protocol.

Additional Sites

Despite adaptations to recruitment protocols, it was still difficult to achieve recruitment targets at some sites. The inclusion criteria (especially the residency requirement of living within a 0.3-mile radius of the study site and having a child between the ages of 3 years and 8 years for the parent-child dyad cohort) limited potential participants to a narrow pool. The sites also varied in terms of population density and zoning. Some study sites primarily included high-density, public housing apartment buildings, whereas others primarily included lower-density single-family residences. Zoning also differed by neighborhood. Some study sites were primarily residential, whereas others included commercial or industrial buildings, limiting recruitment potential. In other sites, geographic features (highways, rivers, etc) limited the number of residential buildings within the recruitment zone. Owing to these recruitment limitations and the associated low enrollment numbers at a few sites and after consultation with the study statistician, four additional sites (three intervention sites and one control site) were added to bolster the study sample size.

Supplemental Sample Cohort

As we anticipated that some participants would move outside of the study zones throughout the life of the study, we specifically recruited NYCHA residences because it is difficult to obtain an NYCHA apartment, and the buildings typically have a low turnover rate. We also only enrolled participants who said they were likely to live in the same residence for the next 4 years. Despite these precautions and accounting for some attrition due to moving, a larger number of participants than expected moved. Between waves 1 and 2 of the study, 183 participants in the adult-only cohort (16% of the cohort at baseline) moved out of the study zone and were no longer eligible to participate.

To maintain the sample size recruited at baseline in subsequent waves, we adapted the protocol to recruit supplemental sample cohorts at each follow-up wave of data collection. The study biostatistician assisted with this decision to develop an appropriate analysis plan. Supplemental sample participants were required to meet the same eligibility criteria as baseline participants. For wave 2, 28% of the adult-only cohort included supplemental sample participants.

Dimension 3: Establishing Operational Flexibility

For the PARCS study, the third key dimension to participant engagement was operational flexibility and the ability to quickly pivot as a team. Field staff needed channels to share relevant information from the field with the management staff who had less direct contact with the participants. In turn, management staff needed flexibility to transform the field teams' insights into protocol adaptations. Furthermore, management staff needed to communicate protocol adaptations to field staff efficiently and be confident that the adaptations would be systematically implemented. The PARCS study sought to increase operational flexibility through (1) switching from teams to one operating unit, (2) case management approach, and (3) reporting mechanisms.

Switch From Teams to One Operating Unit

Initially, the field staff were divided into three teams, each led by a project manager. Each team was responsible for recruitment, follow-up appointments, and participant follow-up for one site at a time. The 3-team structure made it easier to have small group meetings to discuss site-specific insights and challenges and to share site-specific materials. However, having

three site-specific teams was ultimately not flexible enough to meet recruitment needs. With small, site-specific teams, it was not always feasible to schedule appointments with field staff from a specific team at study participants' preferred time and location.

As a result, the research team restructured into one operating unit. This gave project managers more flexibility in recruitment coverage, which, in turn, made it easier to consistently schedule appointments with participants at their preferred time and location and to have sufficient staffing coverage at community events. Having one operating unit increased operational flexibility and made it easier to accommodate last-minute scheduling needs, such as attending a tenant association and meeting or picking up a participant's accelerometer. A centralized approach allowed one project manager to handle scheduling, consistently addressing all scheduling needs and ensuring adequate coverage across all active sites.

Case Management Approach

Another operational challenge was discerning the best way to check in with study participants during their active data collection period and afterward if the materials were not returned. Initially, follow-up calls were made by a project coordinator scheduled to make calls. This approach, however, was unwieldy and did not strengthen the rapport with participants. If a project coordinator did not take meticulous notes, it was easy for a different project coordinator to share redundant or irrelevant information at the next call. If participants needed any follow-up or special attention, it was easy for this information to get lost.

After receiving information about these inefficiencies from the field staff, the staff switched to a case management approach. As part of the switch to case management, when field staff enrolled a new participant, they would indicate which one of them would be the *case manager* and be responsible for participant follow-up. The case manager was then responsible for checking in with the participant 2-3 times over the course of the participant's 7-day active period via text messaging and/or phone calls. The *case manager* was responsible for logging contact attempts and taking notes. By reaching out several times throughout the active data collection period, field staff could remind participants to follow the data collection protocol, answer questions, and troubleshoot any issues. This helped build rapport between the participant and the specific field staff member who enrolled them and thus had a face-to-face connection.

Reporting Mechanisms

By design, there are many external factors that can impact a natural experiment but over which the research team has no control. This makes it critical to establish mechanisms for field staff to report issues as they emerge to project managers. This allows project managers to make protocol adjustments to address the problem and efficiently communicate these changes to the field staff.

For example, in the PARCS study, field staff's early identification of mailing issues was critical for making effective protocol adjustments. As part of the wave 2 data collection protocol, field staff called returning study participants to conduct a brief rescreening survey and explain the wave 2 study protocol. All study materials were subsequently mailed to participants with a preaddressed, stamped return envelope, so that participants could mail back the accelerometer. Some participants reported issues with reliably receiving packages. Other participants reported having difficulty fitting the return envelope with the accelerometer inside their local United States Postal Service (USPS) mailboxes. Upon discovering the latter issue, the team learned that the USPS replaced all the street mailboxes in New York City with a model with a thin slit opening instead of the pull-down drawer in response to rising mailbox theft [40]. The team tried to troubleshoot this with different mailing materials, but the accelerometer dimensions were impossible to fit through the new mailboxes.

In response to these concerns, we added questions to the initial phone survey. First, the participants were asked if they could reliably receive packages at home. If not, participants could provide an alternate address or schedule a time to receive the materials in person. Most participants (96%) reported that they could receive mail at their home address reliably. The call script was also adjusted to let participants know about the blue mailbox design change and to ask if it was still convenient for participants to mail back the materials or if they preferred the materials to be picked up by a project coordinator. As 74% of participants reported that it was still convenient for them to mail the device (often at a nearby post office or workplace), we scheduled individual pick-ups for 26% of participants who requested them.

Dimension 4: Developing Tracking Systems

The last dimension essential to participant engagement has been setting up adaptable measurement and tracking systems. It is essential to develop and track metrics to understand whether protocol adaptations address this concern. In addition to the metrics provided in earlier sections, which tracked the impact of specific protocol adjustments in real time, the team also regularly monitored (1) enrollment and retention and (2) protocol adherence to gauge overall progress and participant engagement.

Tracking Enrollment and Retention

During recruitment periods, project managers tracked the number of participants enrolled per hour worked by each project coordinator on a weekly basis. This helped identify whether anyone needed extra assistance with their recruitment pitch or approach. It also helped to identify whether a site needed an innovative approach or additional resources. A target number of participants to recruit per hour was established, and the staff received monetary incentive bonuses to exceed this target. Enrollment numbers were shared on a biweekly basis to provide frequent performance feedback, help field staff feel a sense of ownership in the project, and provide context for how their work supported the broader project goals.

Field staff used many approaches to connect with returning participants during follow-up waves of data collection, including phone calls, text messages, emails, letters, and flyers. Every contact attempt was tracked. Although it took an average of four contact attempts to connect with returning baseline

participants for the second wave of data collection, the number of contact attempts before connecting ranged from 1 to 30. Tracking participant retention on a weekly basis helped project managers identify the sites to prioritize re-enrollment efforts and supplemental sample recruitment.

Protocol Adherence

Participants were asked to complete an annual survey using the PARCS study app. During the participants' 7-day data collection period, project managers monitored participants' survey completion progress daily through the app's data management interface. This allowed field staff to provide individualized feedback for participants regarding their survey progress and gauge the extent of participant engagement with the study.

After a 7-day study period ended, the participant could no longer access or take the survey via the app. Owing to this and lower-than-anticipated survey completion rates, the research team mailed a paper version of the survey with a dollar bill (as added incentive) to participants who had not yet completed it. In a pilot of this protocol adaptation, 31% of adult participants at baseline who received a paper version of the survey in the mail, completed and mailed it back to us. This demonstrates that protocol adaptation is a useful and worthwhile supporting strategy to maximize survey completion.

Project managers also tracked accelerometer wear adherence and return rates on a weekly basis to further measure participant engagement.

Conclusions

Effective participant engagement strategies are key components of the rigorous study designs needed to further develop the evidence base and better understand the health impacts of built environment interventions. Effective participant engagement practices benefit from strategic adaptations and a research team's ability to communicate, pivot, and iterate. The PARCS study, although ongoing, has centered its participation engagement strategies on the following four dimensions: (1) building trust with communities, (2) adapting the study protocol to meet participants' needs and to reflect their capacity for participation, (3) operational flexibility, and (4) developing tracking systems.

The PARCS study's experience with participant engagement corroborates the best practices from other studies [30-35]. Mistrust and competing priorities emerged as barriers to participation. The study team has frequently adapted and experimented with different engagement strategies to address these barriers. In addition, deploying multiple synergistic strategies is critical to meeting participants' varied needs throughout the course of the study. Careful documentation and tracking systems, where possible, have helped identify engagement problems as they arise and determine the utility of engagement strategies.

Most of the PARCS study's priority dimensions have been designed to specifically address some of the documented participation barriers for low-income neighborhoods and communities of color. As built environment interventions are hypothesized to have the potential to reduce socioeconomic and racial disparities in obesity outcomes and physical activity behaviors, it is critical for studies to include diverse participants to build a relevant evidence base.

A limitation of this paper is that although the research team attempted to track the impact of engagement strategies when possible, the tracking systems were added after the start of the study and were not necessarily designed to formally and empirically test the strategies. Future research studies could consider developing tracking systems *a priori* so that strategies for participant engagement and operational flexibility could be more rigorously tested. This is a potential area for future research.

This paper shares practical lessons about iterative, dynamic strategies to improve participant engagement and operational flexibility based on the experience to date in an ongoing longitudinal study in low-income communities and communities of color. Insights learned from the PARCS study may help other research teams effectively anticipate and respond to participant engagement challenges in future community-based studies.

Acknowledgments

The authors thank the New York City residents who participated in this study and the New York City Department of Parks and Recreation for their support. This study was supported by funding (principal investigator: TH for all projects) from the National Cancer Institute (R01CA206877), the New York State Health Foundation (#16-04236), Bryant Park Corporation, the Robert Wood Johnson Foundation (#76473), and the CDC (5 U48DP006396-02-00).

Authors' Contributions

EF is the principal author of this paper. TH oversaw the conceptual design and development of this study. KW, KE, JD, LT, DC, and TH contributed to the design of the study and writing of the paper.

Conflicts of Interest

None declared.

References

1. Ward ZJ, Bleich SN, Cradock AL, Barrett JL, Giles CM, Flax C, et al. Projected U.S. State-Level Prevalence of Adult Obesity and Severe Obesity. N Engl J Med 2019 Dec 19;381(25):2440-2450. [doi: 10.1056/NEJMsa1909301] [Medline: 31851800]
2. Centers for Disease Control and Prevention. The Health Effects of Overweight and Obesity. The Health Effects of Overweight and Obesity. 2020 Sep 17. URL: https://www.cdc.gov/healthyweight/effects/index.html [accessed 2021-02-22]
3. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011-2012. JAMA 2014 Feb 26;311(8):806-814 [FREE Full text] [doi: 10.1001/jama.2014.732] [Medline: 24570244]
4. Flegal KM, Kruszon-Moran D, Carroll MD, Fryar CD, Ogden CL. Trends in Obesity Among Adults in the United States, 2005 to 2014. JAMA 2016 Jun 07;315(21):2284-2291. [doi: 10.1001/jama.2016.6458] [Medline: 27272580]
5. Skinner AC, Skelton JA. Prevalence and trends in obesity and severe obesity among children in the United States, 1999-2012. JAMA Pediatr 2014 Jun;168(6):561-566. [doi: 10.1001/jamapediatrics.2014.21] [Medline: 24710576]
6. Barrington D, Baquero M, Borrell L, Crawford ND. Racial/ethnic disparities in obesity among US-born and foreign-born adults by sex and education. Obesity (Silver Spring) 2010 Feb;18(2):422-424 [FREE Full text] [doi: 10.1038/oby.2009.212] [Medline: 19590501]
7. Cossrow N, Falkner B. Race/ethnic issues in obesity and obesity-related comorbidities. J Clin Endocrinol Metab 2004 Jun;89(6):2590-2594. [doi: 10.1210/jc.2004-0339] [Medline: 15181028]
8. Zhang H, Rodriguez-Monguio R. Racial disparities in the risk of developing obesity-related diseases: a cross-sectional study. Ethn Dis 2012;22(3):308-316. [Medline: 22870574]
9. Wong RJ, Chou C, Ahmed A. Long term trends and racial/ethnic disparities in the prevalence of obesity. J Community Health 2014 Dec;39(6):1150-1160. [doi: 10.1007/s10900-014-9870-6] [Medline: 24715435]
10. Hales C, Carroll M, Fryar C, Ogden CL. Prevalence of Obesity Among Adults and Youth: United States, 2015-2016. NCHS Data Brief 2017 Oct(288):1-8 [FREE Full text] [Medline: 29155689]
11. Ogden CL, Fakhouri TH, Carroll MD, Hales CM, Fryar CD, Li X, et al. Prevalence of Obesity Among Adults, by Household Income and Education — United States, 2011–2014. MMWR Morb Mortal Wkly Rep 2017 Dec 22;66(50):1369-1373. [doi: 10.15585/mmwr.mm6650a1]
12. NCHS, National Health Interview Survey. Age-sex-adjusted percentage of adults aged 18 and over who met 2008 federal physical activity guidelines for aerobic activity through leisure-time aerobic activity, by race and ethnicity: United States, 2018. 2018 Feb 12. URL: https://public.tableau.com/profile/nhis6957#!/vizhome/FIGURE7_3/Dashboard7_3 [accessed 2021-02-22]
13. Huang T, Drewnosksi A, Kumanyika S, Glass TA. A systems-oriented multilevel framework for addressing obesity in the 21st century. Prev Chronic Dis 2009 Jul;6(3):A82 [FREE Full text] [Medline: 19527584]
14. McCormack GR, Shiell A. In search of causality: a systematic review of the relationship between the built environment and physical activity among adults. Int J Behav Nutr Phys Act 2011 Nov 13;8:125 [FREE Full text] [doi: 10.1186/1479-5868-8-125] [Medline: 22077952]
15. Mayne SL, Auchincloss AH, Michael YL. Impact of policy and built environment changes on obesity-related outcomes: a systematic review of naturally occurring experiments. Obes Rev 2015 May;16(5):362-375. [doi: 10.1111/obr.12269] [Medline: 25753170]
16. Gordon-Larsen P, Nelson MC, Page P, Popkin BM. Inequality in the built environment underlies key health disparities in physical activity and obesity. Pediatrics 2006 Feb;117(2):417-424. [doi: 10.1542/peds.2005-0058] [Medline: 16452361]
17. Williams DR, Collins C. Racial residential segregation: a fundamental cause of racial disparities in health. Public Health Rep 2001;116(5):404-416. [doi: 10.1093/phr/116.5.404] [Medline: 12042604]
18. Sallis JF, Saelens BE, Frank LD, Conway TL, Slymen DJ, Cain KL, et al. Neighborhood built environment and income: examining multiple health outcomes. Soc Sci Med 2009 Apr;68(7):1285-1293 [FREE Full text] [doi: 10.1016/j.socscimed.2009.01.017] [Medline: 19232809]
19. Cohen DA, Han B, Isacoff J, Shulaker B, Williamson S, Marsh T, et al. Impact of park renovations on park use and park-based physical activity. J Phys Act Health 2015 Feb;12(2):289-295 [FREE Full text] [doi: 10.1123/jpah.2013-0165] [Medline: 24956608]
20. Cohen DA, Han B, Isacoff J, Shulaker B, Williamson S. Renovations of neighbourhood parks: long-term outcomes on physical activity. J Epidemiol Community Health 2019 Mar;73(3):214-218. [doi: 10.1136/jech-2018-210791] [Medline: 30606773]
21. Sallis JF, Slymen DJ, Conway TL, Frank LD, Saelens BE, Cain K, et al. Income disparities in perceived neighborhood built and social environment attributes. Health Place 2011 Nov;17(6):1274-1283. [doi: 10.1016/j.healthplace.2011.02.006] [Medline: 21885324]
22. Oreskovic NM, Kuhlthau KA, Romm D, Perrin JM. Built environment and weight disparities among children in high- and low-income towns. Acad Pediatr 2009;9(5):315-321. [doi: 10.1016/j.acap.2009.02.009] [Medline: 19477705]
23. Molina-García J, Queralt A, Adams MA, Conway TL, Sallis JF. Neighborhood built environment and socio-economic status in relation to multiple health outcomes in adolescents. Prev Med 2017 Dec;105:88-94. [doi: 10.1016/j.ypmed.2017.08.026] [Medline: 28863871]

24. Tseng E, Zhang A, Shogbesan O, Gudzune KA, Wilson RF, Kharrazi H, et al. Effectiveness of Policies and Programs to Combat Adult Obesity: a Systematic Review. J Gen Intern Med 2018 Nov;33(11):1990-2001 [FREE Full text] [doi: 10.1007/s11606-018-4619-z] [Medline: 30206789]
25. Ferdinand AO, Sen B, Rahurkar S, Engler S, Menachemi N. The relationship between built environments and physical activity: a systematic review. Am J Public Health 2012 Oct;102(10):e7-e13. [doi: 10.2105/AJPH.2012.300740] [Medline: 22897546]
26. Audrey S, Batista-Ferrer H. Healthy urban environments for children and young people: A systematic review of intervention studies. Health Place 2015 Nov;36:97-117 [FREE Full text] [doi: 10.1016/j.healthplace.2015.09.004] [Medline: 26457624]
27. Sallis JF, Glanz K. The role of built environments in physical activity, eating, and obesity in childhood. Future Child 2006;16(1):89-108. [doi: 10.1353/foc.2006.0009] [Medline: 16532660]
28. Craig P, Cooper C, Gunnell D, Haw S, Lawson K, Macintyre S, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. J Epidemiol Community Health 2012 Dec;66(12):1182-1186 [FREE Full text] [doi: 10.1136/jech-2011-200375] [Medline: 22577181]
29. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
30. Hill KG, Woodward D, Woelfel T, Hawkins JD, Green S. Planning for Long-Term Follow-Up: Strategies Learned from Longitudinal Studies. Prev Sci 2016 Oct;17(7):806-818 [FREE Full text] [doi: 10.1007/s11121-015-0610-7] [Medline: 26453453]
31. Ejiogu N, Norbeck JH, Mason MA, Cromwell BC, Zonderman AB, Evans MK. Recruitment and retention strategies for minority or poor clinical research participants: lessons from the Healthy Aging in Neighborhoods of Diversity across the Life Span study. Gerontologist 2011 Jun;51 Suppl 1:S33-S45 [FREE Full text] [doi: 10.1093/geront/gnr027] [Medline: 21565817]
32. George S, Duran N, Norris K. A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders. Am J Public Health 2014 Feb;104(2):e16-e31. [doi: 10.2105/AJPH.2013.301706] [Medline: 24328648]
33. Abshire M, Dinglas VD, Cajita MIA, Eakin MN, Needham DM, Himmelfarb CD. Participant retention practices in longitudinal clinical research studies with high retention rates. BMC Med Res Methodol 2017 Feb 20;17(1):30 [FREE Full text] [doi: 10.1186/s12874-017-0310-z] [Medline: 28219336]
34. Robinson KA, Dinglas VD, Sukrithan V, Yalamanchilli R, Mendez-Tellez PA, Dennison-Himmelfarb C, et al. Updated systematic review identifies substantial number of retention strategies: using more strategies retains more study participants. J Clin Epidemiol 2015 Dec;68(12):1481-1487 [FREE Full text] [doi: 10.1016/j.jclinepi.2015.04.013] [Medline: 26186981]
35. Brueton VC, Tierney JF, Stenning S, Meredith S, Harding S, Nazareth I, et al. Strategies to improve retention in randomised trials: a Cochrane systematic review and meta-analysis. BMJ Open 2014 Feb 04;4(2):e003821 [FREE Full text] [doi: 10.1136/bmjopen-2013-003821] [Medline: 24496696]
36. de Jong JPJ. Surveying innovation in samples of individual end consumers. European Journal of Innovation Management 2016 2016;19(3):406-423. [doi: 10.1108/ejim-09-2015-0093]
37. Huang TTK, Wyka KE, Ferris EB, Gardner J, Evenson KR, Tripathi D, et al. The Physical Activity and Redesigned Community Spaces (PARCS) Study: Protocol of a natural experiment to investigate the impact of citywide park redesign and renovation. BMC Public Health 2016 Nov 14;16(1):1160. [doi: 10.1186/s12889-016-3822-2] [Medline: 27842531]
38. Cato MS, Wyka K, Ferris EB, Evenson KR, Wen F, Dorn JM, et al. Correlates of accelerometry non-adherence in an economically disadvantaged minority urban adult population. J Sci Med Sport 2020 Aug;23(8):746-752. [doi: 10.1016/j.jsams.2020.01.013] [Medline: 32085979]
39. Swierad EM, Huang TTK. An Exploration of Psychosocial Pathways of Parks' Effects on Health: A Qualitative Study. IJERPH 2018 Aug 08;15(8):1693. [doi: 10.3390/ijerph15081693]
40. Angi Gonzalez. USPS replacing all NYC mailboxes for new 'anti-fishing model'. 2019 Jan 29. URL: https://tinyurl.com/2c9xrdex [accessed 2021-02-22] [WebCite Cache ID https://tinyurl.com/2c9xrdex]

Abbreviations

NYC Parks: New York City Department of Parks and Recreation
NYCHA: New York City Housing Authority
PARCS: Physical Activity and Redesigned Community Spaces
SES: socioeconomic status
USPS: United States Postal Service

Edited by G Eysenbach; submitted 10.03.20; peer-reviewed by R Berkowitz, R Mychasiuk; comments to author 15.06.20; revised version received 30.09.20; accepted 20.01.21; published 24.03.21.

Please cite as:
Ferris EB, Wyka K, Evenson KR, Dorn JM, Thorpe L, Catellier D, Huang TTK
Recruitment and Retention Strategies for Community-Based Longitudinal Studies in Diverse Urban Neighborhoods
JMIR Form Res 2021;5(3):e18591
URL: https://formative.jmir.org/2021/3/e18591
doi:10.2196/18591
PMID:33759799

Review

Investigating the Food and Drug Administration Biotherapeutics Review and Approval Process: Narrative Review

Samuel Bonet Olivencia[1*], BSc, MSc; Farzan Sasangohar[1,2*], BA, BCS, MASc, SM, PhD

[1]Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX, United States

[2]Center for Outcomes Research, Houston Methodist Hospital, Houston, TX, United States

[*]all authors contributed equally

Corresponding Author:
Farzan Sasangohar, BA, BCS, MASc, SM, PhD
Department of Industrial and Systems Engineering
Texas A&M University
3131 TAMU
College Station, TX, 77843
United States
Phone: 1 9794582337
Email: sasangohar@tamu.edu

Abstract

Background: The development, review, and approval process of therapeutic biological products in the United States presents two primary challenges: time and cost. Advancing a biotherapeutic from concept to market may take an average of 12 years, with costs exceeding US $1 billion, and the product may still fail the US Food and Drug Administration (FDA) approval process. Despite the FDA's practices to expedite the approval of new therapies, seeking FDA approval remains a long, costly, and risky process.

Objective: The objective of this paper is to explore the factors and gaps related to the FDA review and approval process that contribute to process inefficiencies and complexities as well as proposed methods and solutions to address such gaps. This paper also aims to investigate the available modeling efforts for the FDA approval process of therapeutic biological products.

Methods: A narrative review of literature was conducted to understand the scope of published knowledge about challenges, opportunities, and specific methods to address the factors and gaps related to the review and approval of new drugs, including therapeutic biological products. Relevant peer-reviewed journal articles, conference proceedings, book chapters, official reports from public policy professional centers, and official reports and guidelines from the FDA were reviewed.

Results: Of the 23 articles identified in this narrative literature review, none modeled the current FDA review and approval process structure to address issues related to the robustness, reliability, and efficiency of its operations from an external point of view. Although several studies summarize the FDA approval process with clarity, in addition to bringing to light the problems and challenges faced by the regulatory agency, only a few attempts have been made to provide solutions for the problems and challenges identified. In addition, although several reform models have been discussed, these models lack the application of scientific methodologies and modeling techniques in understanding FDA as a complex sociotechnical system. Furthermore, tools and methods to assess the efficacy of the models before implementation are largely absent.

Conclusions: The findings suggest the efficacy of model-based systems engineering approaches for identifying opportunities for significant improvements to the FDA review and approval process. Using this holistic approach will serve several investigative purposes: identify influential sources of variability that cause major delays, including individual, team, and organizational decision making; identify the human-system bottlenecks; identify areas of opportunity for design-driven improvements; study the effect of induced changes in the system; and assess the robustness of the structure of the FDA approval process in terms of enforcement and information symmetry.

(JMIR Form Res 2021;5(3):e14563) doi:10.2196/14563

KEYWORDS

biotherapeutics; drug approval; drug review process; model-based systems engineering

Introduction

Background

The introduction of new medicines and treatments into the market is a time- and cost-consuming process that is closely supervised and regulated to ensure the safety and effectiveness of the products. This process takes, on average, 12 years, and the estimated average cost of taking a new drug from concept to market exceeds US $1 billion [1]. After significant expenditure of the manufacturer's time and resources, many drugs fail to achieve approval late in the process.

Regulatory agencies, such as the European Medicines Agency (EMA) and Health Canada, are responsible for promoting and protecting public health in their respective geographic areas through the evaluation and supervision of medicines for humans before their release into the market. In the United States, these regulatory functions fall under the responsibilities of the US Food and Drug Administration (FDA), which is the oldest comprehensive consumer protection agency in the US Federal Government [2]. Factors such as expanded federal regulations, increasing complexity of drugs and devices, and the growth of the pharmaceutical industry have expanded the role of the FDA, which is now one of the largest consumer safety agencies in the world [1].

The increasingly complex regulatory environment and expenses associated with drug development have been criticized for the resultant lag in the release of new pharmaceuticals into the drug market. In addition, the FDA risk aversion approach has forced companies to go overseas and has encouraged medical tourism [3]. A subset of advocacy groups and experts in drug regulation and policies are demanding more rapid development, approval, and release of new products because they consider the current process to be risk averse, slow, and inefficient [4]. In the past, numerous major safety incidents occurred because of drugs released into the market with little or no regulation from the FDA. As a result, weakening or removing FDA regulations for pharmaceuticals is not an option [1]. However, the FDA has created programs to facilitate the development and expedite the approval of drugs that treat serious conditions or fill an unmet medical need [5]. These drugs receive a fast-track designation, which makes them eligible for (1) more frequent meetings and written communication with the FDA; (2) accelerated approval or priority review; and (3) rolling review, meaning that the review application can be submitted and reviewed in sections rather than as a complete application. In addition, the US Congress issued the Prescription Drug User Fee Act (PDUFA) to authorize the FDA to collect fees from applicant companies to invest in resources to accelerate FDA review and approval operations [6]. Despite these efforts, FDA scrutiny remains a long, costly, and risky process, making it challenging for patients to have timely access to potentially useful medications and treatments [7].

As with any new drug, the FDA's long scrutiny affects the review and approval of new biopharmaceutical products, which have become an important sector of the pharmaceutical industry in the United States, the country with the largest market for biopharmaceuticals (around 33% of the global market) [8]. Reports from the National Science Foundation reveal that the US biopharmaceutical sector accounts for the largest single share of all US investments in research and development (R&D) [8]. This fast-growing sector is in a critical position in which therapeutic biological products represent over a third of all new drugs in clinical trials or are awaiting approval from the US FDA [8].

In the past, there has been a clear difference between biotechnology industries and pharmaceutical industries; however, this has become increasingly blurred as many pharmaceutical industries are increasing their presence in biopharma. These large companies have the capital to invest in high-risk research, development, review, and licensing of new therapeutic biological products. However, there is a relevant sector of start-up biotechnology companies that are more vulnerable to the high risks and uncertainties in product development. The biopharmaceutical development process is complex, and small companies must invest significant amounts of capital and human resources. They assume risks in terms of execution, safety, efficacy, and license approval during the regulatory process. The initial financial conditions play a significant role in the financial and strategic planning of the start-up companies. Many of these start-up companies finance their R&D projects by partnering with big pharmaceutical companies or other bioscience corporate partners, securing capital from angel investors, or obtaining government or small business grants [9]. Large pharmaceutical companies, which have the capabilities, operational scale, and capital, create corporate alliances with small biotech companies with the understanding that the larger pharmaceutical entity will have a considerable amount of control and profits over the start-up. Statistics show that there is a 20% likelihood for a bioproduct to progress from the initiation of phase 1 clinical trials all the way to market approval [9]. When partnering with a start-up, companies, sponsors, and investors assume a large portion of the risk and uncertainty associated with the development of the therapeutic biological products. Although there is a risk associated with the upfront investment in R&D and clinical trials, estimated to be in tens of millions of US dollars, there is also a considerable amount of risk regarding the time and cost associated with navigating the FDA approval and licensing process for a biopharmaceutical product [9]. In addition to the associated time and cost complexities, there are other challenges, some unique and others in common with any other new drug, that characterize the release of a therapeutic biological product into the market.

Objectives

The goal of this paper is to explore the factors and gaps related to the FDA drug review and approval process that contribute to process inefficiencies and complexities as well as proposed methods and solutions to address such gaps. The focus of this paper is to understand the constraints and challenges in the drug review and approval process identified by researchers who investigated FDA operations and to identify the models that have been applied to understand, evaluate, analyze, and suggest improvements to the FDA drug review and approval process.

Methods

A narrative review of the literature was conducted to understand the scope of the published peer-reviewed knowledge about challenges, opportunities, and specific methods to address these factors and gaps related to the review and approval of new drugs, including therapeutic biological products.

Search Strategy

First, official reports and guidelines from the FDA official website were retrieved to document the biotherapeutics review and approval process. The authors familiarized themselves with the FDA review and approval process for new drugs, including therapeutic biological products, and identified relevant terms for use in the search of relevant peer-reviewed journal articles, conference proceedings, book chapters, and official sources from public policy professional centers. Several databases including Science Direct, ABI/Inform Complete-ProQuest, MEDLINE (PubMed), and Google Scholar were searched using the terms “FDA” combined with “regulation” AND “model” AND “approval” to identify sources related to the FDA review and approval process. In addition, the term “FDA” was combined with “systems engineering,” “change management,” and “quality by design” for the initial search, to identify sources proposing models and alternative approaches to improve and reform the FDA review and approval process.

Study Selection

The first search yielded 11,296 sources. These sources were screened, and a total of 9400 sources were excluded based on the following exclusion criteria: (1) inaccessibility to full text, (2) not peer-reviewed (exceptions were made for articles coming from public policy professional centers), and (3) duplications. The remaining 1896 sources were assessed for eligibility based on the following inclusion criteria: (1) published between 2000 and 2017; (2) included cost and time considerations; (3) included content about the FDA review and approval process for medical devices, drugs, and biotherapeutics, which are complex products with a longer and generally more costly pathway for approval in comparison with other FDA regulatory processes (medical devices were not excluded from the search to identify work in that area that could be applicable to the review and approval process of drugs and biotherapeutics) [10]; and (4) included content about the operationalization of the FDA review and approval process. On the basis of these criteria, 41 sources were selected for full-text review. The full-text review of the remaining 41 sources was conducted to include sources that were most relevant to the scope of this study: (1) understand the constraints and challenges in the drug review and approval process identified by researchers who investigated FDA operations and (2) identify the models that have been applied to understand, evaluate, analyze, and suggest improvements to the FDA drug review and approval process. After the full-text review, 23 sources, including peer-reviewed articles, conference proceedings, book chapters, and articles from public policy professional centers, were included in the narrative review (Table 1).

Table 1. Sources included in the narrative review.

Sources (reference)	Title
Conko and Madden [7]	Administrative Law and Regulation
Das and Almonor [11]	A Concurrent Engineering Approach for the Development of Medical Devices
Alexander and Clarkson [12]	A Validation Model for the Medical Devices Industry
Tsai and Erickson [9]	Early-Stage Biotech Companies: Strategies for Survival and Growth
Rathore and Winkle [13]	Quality by Design for Biopharmaceuticals
Gernaey and Gani [14]	A Model-Based Systems Approach to Pharmaceutical Product-Process Design and Analysis
Kourti and Davis [15]	The Business Benefits of Quality by Design (QbD)
Medina et al [16]	Design for FDA: A Predictive Model for the FDA's Decision Time for Medical Devices
Medina et al [17]	Supporting Medical Device Development: A Standard Product Design Process Model
Baylor [18]	Regulatory Approval and Compliances for Biotechnology Products
Conner et al [19]	The Biomanufacturing of Biotechnology Products
Lawrence et al [20]	Understanding Pharmaceutical Quality by Design
Aksu et al [21]	QbD Implementation in Biotechnological Product Development Studies
Briggeman et al [22]	The Proper Role of the FDA for the 21st Century
Kinch [23]	2015 in Review: FDA Approval of New Drugs
Thierer and Wilt [24]	The Need for FDA Reform: Four Models
Van Norman [1]	Drugs, Devices, and the FDA: Part 1: An Overview of Approval Processes for Drugs
Van Norman [10]	Drugs and Devices: Comparison of European and US Approval Processes
Williams [4]	Food and Drug Administration Drug Approval Process: A History and Overview
Williams et al [3]	Health Options Foreclosed: How the FDA Denies Americans the Benefits of Medical Research
Djuris and Djuric [25]	Modeling in the Quality by Design Environment: Regulatory Requirements and Recommendations for Design Space and Control Strategy Appointment
Horner et al [26]	Process Modeling in the Biopharmaceutical Industry
Joshi et al [27]	Optimization of Ion Exchange Sigmoidal Gradients Using Hybrid Models: Implementation of Quality by Design in Analytical Method Development

Results

Review and Approval of New Drugs

The overall timeline for the release of a new drug into the market can be divided into 2 main phases: (1) R&D and (2) review, approval, and licensing. The first stage of the R&D phase, which includes preclinical trials and 3 phases of clinical trial, involves an exchange of information between the applicant and the FDA review board. Both parties are in communication, and after each stage, the applicant must provide updated information to the FDA. Even when this flow of updated information is established, the license approval of a new drug application may take a long time because of the prolonged FDA application review process. According to the rules and procedures of the FDA, the review process of a new drug application cannot start unless the application is fully submitted. An exception is made for those new drugs accepted under a rolling review status, where the application can be submitted and reviewed in sections.

The review of a new drug application consists of 6 major steps that include reviewing clinical trial results, planning and execution of the product label, and manufacturing site inspections. In the specific case of therapeutic biological products, the current FDA review and licensing are regulated following the guidelines of *The Program*, a review program created by the reauthorization of the PDUFA Act in 2018 (PDUFA VI), which is valid for the duration of the current version (until the year 2022). *The Program* was created with the intention of increasing the communication and transparency between the FDA review team and the applicant, to increase the efficiency and effectiveness of the first review cycle. In addition, *The Program* decreases the number of review cycles necessary to approve a biologics license agreement (BLA), which is a highly complex application [28]. These revisions also provide additional review clock time for the agency to meet with the applicant during review as well as to address review activities that occur late in the review cycle [28]. According to the timeline established in the Center for Drug Evaluation and Research (CDER) 21st Century Review Process Desk Reference Guide, the total estimated time for a standard review of a BLA under *The Program* is 12 months, and it is 8 months for a priority review [29]. All these efforts to expedite and streamline the process come with a cost that is pushed to the applicant. PDUFA VI authorizes the FDA to collect fees from companies that submit BLAs. For example, an application including clinical

trials may have a fee of approximately US $2.5 million dollars [6].

In addition to the time and money constraints in the process, the review and approval of a drug involve multiple FDA resources and constant communication with the applicant through phone calls, emails, and meetings. The personnel assigned to review an application vary according to the type of submission and product. In general, an application is reviewed by a team of professionals from different disciplines. The review team has to deal with the flow of not only new submissions but also resubmissions. A company can resubmit an application to answer all the deficiencies indicated by the FDA in the initial review [30]. The resubmissions may put a strain on FDA's normal operations by sharing resources between both types of submissions.

Other constraints and challenges have been identified by researchers who investigated FDA's operations [1,4,10,18,19,23]. Van Norman [1] points out the complexity of the approval process and emphasizes that the main challenges for pharmaceuticals are in terms of cost and time. Similarly, Williams [4] emphasizes the criticism and controversy that the FDA has confronted because of the lengthy approval process and that the FDA has been accused of conflict of interest because of the user fees collected from sponsors and drug manufacturers to support the drug approval process. In addition, the author discusses critics' claims that FDA's operations are slower and less efficient than that of the EMA in the European Union (EU), even when there is little evidence available to support the criticism [4]. Van Norman [10] compared the European and US approval processes with the purpose of presenting their similarities and differences as well as the perceived challenges faced by each. His work emphasizes how both approval processes are similar, except that the US approval process is completely centralized, whereas the EU has 4 possible pathways for drug approval: (1) centralized through the EMA (mandatory for some classes of drugs, such as those used in the treatment of oncological diseases and diabetes, among others), (2) national (each EU state has its own procedure), (3) by mutual recognition (drugs approved in one EU state can obtain marketing authorization in another EU state), and (4) decentralized (manufacturers can simultaneously apply for authorization in more than one EU state). In addition, Van Norman [10] presents data that weaken the claim that the FDA is significantly slower than the EMA.

Another relevant challenge faced by the FDA is the lack of transparency in nonpublished drug trial data. According to Van Norman [10], this issue results in challenges associated with the production of systematic reviews and meta-analyses that are essential to public health and safety. Promoting information symmetry (information equally accessible to all parts involved, including regulators, industry representatives, and consumers) must be a central function of the FDA. This issue suggests the need for an assessment of the current review and approval structure. In addition, it is necessary to investigate methods to support individual, team, and organizational decision making to balance the process structure in terms of enforcement and information.

In a similar manner, pharmaceutical quality oversight has been a major issue that the FDA has addressed over the years, establishing strict regulatory standards to ensure the safety and efficacy of the products. Modeling to comply with quality regulation standards has been another research area of interest in the academic community. Specifically, importance has been given to formulation and process modeling, addressing the quality by design (QbD) concept, first introduced in 2004 as part of the Pharmaceutical cGMPs (Current Good Manufacturing Practices) for the 21st Century Initiative [13,25,27]. QbD assists both the industry and the FDA in implementing a scientific approach by targeting the desired product quality throughout the design and development process. This is different from the traditional quality by testing methodology as it encourages the definition of a design space early in the process development. The design space works as an acceptable operating range for the critical process parameters. Changes in the process within the design space are acceptable by the FDA. Any movement outside the design space is considered to be a change that must be approved by the FDA. Consequently, a regulatory postapproval process must be initiated by the manufacturer and submitted to the FDA for review and approval [13].

The QbD approach was implemented in other FDA review programs before being adopted by the Office of Biotechnology Products because of the complexity of its application in the development and manufacturing process of biotechnology products. Since the implementation of QbD principles for pharmaceutical development, multiple benefits have been identified, but certain challenges have been raised, for example, (1) determining common terminology between the industry and the FDA [20], (2) the provision of training programs to industry representatives [21], (3) lack of understanding and trust among all stakeholders involved [21], and (4) the associated costs of implementing QbD in product development and regulatory processes [21]. One of the major challenges in the implementation of QbD has been to manage the surveillance of legacy products approved before the implementation of the QbD principles [13]. The integration of QbD into the drug development process adds a new complete level of interconnections and communications between the industry and the FDA. A common understanding of QbD and the steps involved is necessary to facilitate communication between both parties [15,20]. Therefore, modeling and architecting the FDA network flow structure will require addressing the inter- and intraconnections in the drug application review process. In addition, such modeling approaches must provide a clear representation that leads to a common understanding of the QbD approach, highlighting the elements relevant to the entity in charge of the risk-based drug development process (industry) as well as the entity responsible for reviewing and monitoring the drug application submissions (FDA).

Modeling the FDA Regulatory Process

Models and Points of View

Most of the reviewed academic literature that modeled FDA regulatory processes presents efforts in modeling from the applicant's point of view. The purpose of modeling from this view is to enhance the chances of compliance with the regulatory

requirements and recommendations. Most work in this area applies to the development of medical devices, with a special focus on design [11,12,16,17]. For example, Medina et al [17] developed a standard product design process model to support medical device development. Even when the work includes FDA regulations that complement the medical device development process, the focus is directed toward modeling the process to comply with FDA regulations rather than addressing the efficiency and robustness of the FDA regulatory process. A relevant aspect of their work is the application of model-based systems engineering (MBSE) tools to model the medical device design and development process. Medina et al [16] applied unified modeling language to model the relationships among the different elements (classes) in the development process. This research was later extended to the development of a predictive tool to estimate the FDA decision time for medical device approval, with the purpose of estimating the product's time to market [16]. The relevance of applying an MBSE approach to the FDA approval process is because of its suitability for analyzing complex systems as networks of interrelated elements that include people, facilities, policies, laws, regulations, internal and external institutions, and technologies, among other elements. In a system as complex as the FDA drug approval process, the application of an MBSE approach may provide the following benefits: (1) facilitate communication among the various stakeholders involved; (2) provide a set of models (abstractions to manage size and complexity) that serve as a tool to analyze the effect of changes to the system; (3) allow compare and contrast analysis of the *as-is* and *to-be* solutions; and (4) allow the exploration of multiple system alternatives concurrently with minimal risk, among other benefits [31].

A review of the available literature reveals a lack of research on the approval process from the regulatory agency's point of view. In addition, only a few studies have applied modeling approaches with direct application to pharmaceutical products and process design and analysis [13,27]. Most research efforts were industry-driven and directed toward the incorporation of modeling tools into drug development and production practices, with the purpose of enhancing the process by reducing the cost and time to market of the products [14,26].

Reform Models and Change Management

In a stochastic system typical of the FDA, regulation procedures can undergo changes at any time. The pressure that the FDA is receiving to relax the scrutiny process and expedite the time to market of critical pharmaceuticals increases the urgency to enhance and reform the regulatory agency approval process. Following this line of thinking, Thierer and Wilt [24] presented a summary of 4 models to reform the FDA approval process and change the way medical products are brought to the market. These 4 models are flexible approaches presented under the premise that the implementation of a comprehensive reform would benefit both the innovators and the patients.

The first model, developed by Williams et al [3], targets the approval of medical devices by suggesting the provision of regulatory authority to multiple private parties that will compete with the FDA on the price, quality, and timeliness of approvals. This model not only creates a competition for trust between the FDA and the private regulatory bodies but also requires the FDA to fulfill a new role of establishing quality standards and good manufacturing practices that must be monitored by private regulatory parties. As established by the authors, this reform model can also be applied to the approval of drugs [24]. This idea was first presented in 2000 by Henry I Miller [32], who proposed the creation of nongovernmental drug-certifying bodies, changing the role of the FDA from being the *certifier of products* to being the *certifier of the certifiers* [33]. An implementation of a model such as this will create a significant change in the drug approval process as new stakeholders will be integrated into the review process, whereas the FDA's responsibility would be limited to a higher-level review. If analyzed as a network, this will require a new set of interrelations and modified information flow.

The second reform model was discussed by Klein and Tabarrok [33], the objective of which was to eliminate the FDA monopoly on drug approval by allowing manufacturers to market their products in the United States once they have gained approval in other major markets, such as the EU [24,33]. The term *international reciprocity* is used to refer to this approach. Similar to the first approach, this model requires the FDA to compete with other international regulatory agencies for the business of drug manufacturers.

The third model presented is more in line with the perspective of the current critics of FDA practices. The model proposes not having to wait until further clinical trials to release a drug for use when safety and efficacy have been demonstrated in the initial clinical trials. A proposal which can be referred to as *free-to-choose medicine*, developed by Conko and Madden [7], suggests a dual-track system for patients and doctors consisting of (1) deciding to stick with the current FDA procedures or instead (2) selecting a *free-to-choose track* option that provides freedom to the patient, with the advice of their doctors, to make an informed choice of using an experimental drug for which safety and efficacy have been demonstrated but FDA approval has not yet been obtained [7,24]. The implementation of this dual-track system requires a commitment from the FDA and the manufacturing companies to promote information symmetry. This reform is similar to what the Independent Institute calls *the sensible alternative*, which supports voluntary certification. This alternative proposes to keep the FDA as a voluntary institution where companies submit their application because of a belief in the integrity and cooperation of the agency, not because it is mandatory. Drugs approved under this alternative must be labeled as *not FDA approved* [33].

The fourth model, established by Briggeman et al [22], was developed based on the idea that the FDA has exceeded its authority by not only assuring safety and efficacy but also making judgments about the benefits and risks of the drugs. According to the authors, that responsibility lies with doctors and patients based on their experiences with the drug [22]. The reform model suggests that the FDA must readopt the regulation model they followed in the 80s and 90s. In this model, the regulatory agency is at the top of the funnel, setting standards to measure a drug's effectiveness based on the pharmacologic activity on the disease. The responsibility of determining the

utility of the drug is deferred to the doctors, who must make decisions regarding the adoption of a drug based on real-world experiences, the characteristics of each patient, and additional clinical trials sponsored by clinics and biopharmaceutical industries [22].

The implementation of any of these 4 reform models may imply a shift in the current activities and responsibilities of the FDA and, therefore, may change the organizational structure of the regulatory agency. The addition of new internal or external regulatory sources as well as the elimination of current regulatory bodies within the FDA will have a variety of consequences in the agency network, which must be addressed and measured for effectiveness. Current FDA publications and reports on change management were created to provide guidance to industry applicants on how to manage changes in their processes and still comply with the quality regulations [20,25]. To deal specifically with organizational changes, the FDA has developed a set of manuals of policies and procedures for each of its offices. For example, the Manual of Policies and Procedures for the CDER provides policies and procedures for submitting, evaluating, coordinating, reviewing, and approving organizational changes, including the addition or elimination of organizational components [30]. However, to the best of our knowledge, no FDA publications or academic research has presented efforts to develop a flexible system-based model that shows the interconnections and information flow among the main elements of the approval process. A contribution in this area will provide the FDA with a diagnostic tool to efficiently address the change management affairs.

Discussion

Principal Findings

The FDA has made efforts to improve their internal drug review and approval operations through the creation of expedited review and approval tracks and the establishment of special review programs such as *The Program*, created under the PDUFA. Although these efforts have reduced the time to market of drugs, the drug approval process remains a costly and slow process. These challenges have a major impact on small-to-medium biopharma start-ups that do not have the initial capital investment required and must resort to alternate financial agreements that are not necessarily competitive for them. More importantly, it remains challenging for chronically ill patients to have timely access to novel alternative medications and treatments that could save their lives.

Suggesting changes to the review and approval of therapeutic biological products is a challenging task because of the complexity of the process. Each therapeutic biological product has different properties, meaning that each FDA review and approval process is different. In addition, interrelationships and interdependencies exist among the different stages of the process, and external and internal factors may act as influential sources of variability, causing major delays. Steps in the process may serve as bottlenecks, halting the review process. In addition, application errors and lack of required documentation on the part of the applicant can cause major delays in the review process.

To the best of our knowledge, none of the sources identified in this narrative literature review have modeled the current FDA review and approval process structure to address issues related to the robustness, reliability, and efficiency of its operations. Although several studies summarize the FDA approval process with clarity, in addition to bringing to light the problems and challenges faced by the regulatory agency, only a few attempts have been made to provide solutions for the problems and challenges identified. In addition, although several reform models have been discussed, these models lack the application of scientific methodologies and modeling techniques to understand FDA as a complex sociotechnical system. Furthermore, tools and methods to assess the efficacy of the models before implementation are largely absent. The implementation of any of these models would not only impact the FDA's authority but may also imply a change in the current structure of the regulatory agency. Changes proposed to the approval process must be accompanied by a dynamic model of the FDA regulatory structure under the new changes suggested.

Findings from this narrative review suggest an opportunity to employ MBSE approaches to provide a systems-oriented descriptive model of the FDA approval process for therapeutic biological products as a service network. Using this holistic approach will serve several investigative purposes: (1) identify influential sources of variability that cause major delays, including individual, team, and organizational decision making; (2) identify the human-system bottlenecks; (3) identify areas of opportunity for design-driven improvements; (4) study the effect of induced changes in the system; and (5) assess the robustness of the structure of the FDA approval process in terms of enforcement and information symmetry. However, adopting this approach may pose several challenges owing to the complexity of investigating internal FDA processes externally. Researchers external to the FDA have access to the official FDA guidelines such as the CDER 21st Century Review Process Desk Reference Guide, which provides a summarized overview of the drug approval process [29]. However, as the approval process of each drug may be different because of certain characteristics of the products and the manufacturing processes, modeling the approval process network based only on those guidelines would not reflect the intrinsic variability in the FDA operations. Consequently, any generated model may have poor external validity. Although this can be seen as a limitation, it also represents an opportunity for future collaboration between the FDA and the academia. The integration of external research and internal FDA efforts could facilitate the development of novel techniques and methodologies with practical applications that will benefit the regulator, industry, and patient population.

Limitations

Although this narrative review provides a broad, critical, and objective analysis of the current knowledge regarding the constraints and challenges in the FDA drug review and approval process and the approaches suggested in the literature to reform the FDA review and approval process, the search process and the number of sources included are not as comprehensive as it would be in a systematic literature review. The authors of this paper have expertise in the area of human and health care systems engineering; therefore, the review is limited to analyzing

the problem from a systems engineering perspective, shedding light on how systems engineering methods and approaches can be applied to obtain a better understanding of the FDA review and approval process, and to verifying, validating, and complementing the proposals of experts from other areas such as public policy.

Conclusions

The FDA review and approval of new drugs, including biotherapeutics, is a long and complex process. The process requires the involvement of multiple stakeholders internal and external to the FDA, in addition to the complexity of the interrelationships and interdependencies that exist among the different stages of the process. Literature in this area has identified challenges in the process related to cost, financial considerations, resource constraints, need for balance in terms of enforcement and information transparency, and the need for clarity in the FDA review and approval process and the requirements of the applicants. Although the FDA has implemented efforts to expedite the review and approval process, safety- and efficacy-related concerns about relaxing the FDA regulation over new drugs remain. Reform models and approaches have been proposed by experts in the area; however, these proposals lack the application of scientific methodologies and modeling techniques in understanding FDA as a complex sociotechnical system to obtain an unbiased assessment of the models' efficacy before implementation. MBSE approaches have been successfully used to model the FDA regulatory process for medical devices from the applicant's point of view. There is an area of opportunity to use MBSE approaches to model the review and approval process of new drugs from the regulatory agency's point of view to address issues related to the robustness, reliability, and efficiency of its operations.

Conflicts of Interest

None declared.

References

1. Van Norman GA. Drugs, Devices, and the FDA: Part 1: An Overview of Approval Processes for Drugs. JACC Basic Transl Sci 2016 Apr;1(3):170-179 [FREE Full text] [doi: 10.1016/j.jacbts.2016.03.002] [Medline: 30167510]
2. The history of FDA's fight for consumer protection and public health. U.S. Food and Drug Administration. 2018. URL: https://www.fda.gov/AboutFDA/History/default.htm [accessed 2021-02-12]
3. Williams R, Joffe M, Slonim A. Health options foreclosed: how the FDA denies Americans the benefits of medical research. The Mercatus Center at George Mason University. URL: https://www.mercatus.org/publications/regulation/health-options-foreclosed-how-fda-denies-americans-benefits-medical-research [accessed 2021-02-12]
4. Williams CT. Food and Drug Administration Drug Approval Process: A History and Overview. Nurs Clin North Am 2016 Mar;51(1):1-11. [doi: 10.1016/j.cnur.2015.10.007] [Medline: 26897420]
5. Fast track. U.S. Food and Drug Administration. 2018. URL: https://www.fda.gov/forpatients/approvals/fast/ucm405399.htm [accessed 2021-02-12]
6. Prescription drug user fee amendments. U.S. Food and Drug Administration. 2018. URL: https://www.fda.gov/forindustry/userfees/prescriptiondruguserfee/ [accessed 2021-02-12]
7. Conko G, Madden B. Free to Choose Medicine 2000;14(3):439-460.
8. International TA. The biopharmaceutical industry in the United States. n. URL: https://www.selectusa.gov/pharmaceutical-and-biotech-industries-united-states [accessed 2021-02-12]
9. Tsai W, Erickson S. Early-stage biotech companies: strategies for survival and growth. Biotechnol Healthc 2006 Jul;3(3):49-53 [FREE Full text] [Medline: 23424362]
10. Van Norman GA. Drugs and Devices: Comparison of European and U.S. Approval Processes. JACC Basic Transl Sci 2016 Aug;1(5):399-412 [FREE Full text] [doi: 10.1016/j.jacbts.2016.06.003] [Medline: 30167527]
11. Das SK, Almonor JB. A concurrent engineering approach for the development of medical devices. International Journal of Computer Integrated Manufacturing 2000 Jan;13(2):139-147. [doi: 10.1080/095119200129984]
12. Alexander K, Clarkson PJ. A validation model for the medical devices industry. Journal of Engineering Design 2002 Sep;13(3):197-204. [doi: 10.1080/09544820110108890]
13. Rathore AS, Winkle H. Quality by design for biopharmaceuticals. Nat Biotechnol 2009 Jan;27(1):26-34. [doi: 10.1038/nbt0109-26] [Medline: 19131992]
14. Gernaey KV, Gani R. A model-based systems approach to pharmaceutical product-process design and analysis. Chemical Engineering Science 2010 Nov;65(21):5757-5769. [doi: 10.1016/j.ces.2010.05.003]
15. Kourti T, Davis B. The business benefits of quality by design (QbD). Pharm Eng 2012;32(4):1-10.
16. Medina L, Dávila S, Kremer G, Wysk R. Design for FDA: a predictive model for the FDA's decision time for medical devices. In: ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference; 2013:4-7.
17. Medina LA, Kremer GEO, Wysk RA. Supporting medical device development: a standard product design process model. Journal of Engineering Design 2013 Feb;24(2):83-119. [doi: 10.1080/09544828.2012.676635]

18. Baylor N. Regulatory approvalcompliances for biotechnology products. Inhimasaki C. editor. Biotechnology Entrepreneurship. Waltham, MA: Academic Press; 2014:A-349.
19. Conner J, Wuchterl D, Lopez M, Minshall B, Prusti R, Boclair D, et al. The biomanufacturing of biotechnology products. Inhimasaki C. editor. Biotechnology Entrepreneurship. Waltham, MA: Academic Press; 2014:A-385.
20. Yu LX, Amidon G, Khan MA, Hoag SW, Polli J, Raju GK, et al. Understanding pharmaceutical quality by design. AAPS J 2014 Jul;16(4):771-783 [FREE Full text] [doi: 10.1208/s12248-014-9598-3] [Medline: 24854893]
21. Aksu B, Sezer A, Ye?en G, Kusçu L. QbD Implementation in biotechnological product development studies. In: Chen T, Chai S. editors. Special Topics in Drug Discovery. InTechOpen; 2016:133-156.
22. Briggeman J, Roberts E, Gulfo J. The Proper Role of the Fda for the 21st Century. SSRN Journal 2016 Feb 16. [doi: 10.2139/ssrn.3191411]
23. Kinch MS. 2015 in review: FDA approval of new drugs. Drug Discov Today 2016 Jul;21(7):1046-1050. [doi: 10.1016/j.drudis.2016.04.008] [Medline: 27109618]
24. Thierer AD, Wilt MP. The Need for FDA Reform: Four Models. SSRN Journal 2016 Sep 16. [doi: 10.2139/ssrn.2839384]
25. Djuris J, Djuric Z. Modeling in the quality by design environment: Regulatory requirements and recommendations for design space and control strategy appointment. Int J Pharm 2017 Dec 30;533(2):346-356. [doi: 10.1016/j.ijpharm.2017.05.070] [Medline: 28579542]
26. Horner M, Joshi S, Waghmare Y. Process modeling in the biopharmaceutical industry. In: Pandey P, Bharadwaj R. editors. Predictive Modeling of Pharmaceutical Unit Operations. Elsevier; 2017:383-425.
27. Joshi VS, Kumar V, Rathore AS. Optimization of ion exchange sigmoidal gradients using hybrid models: Implementation of quality by design in analytical method development. J Chromatogr A 2017 Mar 31;1491:145-152. [doi: 10.1016/j.chroma.2017.02.058] [Medline: 28259457]
28. Program for enhanced review transparency and communication for NME NDAs and original BLAs. U.S. Food and Drug Administration. 2017. URL: https://www.fda.gov/ForIndustry/UserFees/PrescriptionDrugUserFee/ucm327030.htm [accessed 2021-02-12]
29. CDER 21st century review process: desk reference guide. U.S. Food and Drug Administration. 2014. URL: https://www.fda.gov/media/78941/download [accessed 2021-02-12]
30. CDER Manual of Policies and Procedures | MAPP. U.S. Food and Drug Administration. 2018. URL: https://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ManualofPoliciesProcedures/default.htm
31. What is systems engineering? Z-Guides. UK Chapter International Council on Systems Engineering (INCOSE). 2009. URL: http://www.incoseonline.org.uk/Documents/zGuides/Z1_What_is_SE.pdf [accessed 2021-02-12]
32. Miller HI. To America's Health: A Proposal to Reform the Food and Drug Administration. Washington, DC: Hoover Press; 2000.
33. Klein D, Tabarrok A. Reform options. FDAReview. 2016. URL: http://www.fdareview.org/09_reform.php#3

Abbreviations

BLA: biologics license agreement
CDER: Center for Drug Evaluation and Research
EMA: European Medicines Agency
EU: European Union
FDA: Food and Drug Administration
MBSE: model-based systems engineering
PDUFA: Prescription Drug User Fee Act
QbD: quality by design
R&D: research and development

Edited by G Eysenbach; submitted 01.05.19; peer-reviewed by A Ayala, P Petaipimol; comments to author 16.09.19; revised version received 09.08.20; accepted 17.01.21; published 04.03.21.

Please cite as:
Bonet Olivencia S, Sasangohar F
Investigating the Food and Drug Administration Biotherapeutics Review and Approval Process: Narrative Review
JMIR Form Res 2021;5(3):e14563
URL: https://formative.jmir.org/2021/3/e14563
doi:10.2196/14563
PMID:33661119

Viewpoint

Mind Your Data: Privacy and Legal Matters in eHealth

Catharina M L Zegers[1,2,3], PhD; Annemieke Witteveen[1,4], PhD; Mieke H J Schulte[1,5], PhD; Julia F Henrich[1,6], PhD; Anouk Vermeij[1,7], PhD; Brigit Klever[1,8], PhD; Andre Dekker[1,3], Prof Dr

[1]Digital Society Health & Well-being, The Hague, Netherlands

[2]Institute of Data Science, Maastricht University, Maastricht, Netherlands

[3]Department of Radiation Oncology (MAASTRO), GROW - School for Oncology and Development Biology, Maastricht University Medical Centre, Maastricht, Netherlands

[4]Department of Biomedical Signals and Systems, TechMed Centre, University of Twente, Enschede, Netherlands

[5]Department of Clinical, Neuro and Developmental Psychology, Vrije Universiteit, Amsterdam, Netherlands

[6]Unit of Health, Medical and Neuropsychology, Institute of Psychology, Leiden University, Leiden, Netherlands

[7]Department of Cognitive Neuropsychology, Tilburg University, Tilburg, Netherlands

[8]University Medical Center, University of Groningen, Groningen, Netherlands

Corresponding Author:
Catharina M L Zegers, PhD
Institute of Data Science
Maastricht University
Paul-Henri Spaaklaan 1
Maastricht, 6229 ER
Netherlands
Phone: 31 43 388 4996
Email: karen.zegers@maastrichtuniversity.nl

Abstract

The health care sector can benefit considerably from developments in digital technology. Consequently, eHealth applications are rapidly increasing in number and sophistication. For successful development and implementation of eHealth, it is paramount to guarantee the privacy and safety of patients and their collected data. At the same time, anonymized data that are collected through eHealth could be used in the development of innovative and personalized diagnostic, prognostic, and treatment tools. To address the needs of researchers, health care providers, and eHealth developers for more information and practical tools to handle privacy and legal matters in eHealth, the Dutch national Digital Society Research Programme organized the “Mind Your Data: Privacy and Legal Matters in eHealth” conference. In this paper, we share the key take home messages from the conference based on the following five tradeoffs: (1) privacy versus independence, (2) informed consent versus convenience, (3) clinical research versus clinical routine data, (4) responsibility and standardization, and (5) privacy versus solidarity.

(JMIR Form Res 2021;5(3):e17456) doi:10.2196/17456

KEYWORDS

data; privacy; eHealth

The Digital Society Program

The Association of Universities in the Netherlands (VSNU) has brought together scientists from all 14 universities in the Netherlands to address the pressing questions raised by the emergence of a digital society. Exploring the responsible use of innovative digital technologies in the health care sector has high societal priority because of the potential to significantly improve health care and reduce health care costs. The Health & Well-Being program line of the VNSU Digital Society Research Programme aims to develop, evaluate, and implement integrated and personalized digital health care solutions, while addressing the societal challenges raised by the digitalization of health care.

Decelerating factors in the development and implementation of eHealth are a lack of knowledge, information, and practical tools with respect to handling privacy and legal matters. To discuss these factors, the Digital Society Health & Well-Being team organized a conference on September 26, 2019 titled “Mind Your Data: Privacy and Legal Matters in eHealth” with the aim to learn from each other’s approaches to tackle privacy and legal matters in the development of eHealth.

The conference hosted five speakers who were selected based on their unique backgrounds (law, eHealth, data science, philosophy, and mobile health [mHealth]), vision, and expertise on privacy and legal issues in eHealth. Marie-José Bonthuis is an external privacy lawyer who is connected to the Medical Biobank Lifelines and to the University Medical Center Groningen. Furthermore, she is an expert in the Health Research Infrastructure initiative (Health-RI) service desk for ethical, legal, and societal questions related to personalized medicine and next-generation sequencing. Dr. Bonthuis presented a talk titled "Overview of data protection principles in research: bringing practice and legislation together." Niels Chavannes is a professor of Public Health and Primary Care at Leiden University Medical Centre, a general practitioner, and the founder of the National eHealth living lab (NeLL). Professor Chavannes presented a talk titled "Clinical implementation of successful eHealth initiatives: ethical and legal issues." Andre Dekker is the professor of Clinical Data Science at Maastricht University, Maastricht University Medical Center+, and MAASTRO Clinic. Professor Dekker presented a talk titled "The personal health train: privacy preserving learning from health data." Peter Paul Verbeek is the professor of Philosophy of Technology and scientific codirector of DesignLab of the University of Twente. In addition, he is an honorary professor of Techno-Anthropology at Aalborg University and chair of the UNESCO World Commission for the Ethics of Science and Technology. Professor Verbeek shared his perspective in a talk titled "Privacy and beyond: inclusive digitalisation and the dynamics of privacy." Finally, Edward Watkins is the professor of Experimental and Applied Clinical Psychology at the University of Exeter. Professor Watkins presented a talk titled "ECoWeB – mental health app for young people data and governance issues."

More than 100 participants from a wide range of organizations (universities, medical centers, knowledge institutes, private parties, citizens, and government) attended the conference. Three independent authors noted down specific points that were expressed during the presentations, panel discussion, and eHealth forum. These notes were compared, sorted in categories, and juxtaposed in a way that the ethical challenges clearly emerged. Solutions provided by speakers were described; otherwise, clarification was provided by the authors of the paper. This resulted in our summary of the most prominent ethical-, technical-, and research-related issues in eHealth and their potential solutions.

Privacy Versus Independence

There is no straightforward answer for the best way to address privacy issues in eHealth. For each eHealth application, there should be a balance between individual privacy and potential individual or societal benefit. Data protection is all about contextual integrity; that is, using data responsibly within a specific context. Take for example the development of an mHealth approach to assess and enhance emotional competence for well-being in the young (ECoWeB project) [1]. In this project, young people expressed their reluctance to share passive sensor data and preferred to be in control of the data they would like to release. In this situation, an active approach to gather data (eg, by questionnaires) is most suitable. By contrast, the use of cameras, a more privacy-intrusive method, is more likely to be accepted in smart homes for people with dementia, since the benefit for the user is larger, as people could have the opportunity to stay at home for a longer time [2]. This is in agreement with Wilkowska et al [3], who showed that female patients and healthy adults insist on higher security and privacy standards, whereas people with ailing health consider privacy of lower importance. End users seem to be willing to tradeoff part of their privacy for the benefit of, for example, independence.

To gain insight into this tradeoff, user preferences, and needs, it is essential to include the end user in the design of eHealth at an early stage of development. This should provide an understanding as to what extent the user is willing to share data and for what purpose. During the conference, this was exemplified by the ECoWeB project [1] in which co-design with young people was critical. The data collection and recruitment are fully online processes, and the use of data is transparent and clear to the users. Previous research has shown that the uptake of eHealth and mHealth is only successful when they are built to fulfill a certain need of either patients or health care providers [4]. The early involvement of clinicians and patients will encourage adoption and maximize the positive impact of an intervention [5]. An example of a dedicated approach to develop and evaluate eHealth is the Centre for eHealth and Wellbeing Research roadmap developed by van Gemert et al [6], which focuses on user participation and process evaluation.

Informed Consent Versus Convenience

eHealth research generally includes an informed consent procedure for use and accessibility of data. This can potentially be done by digital authentication, including, for example, parental consent and age verification. However, during the conference's panel discussion, the issue about how elaborate digital informed consent should be arose. The panelists concluded that there should be a balance between simple, convenient, and easy to understand versus fully complete. This tradeoff is similar to a paper-based consent procedure. Nevertheless, there seems to be a striking difference between the requirements for informed consent of eHealth in comparison to commercial applications. An editorial published in *Nature* [7] also addressed this issue, stating that the consent for commercial mobile applications is often not more than a box to tick, with terms and conditions that are hardly ever read by the users.

In addition to informed consent, it is highly important to address the expectations of the eHealth app. This includes information on the procedure for incidental findings, such as whether or not the user wants to be actively informed or what can be expected with regard to automated messaging/triggering the health care provider for actions in the case of a monitoring app. Providing this additional information might limit possible overexpectations of users of the app.

Clinical Research Versus Clinical Routine Data

Data obtained after informed consent are only available from a small population of people that are registered for clinical research or use a specific eHealth app. By contrast, general registries collect data from a large number of patients, but the information is limited to demographics and a small selection of clinical variables. Another data source is clinical routine data, which contains the largest amount of clinically relevant information. One could think of a "patients like me" approach, where we can learn from existing data worldwide to find a similar patient. Unfortunately, clinical routine data are very hard to collect centrally because they are stored in individual local databases. One of the potential solutions is the use of distributed learning. The Personal Health Train is an example of this, where the data remain at the source (eg, the hospital) in Findable, Accessible, Interoperable, Reusable (FAIR) data stations, and the analysis method (eg, the algorithm) is transferred to the data. This method has been successfully implemented such as for predicting the 2-year survival of lung cancer patients using clinical data of 20,000 patients [8,9].

To allow for the secondary use of clinical personal data, data should be made nonidentifiable or anonymous [10], however, it is unclear when data are truly anonymous. Complete anonymity cannot be guaranteed if combinations of demographic data are given, and thus the question as to when data are truly anonymous remains. Many datasets that appeared to be anonymous have been released and individuals were reidentified [11]. Reidentification is generally performed by media or researchers, with the aim to show that shared data are unsafe, to publish new algorithms, and show weaknesses in the databases. An example was the reidentification of an individual from an adverse events database of Health Canada, and the media were able to reidentify a deceased woman based on a match of age, location, and date of death [12]. Algorithms such as k-anonymity can be used to describe the level of anonymity of datasets that plays a role in the definition of (non)identifiable personal data [13]. In addition, to determine the likelihood of an individual to be correctly reidentified, Rocher et al [11] proposed and validated a statistical model that was able to reidentify individuals even if the dataset was heavily incomplete. As such, if a dataset has been completely anonymized, it would be impossible to find the data of an individual who would like to withdraw consent to use their data. This is problematic if participants have been informed that they can withdraw their consent at any time [7]. An additional problem occurs in research scenarios that require data to be linked across different entities (eg, linking medical data from a hospital to socioeconomic data from Statistics Netherlands). These scenarios demand separate solutions that address privacy while still enabling subject-level linking based on common information [14].

Responsibility and Standardization

There are no straightforward answers to the questions of who is responsible for digital health apps, and how to guarantee maximal privacy and compliance with legislation. Together with multiple partners, the VSNU developed a Code of Conduct for research integrity in the Netherlands in 2018. The responsibilities have been defined at multiple levels, from the individual researcher to the boards of research institutions and the institutions as a whole [15]. Researchers can also use the General Data Protection Regulation (GDPR) carwash [16], a prototype of a practical flowchart that aims to guide researchers through the necessary steps to work with personal data in a GDPR-compliant manner.

Moreover, it is hard for a user to determine which app is qualitatively good. The availability of health apps is increasing rapidly. Pereira-Azevedo and Venderbos [4] estimated that over 300,000 medical apps were available in 2018. Medical App Checker developed by the Royal Dutch Medical Association was established as an initiative to evaluate mobile medical apps [17]. Medical App Checker consists of several checklists, including one for evaluating the protection and security of personal data. Additionally, an international norm (ISO standard) is in development for health and wellness apps. This standard could be used to certify apps that meet the norms for safety, reliability, and user-friendliness based on existing quality requirements and legislation. The latter is in flux as the new EU Medical Device Regulation, which takes effect on May 2021, will set more stringent demands on medical applications.

Finally, when moving from research toward the clinical implementation of eHealth, Dutch professional communities (medical specialists, medical physics, and clinical informatics) have expressed in their vision statements that they will take their responsibilities in the stimulation of the development and use of eHealth, and to assure its quality and safety.

Privacy and Solidarity

Technological innovations change our society rapidly and the interaction of humans with these digital innovations may also influence our perception of societal values such as privacy. The complex interactions of how innovations influence the ethical frameworks with which they are valued can be exemplified with a Google Glass study. In this study, a technological mediation approach was used to focus on the dynamics of the interaction between technologies and human values. Online discussions about Google Glass technology were investigated to evaluate how people articulate new meanings of the value of privacy [18].

Additionally, there are cultural differences in the way we value privacy, especially on a global scale. To account for this dynamism of values, value-sensitive and responsible design approaches should be adopted. There is also a movement toward solidarity and data donorship. Toward this end, a culturally sensitive balance should be sought between sharing ("give data and save lives") and protection (eg, potential threat of commercial exploitation) of data.

Conclusion

The information presented and discussed at the conference highlighted the many tradeoffs in eHealth with regard to privacy

and legal questions. To prevent potential decelerating factors in the development and implementation of eHealth, we need to be aware of these tradeoffs between (i) privacy and independence, (ii) informed consent and convenience, (iii) clinical research and clinical routine data, (iv) responsibility and standardization, and (v) privacy and solidarity. Furthermore, we need to make use of the available knowledge and tools on a national and international level, think carefully about the design of the application, and include end users at an early stage of development to reach the full potential of the eHealth technology. Clearly, there are risks associated with developments in eHealth, but rather than avoiding risks and stalling innovation, we should attempt to minimize risks while providing the greatest possible benefits to society.

Acknowledgments

We would like to thank the VSNU, the Digital Society program coordinators (Prof I Lagendijk, Prof M de Rijke, and Prof S Wyatt), our colleagues of the Health & Well-Being team and co-organizers of the conference (Prof A Brombacher, Prof A Evers, Prof E Feskens, Prof H Hermens, Prof L van Gemert-Pijnen, Prof N Maurits, Prof H Riper, Prof M Sitskoorn, Dr I Kalinauskaite, Dr J van Soest, Dr R Fijten, Dr R van der Vaart, Dr S van Dijk, Dr M Simons, Dr M Tabak, and Dr K Gehring), the speakers, including external privacy lawyer Marie-José Bonthuis, Professor of Public Health and Primary Care at Leiden University Medical Centre; general practitioner and the founder of the National eHealth living lab (NeLL) Niels Chavannes; Professor of Clinical Data Science in Maastricht Andre Dekker; Professor of Philosophy of Technology at the University of Twente and chairman of UNESCO's World Commission on the Ethics of Scientific Knowledge and Technology Peter Paul Verbeek; and Professor of Experimental and Applied Clinical Psychology at the University of Exeter Edward Watkins. We also thank all (forum) participants at the event.

Conflicts of Interest

None declared.

References

1. ECoWeB Project. URL: http://www.ecowebproject.eu/ [accessed 2019-10-17]
2. Wrede C, Jansen L, Havinga PJM, Croockewit JH, van Gemert-Pijnen JEWC. 2019 Presented at: 14th International Conference on Persuasive Technology; April 9, 2019; Limassol, Cyprus p. 43-45. [doi: 10.1007/978-3-030-17287-9]
3. Wilkowska W, Ziefle M. Privacy and data security in E-health: requirements from the user's perspective. Health Informatics J 2012 Sep 24;18(3):191-201 [FREE Full text] [doi: 10.1177/1460458212442933] [Medline: 23011814]
4. Pereira-Azevedo NM, Venderbos LDF. eHealth and mHealth in prostate cancer detection and active surveillance. Transl Androl Urol 2018 Feb;7(1):170-181. [doi: 10.21037/tau.2017.12.22] [Medline: 29594031]
5. Yang W, Shah L, Spaulding E, Wang J, Xun H, Weng D, et al. The role of a clinician amid the rise of mobile health technology. J Am Med Inform Assoc 2019 Nov 01;26(11):1385-1388 [FREE Full text] [doi: 10.1093/jamia/ocz131] [Medline: 31373364]
6. van Gemert-Pijnen JEWC, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
7. Editorial. Time to discuss consent in digital-data studies. Nature 2019 Aug;572(7767):5 [FREE Full text] [doi: 10.1038/d41586-019-02322-z] [Medline: 31367033]
8. Jochems A, Deist TM, van Soest J, Eble M, Bulens P, Coucke P, et al. Distributed learning: Developing a predictive model based on data from multiple hospitals without data leaving the hospital - A real life proof of concept. Radiother Oncol 2016 Dec;121(3):459-467 [FREE Full text] [doi: 10.1016/j.radonc.2016.10.002] [Medline: 28029405]
9. van Soest J, Sun C, Mussmann O, Puts M, van den Berg B, Malic A, et al. Using the personal health train for automated and privacy-preserving analytics on vertically partitioned data. Stud Health Technol Inform 2018;247:581-585. [Medline: 29678027]
10. Chevrier R, Foufi V, Gaudet-Blavignac C, Robert A, Lovis C. Use and understanding of anonymization and de-identification in the biomedical literature: scoping review. J Med Internet Res 2019 May 31;21(5):e13484 [FREE Full text] [doi: 10.2196/13484] [Medline: 31152528]
11. Rocher L, Hendrickx JM, de Montjoye YA. Estimating the success of re-identifications in incomplete datasets using generative models. Nat Commun 2019 Jul 23;10(1):3069. [doi: 10.1038/s41467-019-10933-3] [Medline: 31337762]
12. El Emam K, Rodgers S, Malin B. Anonymising and sharing individual patient data. BMJ 2015 Mar 20;350:h1139 [FREE Full text] [doi: 10.1136/bmj.h1139] [Medline: 25794882]
13. Samarati P, Sweeney L. Protecting privacy when disclosing information: k-anonymityits enforcement through generalization and suppression. Data Privacy Lab De-identification Project.: Carnegie Mellon University; 2018. URL: https://dataprivacylab.org/dataprivacy/projects/kanonymity/index3.html [accessed 2021-02-23]

14. Sun C, Ippel L, van Soest J, Wouters B, Malic A, Adekunle O, et al. A privacy-preserving infrastructure for analyzing personal health data in a vertically partitioned scenario. Stud Health Technol Inform 2019 Aug 21;264:373-377. [doi: 10.3233/SHTI190246] [Medline: 31437948]
15. Netherlands Code of Conduct for Research Integrity 2018 [FREE Full text] [doi: 10.17026/dans-2cj-nvwu]
16. Bonthuis MJ. 50 misverstanden over de Algemene Verordening Gegevensbescherming. Deventer, Netherlands: Wolters Kluwer; 2020.
17. App Checker. Royal Dutch Medical Association (KNMG). URL: https://www.knmg.nl/advies-richtlijnen/knmg-publicaties/medische-app-checker.htm [accessed 2019-11-18]
18. Kudina O, Verbeek P. Ethics from within: Google Glass, the Collingridge dilemma, and the mediated value of privacy. Sci Technol Hum Values 2018 Aug 21;44(2):291-314. [doi: 10.1177/0162243918793711]

Abbreviations

ECoWeB: emotional competence for well-being in the young
FAIR: Findable, Accessible, Interoperable, Reusable
GDPR: General Data Protection Regulation
Health-RI: Health Research Infrastructure Initiative
mHealth: mobile health
VSNU: Association of Universities in the Netherlands

Edited by G Eysenbach; submitted 13.12.19; peer-reviewed by N Seeman, DV S, M Abdelhamid; comments to author 30.03.20; revised version received 01.05.20; accepted 17.01.21; published 17.03.21.

Please cite as:
Zegers CML, Witteveen A, Schulte MHJ, Henrich JF, Vermeij A, Klever B, Dekker A
Mind Your Data: Privacy and Legal Matters in eHealth
JMIR Form Res 2021;5(3):e17456
URL: https://formative.jmir.org/2021/3/e17456
doi:10.2196/17456
PMID:33729163

Original Paper

Use of Telehealth for Domiciliary Follow-up After Hematopoietic Cell Transplantation During the COVID-19 Pandemic: Prospective Pilot Study

Alberto Mussetti[1,2], MD; Maria Queralt Salas[1,2], MD; Maria Condom[1,2], MD; Maite Antonio[2,3], MD; Cristian Ochoa[2,4,5], DPhil; Iulia Ivan[1,2], MD; David Jimenez Ruiz-De la Torre[1,2], MD; Gabriela Sanz Linares[1,2], MD; Belen Ansoleaga[1,2], MD; Beatriz Patiño-Gutierrez[1], RN; Laura Jimenez-Prat[1], RN; Rocio Parody[1,2], MD; Ana Sureda-Balari[1,2], MD

[1]Clinical Hematology Department, Institut Català d'Oncologia-Hospitalet, L'Hospitalet de Llobregat, Barcelona, Spain

[2]Institut d'Investigació Biomèdica de Bellvitge, Barcelona, Spain

[3]Oncohematogeriatrics Unit, Institut Català d'Oncologia-Hospitalet, Barcelona, Spain

[4]Psycho-Oncology Unit, Institut Català d'Oncologia, ICOnnecta't Health Program, Hospital Duran i Reynals, L'Hospitalet de Llobregat, Barcelona, Spain

[5]Clinical Psychology and Psychobiology Department, Faculty of Psychology, University of Barcelona, Barcelona, Spain

Corresponding Author:
Alberto Mussetti, MD
Clinical Hematology Department
Institut Català d'Oncologia-Hospitalet
Hospital Duran i Reynals
Avinguda de la Gran Via de l'Hospitalet, 199-203
L'Hospitalet de Llobregat, Barcelona, 08908
Spain
Phone: 34 622188475
Email: amussetti@iconcologia.net

Abstract

Background: Patients who have recently received a hematopoietic cell transplant (HCT) are at higher risk of acute complications in the first weeks after discharge, especially during the COVID-19 pandemic.

Objective: The aim of this study was to test the use of a telehealth platform for the follow-up of HCT patients during the first two weeks after discharge.

Methods: In total, 21 patients who received autologous or allogeneic HCT for hematological malignancies were screened from April 30, 2020, to July 15, 2020. The telehealth platform assisted in the daily collection of vital signs as well as physical and psychological symptoms for two weeks after hospital discharge. The required medical devices (oximeter and blood pressure monitor) were given to patients and a dedicated smartphone app was developed to collect this data. The data were reviewed daily through web-based software by a hematologist specializing in HCT.

Results: Only 12 of 21 patients were able to join and complete the study. Technological barriers were the most frequent limiting factor in this study. Among the 12 patients who completed the study, adherence to data reporting was high. The patients' experience of using such a system was considered good. In two cases, the system enabled the early recognition of acute complications.

Conclusions: This pilot study showed that telehealth systems can be applied in the early posttransplant setting, with evident advantages for physicians and patients for both medical and psychological aspects. Technological issues still represent a challenge for the applicability of such a system, especially for older adult patients. Easier-to-use technologies could help to expand the use of telehealth systems in this setting in the future.

(JMIR Form Res 2021;5(3):e26121) doi:10.2196/26121

KEYWORDS

SARS-CoV-2; COVID-19; hematology; hematopoietic cell transplantation; telemedicine; mortality; surveillance; monitoring; stem cell; transplant; app; medical device

Introduction

Patients receiving a hematopoietic cell transplant (HCT) have a high risk of developing severe acute toxicities in the early posttransplant period. An increased mortality risk has been observed in HCT patients who go on to develop COVID-19, with an estimated mortality rate between 20%-40% [1,2]. Although other hematological procedures can be postponed in case of emergency, there is an international agreement to not defer transplant procedures, especially for rapidly progressing diseases such as acute leukemias and aggressive lymphomas. Due to this, it is necessary to improve the domiciliary clinical monitoring of patients who received transplants to rapidly detect clinical deterioration. In addition, unnecessary in-person visits that might increase the risk of intrahospital contagion should be reduced. During the last few years, smart devices for assessing vital signs or physical activity have emerged as breakthrough innovations in the oncological setting [3-5]. Digital technologies allow clinicians to perform real-time monitoring of a patient's clinical status. In patients with COVID-19, the use of devices such as digital oximeters allow for early detection of clinical deterioration and a safer domiciliary follow-up. This is of paramount importance for patients who have received transplants, who have higher COVID-19–related mortality than the general population.

Methods

The aim of this study (SMARTCOVID19 study) is to report the feasibility of a real-time patient monitoring system through the use of a smartphone app and mobile health care devices. The institutional review board of Institut Catalá d'Oncologia - Hospitalet approved the study. Inclusion criteria were the following: those aged >18 years who had received autologous or allogeneic HCT while hospitalized and have a smartphone with an operating system able to run the SMARTCOVID19 app. Those who did not have adequate social support were excluded from the study.

Patient education regarding the use of the platform was provided by a hematologist at the time of enrollment, which was 1-2 days before hospital discharge.

Vital signs (heart rate, oxygen saturation, and arterial blood pressure) were collected daily using clinically validated oximeters (Onyx II, Nonin Inc) and a blood pressure monitor (iHealth Track, iHealth Labs), while temperature was measured using domiciliary thermometers. Patients were educated on how to measure their respiratory frequency. A checklist of clinical symptoms was completed daily (presence of cough, myalgia, headache, fatigue, dyspnea, emesis, odynophagia, rhinitis, conjunctivitis, and chest pain). An analog visual scale "thermometer" (0-10) used to detect potential cases of anxiety or depressive disorders was completed by patients daily. Scores of >6 resulted in an automatic referral to a psycho-oncologist through the platform, who would contact the patient and evaluate whether psychological support was needed via videoconference. A chat service was available for nonurgent communications. All data were reported to an online platform through a smartphone app ("Saludencasa," Fundación Trilema) compatible with Apple (iOS Version 9 or higher) and Android systems (Version 6 or higher). A hematologist with experience with HCTs reviewed all patient data daily. Programmed alarms were set in the event of any of the following situations: fever >38 C, oxygen saturation <92%, tachycardia >125 beats per minute, hypotension (systolic blood pressure <90 mm Hg, diastolic blood pressure <60 mm Hg), altered mental status, and persistent emesis or diarrhea (lasting more than 48 hours). In case of alarm activation, the hematologist contacted the patient by phone to evaluate the need for an in-person visit and determine the clinical management steps that were considered most appropriate. Outpatient monitoring started from the day of hospital discharge 2 days and continued for 14 days. The study accrual period was April 30, 2020, to July 15, 2020. Data were collected prospectively and all patients signed informed consent forms. Finally, two weeks after the end of the recruitment period, patients were contacted by phone and asked to answer a satisfaction questionnaire.

Results

During the study period, 16 of 21 patients who received transplants were successfully recruited into the study (76% feasibility). Reasons for patients not being enrolled were the following: language incompatibility (1 patient), no consent (1 patient), and no compatible smartphone (3 patients). Of the 16 enrolled patients, the median age was 50 years (range 22-70 years), 38% (n=6) were female, and 94% (n=15) had lymphatic diseases. In addition, 38% (n=6) of HCTs were autologous and 62% (n=19) were allogeneic.

Of the 16 enrolled patients, 4 were not able to use the app due to an inability to use smartphone apps in general. Of the remaining 12 patients, average adherence to reporting study data (median number of days reported across all patients during the planned 14-day study period) was as follows: 89% for temperature, 90% for oxygen saturation, 70% for respiratory frequency, 85% for cardiac frequency, 89% for blood pressure, 65% for symptoms reporting, and 71% for emotional distress.

Automatic alarms were activated only 3 times: twice for the presence of clinical symptoms and once for emotional distress. Only one patient spoke with the psycho-oncologist via videoconference. In total, 4 patients used the chat service to communicate with hospital personnel.

Despite the feasibility nature of the study, data collected with the digital system helped the clinician to recognize calcineurin inhibitor–induced arterial hypertension in one patient and acute cutaneous graft-versus-host disease (grade I) in another patient.

Only two patients in this cohort were readmitted within 14 days of discharge, both due to grade 4 odynophagia related to herpes simplex virus 1/2 reactivation.

The patients' responses to the survey questions about their experience with the telehealth system are reported in Table 1.

Table 1. Patients' responses to survey questions about their experience with the telehealth system (n=12).

Question	Mean score (scored from 1-5, where 1=disagree and 5=agree)
Overall satisfaction with the telehealth system	4.67
Did you feel safer at home with the use of the telehealth system?	4.67
Do you think that using such a device has improved your domiciliary follow-up?	4.67
Was the app easy to use?	4.50
When you feel well, would you be comfortable substituting in-person visits with telemedicine?	3.50

Discussion

Our prospective study showed that the use of mobile health care devices and smartphone apps for self-reported outcomes is feasible in the post-HCT setting. In a study conducted by Nawas et al [6], it was found that telehealth evaluations could be useful in the early peritransplant period, with a high satisfaction rate among patients. In our study, we found that patients felt safer when using the telehealth system. However, only a few patients would completely substitute an in-person visit with telehealth monitoring. This suggests that an in-person visit with a medical doctor is still the preferred follow-up modality for patients who have received transplants. Apart from vital signs monitoring and blood test results, which could be evaluated without the patient present, other aspects of medical care cannot be replaced by telehealth. Human contact and empathy are still largely needed during a visit. With respect to feasibility and adherence, two observations emerged from our study. The first is that only 57% of the potential patients used the telehealth system in the end. Technological barriers such as incompatible smartphones or inexperience in using smartphone apps were the main causes for study failure. This problem was more frequently observed in older adult patients living alone or without caregivers able to help them use the devices. A solution to this could be to automate data collection through the use of devices that automatically monitor and report this data. Newer wearable devices could allow for real-time monitoring of patients' heart rate, oxygen saturation, and physical activity in an automated manner. A second technological issue became apparent when evaluating the reporting adherence of the people who succeeded in using the platform. In this case, adherence to reporting clinical parameters such as respiratory frequency or symptoms was low. The use of currently available wearable devices could also resolve this adherence issue. In addition to technical issues, it is possible that psychological barriers could have contributed to reducing study adherence. Better patient education on how to use this system could also improve adherence, and should always be considered when applying digital medicine. Finally, the quality of telehealth monitoring should be complemented and improved by the collection of other clinically relevant parameters. For example, a virtual physical examination could be carried out through the use of high-quality video calls, digital stethoscopes, and weight scales [6,7]. In our study, the system worked well for detecting acute complications such as infections and dehydration. It is useful for monitoring the early posttransplant period, when acute complications are more frequent. For long-term follow-up, this system should be implemented with other technologies (eg, ones that allow visual communication between physicians and patients).

A study limitation was the small number of patients that completed the study. However, the study population was considered sufficient for a pilot study. In fact, the results obtained should be used to improve the design of the next study, which will recruit a larger number of patients. In addition, the uptake and use of newer technologies could be influenced by country or hospital resources. Such a system would not be feasible to implement within hospitals or health care systems that cannot afford such expenses. This applicability issue is common to the majority of studies using newer technologies.

In conclusion, telehealth monitoring could potentially improve patient follow-up in terms of both physical and psychological outcomes. This is especially true whenever an external cause (such as the COVID-19 pandemic) impedes in-person visits. Technological issues still represent a barrier to the wider application of telehealth monitoring systems in a medical setting and these issues should be considered for future studies.

Acknowledgments

We acknowledge all patients and their families, as well as all health care workers and scientists helping in the fight against the SARS-CoV-2 outbreak. We thank CERCA Programme/Generalitat de Catalunya for institutional support. In addition, we thank Tarsila Ferro Garcia for her help in finding the resources required to develop the study. We also thank Fundación Trilema for their free support in creating the platform for the study.

Authors' Contributions

AM contributed to conception and design and data analysis, and prepared the first draft of the manuscript. All authors contributed to collection and assembly of data, interpretation, and manuscript revision. All authors approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Shah G, DeWolf S, Lee Y, Tamari R, Dahi P, Lavery J, et al. Favorable outcomes of COVID-19 in recipients of hematopoietic cell transplantation. J Clin Invest 2020 Dec 01;130(12):6656-6667 [FREE Full text] [doi: 10.1172/JCI141777] [Medline: 32897885]
2. Passamonti F, Cattaneo C, Arcaini L, Bruna R, Cavo M, Merli F, ITA-HEMA-COV Investigators. Clinical characteristics and risk factors associated with COVID-19 severity in patients with haematological malignancies in Italy: a retrospective, multicentre, cohort study. Lancet Haematol 2020 Oct;7(10):e737-e745 [FREE Full text] [doi: 10.1016/S2352-3026(20)30251-9] [Medline: 32798473]
3. Patel M, Cooney MM, Caimi PF, Randall RE, Metheny LL, Mendiratta P, et al. Pilot study for the use of wearable devices for patients undergoing oral chemotherapy. JCO 2020 May 20;38(15_suppl):e14103-e14103. [doi: 10.1200/JCO.2020.38.15_suppl.e14103]
4. Liao Y, Thompson C, Peterson S, Mandrola J, Beg MS. The Future of Wearable Technologies and Remote Monitoring in Health Care. American Society of Clinical Oncology Educational Book 2019 May(39):115-121. [doi: 10.1200/edbk_238919]
5. Cox SM, Lane A, Volchenboum SL. Use of Wearable, Mobile, and Sensor Technology in Cancer Clinical Trials. JCO Clinical Cancer Informatics 2018 Dec(2):1-11. [doi: 10.1200/cci.17.00147]
6. Nawas MT, Landau HJ, Sauter CS, Featherstone CA, Kenny SA, Rodriguez ES, et al. Pilot Study of Telehealth Evaluations in Patients Undergoing Hematopoietic Cell Transplantation. Biol Blood Marrow Transplant 2020 Jun;26(6):e135-e137. [doi: 10.1016/j.bbmt.2020.02.004] [Medline: 32070723]
7. Lupo-Stanghellini MT, Messina C, Marktel S, Carrabba MG, Peccatori J, Corti C, et al. Following-up allogeneic transplantation recipients during the COVID-19 pandemic. Lancet Haematol 2020 Aug;7(8):e564-e565 [FREE Full text] [doi: 10.1016/S2352-3026(20)30176-9] [Medline: 32450053]

Abbreviations

HCT: hematopoietic cell transplant

Edited by G Eysenbach; submitted 28.11.20; peer-reviewed by J Mariotti, H Mehdizadeh, P Tripathi; comments to author 30.12.20; revised version received 31.12.20; accepted 16.01.21; published 12.03.21.

Please cite as:
Mussetti A, Salas MQ, Condom M, Antonio M, Ochoa C, Ivan I, Jimenez Ruiz-De la Torre D, Sanz Linares G, Ansoleaga B, Patiño-Gutierrez B, Jimenez-Prat L, Parody R, Sureda-Balari A
Use of Telehealth for Domiciliary Follow-up After Hematopoietic Cell Transplantation During the COVID-19 Pandemic: Prospective Pilot Study
JMIR Form Res 2021;5(3):e26121
URL: https://formative.jmir.org/2021/3/e26121
doi:10.2196/26121
PMID:33600351

Corrigenda and Addenda

Correction: 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:

Correction of: https://formative.jmir.org/2020/12/e24776/

(JMIR Form Res 2021;5(3):e28357) doi:10.2196/28357

In "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) three errors were noted.

In the originally published paper, the "greater than or equal to" symbol (≥) was missing in three places in tables due to an XML conversion error. The following corrections have been made:

In Table 1, under "Age group (years) (reference: 18-39)," "60" has been corrected to "≥60".

In Table 1, under "Household size (reference: 1 member)," "6" has been corrected to "≥6".

In Table 2, under "Age group (years) (reference: 18-39)," "60" has been corrected to "≥60".

The correction will appear in the online version of the paper on the JMIR Publications website on March 5, 2021, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 02.03.21; this is a non–peer-reviewed article;accepted 03.03.21; published 05.03.21.

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 Correction: Psychological Impacts of COVID-19 During the First Nationwide Lockdown in Vietnam: Web-Based, Cross-Sectional Survey Study
JMIR Form Res 2021;5(3):e28357
URL: https://formative.jmir.org/2021/3/e28357
doi:10.2196/28357
PMID:33667175

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org