JMIR Formative Research

Impact Factor (2022): 2.2
Volume 4 (2020), Issue 7 ISSN 2561-326X Editor in Chief: Gunther Eysenbach, MD, MPH, FACMI

Contents

Original Papers

Corrigenda and Addenda

Original Paper

Patient and Parent Perspectives on Improving Pediatric Asthma Self-Management Through a Mobile Health Intervention: Pilot Study

Michelle Nichols[1], PhD, RN; Sarah Miller[1], PhD, RN; Frank Treiber[1], PhD; Kenneth Ruggiero[1], PhD; Erin Dawley[2], BA; Ronald Teufel II[2], MD, MSCR

[1]College of Nursing, Medical University of South Carolina, Charleston, SC, United States

[2]College of Medicine, Medical University of South Carolina, Charleston, SC, United States

Corresponding Author:
Michelle Nichols, PhD, RN
College of Nursing
Medical University of South Carolina
99 Jonathan Lucas St
MSC 160
Charleston, SC, 29425
United States
Phone: 1 8437921802
Email: nicholmg@musc.edu

Abstract

Background: Asthma is a common chronic pediatric disease that can negatively impact children and families. Self-management strategies are challenging to adopt but critical for achieving positive outcomes. Mobile health technology may facilitate self-management of pediatric asthma, especially as adolescents mature and assume responsibility for their disease.

Objective: This study aimed to explore the perceptions of youths with high-risk asthma and their caregivers on the use of a smartphone app, Smartphone Asthma Management System, in the prevention and treatment of asthma symptoms, possible use of the app to improve self-management of asthma outside traditional clinical settings, and the impact of asthma on everyday life to identify potential needs for future intervention development.

Methods: Key informant interviews were completed with parent-child dyads post participation in an asthma management feasibility intervention study to explore the perceptions of users on a smartphone app designed to monitor symptoms and medication use and offer synchronous and asynchronous provider encounters. A thematic qualitative analysis was conducted inductively through emergent findings and deductively based on the self-determination theory (SDT), identifying 4 major themes.

Results: A total of 19 parent-child dyads completed the postintervention interviews. The major themes identified included autonomy, competence, relatedness, and the impact of asthma on life. The participants also shared their perceptions of the benefits and challenges associated with using the app and in the self-management of asthma. Both children and parents conveyed a preference for using technology to facilitate medication and disease management, and children demonstrated a strong willingness and ability to actively engage in their care.

Conclusions: Our study included support for the app and demonstrated the feasibility of enhancing the self-management of asthma by youth in the community. Participant feedback led to intervention refinement and app improvements, and the use of the SDT allowed insight into motivational drivers of behavioral change. The use of mobile apps among high-risk children with asthma and their parents shows promise in improving self-management, medication adherence, and disease awareness and in reducing overall disease morbidity.

(JMIR Form Res 2020;4(7):e15295) doi:10.2196/15295

KEYWORDS

asthma; mobile health; ecological momentary assessment; adolescents; medication adherence; self-management; mobile phone

Introduction

Background

Asthma is one of the most common chronic pediatric conditions in the United States [1,2]. Asthma prevalence estimates highlight the racial disparities faced by minority youth, with higher prevalence among non-Hispanic (NH) black children (13.4%) compared with their NH white counterparts (7.8%) [1]. Higher percentages of minority children with asthma also have suboptimal asthma control [3]. Poorly controlled asthma results in substantial burden to youth and their caregivers as it negatively impacts the quality of life, school attendance, and parental missed workdays and increases mortality among youth [4-6]. Identifying mechanisms that can reduce the burden of asthma and improve symptom control through education, skill development, and accurate and consistent use of inhaler medications may reduce morbidity and mortality among children with high-risk asthma [7].

Asthma control among adolescents can be affected by numerous factors, including knowledge about self-management of asthma, self-efficacy, medication adherence, exposure to smoke and other triggers, and access to medical care. As youth transition to older adolescence, they typically become increasingly responsible for self-management of asthma, making this transitional period an important time to act as life-long health behaviors are established [7]. The increasing use of technology across all demographic groups has led to the proliferation of mobile health (mHealth) apps for the management of asthma. Despite expanded availability and use, gaps remain in the development and implementation of mobile technology to enhance the self-management of asthma [8,9]. Most apps are not rigorously developed with patient engagement or feedback [8,9], which is an important consideration when designing and evaluating digital interventions to meet the needs of patients with increased morbidity and mortality, such as high-risk asthma [10]. Recent studies on the use of mobile apps to promote self-management among youth with diabetes highlight the benefits and continued need for patient-centered design processes at each phase of the intervention design process [11,12]. A systematic review by Ramsey et al [13] exploring publicly available mobile apps for asthma management noted a continued need for efficacy testing of apps and for intervention design to be grounded in evidence-based strategies. In addition, few apps are developed or implemented using established theories for health behavior or behavioral change, which can decrease the likelihood of an app influencing the skill development and ultimately, the desired health behavior change [14,15]. To address this gap, self-determination theory (SDT) was used to inform the development and refinement of the app and served as the overarching framework for the postintervention qualitative data collection and analysis within this study. The members of this research team have successfully used SDT and other behavioral theories to guide app development and refinement in previous studies [16,17].

Within SDT, relatedness (specifically shared experience) has been associated with positive health behaviors in youth with asthma. A study by Geryk et al [18] explored the preferences of parents and clinicians for an mHealth app using an iPod Touch to promote self-management of asthma in adolescents. The results showed receptivity toward facilitating a 2-way asthma care communication approach; however, their study did not include the perspectives of youth with asthma. They highlighted the benefit of integrating a theoretical framework within the implementation research designs to increase adolescent self-management, medication adherence, and asthma control and suggested including the ability to tailor interventions according to participant feedback [18]. Recognizing the need to increase engagement with youth who have asthma, Sleath et al [19] studied the use of an asthma question prompt list for youth with asthma, which they found encouraged youth randomized to the interventional arm to get actively involved during clinic visits [19]. The objective of this study was to engage youth and adolescents with high-risk asthma and their caregivers via key informant interviews (KIIs) to explore their perceptions of the use of a smartphone app in the prevention and treatment of asthma symptoms and the potential increase of asthma self-management outside clinical settings. Our team worked with software developers to design and implement a Smartphone Asthma Management System (SAMS) app in 2 phases for youth with high-risk asthma. Phase 1 focused on assessing the feasibility and acceptability of collecting real-time asthma information, and phase 2 included the intervention of asynchronous video assessment of inhaler use techniques and synchronous direct-to-consumer telehealth encounters. Both phases included a qualitative assessment of the impact of asthma on daily life to better understand the unique experiences of participants and to further identify the potential needs of youth and parents of children with asthma. By identifying and understanding the larger scope of the impact of asthma on quality of life, motivational strategies to promote behavioral change and activation can be leveraged for future interventional expansion [20]. Soliciting input through initial feasibility testing of youth and their parents on the acceptability, feasibility, and adaptability of mHealth apps during all phases of development is critical to understand the potential long-term effectiveness and viability of using this technology in managing high-risk pediatric asthma. The results from our feasibility study will inform intervention modifications for a future full-scale randomized clinical trial.

Theoretical Framework

Originating from the work of Deci and Ryan [21], the SDT explores factors that motivate individuals and allows insight into how to best tailor interventions that optimize action. Serving as a framework to understand and explore intrinsic and extrinsic motivation, the SDT can be used to explain factors that may drive health behaviors [22]. A central tenet of the SDT is the distinction of motivation, whether controlled or autonomous. Within a controlled motivational domain, individuals may feel a sense of obligation or duty to act in response to an employer or teacher's directive, whereas within autonomous motivation, the individual acts fully on his or her own volition [21].

The SDT comprises 3 main constructs: autonomy, competence, and relatedness [21]. Autonomy allows for self-direction and personal choice on one's actions and key decisions throughout life [21]. Competence includes self-efficacy and empowerment

with skills (including health behaviors) as well as mastery over things important to the individual. Relatedness is the sense of belonging or connectedness one has to others, along with the feeling of being cared for [21]. Through recognition of each of these constructs, clinicians can develop interventions conducive to enhance self-motivation and self-management, in turn leading to improved health outcomes [21].

Methods

Study Sample

Children and adolescents, aged between 8 and 17 years, with high-risk asthma and their parents were invited to participate in our iterative design feasibility study to ensure that the final product is end user guided. The inclusion criteria were as follows: (1) patient with high-risk asthma, defined as a patient having an emergency department or hospital visit for asthma in the past 12 months or the child's primary asthma provider in the outpatient setting having the opinion that the youth is at high risk for future asthma-related visits to the emergency department or hospital owing to the concern of noncompliance or poorly controlled asthma; (2) patient is prescribed a controller and rescue medication compatible with Bluetooth inhaler cap; (3) the caregiver or youth owns and is able to use a smartphone compatible with Bluetooth devices; (4) participant is English speaking; (5) patient with an identified primary care provider with at least one visit to the provider in the last year; and (6) patient with at least 1 identified caregiver present for enrollment, willingness, and availability to participate in study visits. This study was approved by the institutional review board of the Medical University of South Carolina. The participants were recruited from 2 children's hospitals in South Carolina and through provider referrals to participate in our iterative design process. The qualitative data reported in this study represent all enrollees in both phases of the iterative design process.

Mobile Health Platform

Iterative Design

The first phase of our iterative design included developing a process for real-time monitoring of the ecological momentary assessment (EMA) of asthma, asthma symptoms and other contextual factors such as emotional state, medication use via Bluetooth inhaler devices, and enrollee engagement via KII to inform the provider interactions in phase 2. The feasibility and acceptability data from phase 1 have been previously reported [23]. The second phase included the end user–guided provider interactions of asynchronous video capture of inhaler use techniques and synchronous direct-to-participant smartphone-based telehealth visits. Each study participant completed 1 video capture session and 1 telehealth visit during the intervention phase to assess feasibility, acceptability, and preferences toward these approaches, including their perceived benefits on the educational content from the personalized inhaler feedback received. Modifications to the data collection processes were made to include the collection of validated youth and parent surveys, including SDT metrics and removal of Bluetooth-enabled inhalers. The devices were removed for pragmatic reasons, including insufficient access to Bluetooth devices; initial participant and study team feedback regarding concerns over the accuracy of Bluetooth data; and their potential to add to the technical complexity for phase 2 enrollees, whom we wanted to focus on the newly developed provider interactions. The participants received SMS text message reminders to encourage completion of the EMA questions, video capture, and participation in the telehealth visit. Throughout the iterative design, our team considered the importance of developing competency (eg, the ability of youth to recognize symptoms and develop inhaler use techniques), autonomy (eg, telehealth visits performed by an educator trained in motivational interviewing), and relatedness (eg, encouraging but not requiring parental involvement in asthma management and developing remote provider encounters).

Real-Time Asthma Assessment

All enrollees had the SAMS app downloaded to their personal mobile phones and were requested to report EMAs daily. The EMA included 8 questions on their asthma symptoms, mood, or affect. Participants were asked to track asthma medication use, including controller and rescue inhalers, for a 2-month feasibility intervention period. EMAs reduce recall bias through real-time daily response reporting [24]; however, the frequency of response entries has the potential to increase participant burden [25]. All enrollees could view data on recorded medication use over their smartphones. Enrollees 1 to 14 recorded medication use with Bluetooth inhaler caps, and enrollees 15 to 22 recorded medications manually. Reminders were sent if real-time assessment data were missing for 3 days. Further details on the development, feasibility, and acceptability of the real-time assessments on SAMS are reported elsewhere [23].

End User–Guided Provider Interactions

Provider interactions, including asynchronous video capture of inhaler use technique and synchronous direct-to-participant smartphone-based telehealth visits, were discussed with all enrollees during their interviews. As the iterative design process progressed incorporating participant feedback, the protocols and programming were developed and refined for enrollees 14 to 22 to participate in these provider interactions during their 2-month feasibility trial. The enrollees were requested to participate at least once in both provider interactions over the 2-month study period. The asynchronous inhaler use technique was performed via the enrollee's smartphone by a tab within the SAMS app. Once activated, the enrollee would receive on-screen prompts regarding inhaler use technique, with a counter designed to be consistent with national guidelines. The video was encrypted and transferred to a secure server for review by an asthma educator who was a respiratory therapist, with feedback provided via SMS text, during the telehealth visit, or both. Synchronous telehealth visits were performed by enrollees using an embedded link to Vidyo, a telehealth video conferencing system, present in the SAMS app. An asthma educator trained in motivational interviewing performed these visits. Telehealth visits were requested at week 4 of the study and could be performed at a convenient time for enrollees.

Data Collection and Analysis

Following the completion of the 2-month interventional phase of the study, dyadic KIIs were conducted with the enrolled youth and caregivers using a semistructured interview guide (Table 1). The interview guide was structured to explore the constructs of the SDT and to identify the perceptions of participant' on the challenges and benefits associated with SAMS, including the potential burden associated with the use of EMA at home and in the community. Participants were also asked about the impact of asthma on their daily lives. Using an iterative interventional design approach, 2 waves of participants were enrolled. Data from phase 1 were collected, analyzed, and used to refine the intervention content and delivery for phase 2. Participants from both phases were invited to complete postintervention KIIs. The validated survey tools of youth and caregivers were collected in phase 2 and focused on asthma control, medication adherence, and SDT constructs (eg, self-efficacy, which is akin to competence) to better understand the study population and change of the measure from enrollment to the final visit.

KIIs were audio recorded, professionally transcribed, and reviewed for accuracy. The transcripts were then uploaded to NVivo 11.4.1 for data analysis. An experienced qualitative research methodologist analyzed the transcripts using a thematic content analysis approach, allowing identification of implicit and explicit themes found in the data [26-28]. A theoretical framework was used to situate findings, and researchers created a codebook containing operational definitions and examples, which enabled a systematic approach to data analysis and interpretation [26,29]. Data were deductively coded according to the SDT and the challenges and benefits of SAMS and then inductively through emergent codes [30]. Following line-by-line coding of raw text, the results were confirmed by members of the research team for confirmability of interpretation [31]. An audit trail of data coding and interpretation was maintained to increase the trustworthiness of the findings [32].

Table 1. Key informant interview guide.

Objective	Parent-directed interview prompts	Child-directed interview prompts	Dyad-directed interview prompts (parent and child)
Background information	• Tell me about his/her asthma medications before joining this study. What was working and what wasn't? How about now that he/she has been in the research study? What has changed?	• Can you tell me how old you are, what grade you are in, and whether you have your own cell phone or not? If not, whose cell phone did you use for this study? • About how many kids in your class at school would you say have their own cell phones? Do most kids your age have them, some, or only a few?	• N/A[a]
Impact of asthma on life	• What is it like for you having a child with asthma? • How does your child having asthma affect your work? • What about you, (parent/caregiver), how has his/her asthma affected your schedule?	• What is it like to live with asthma? • How does your asthma affect your ability to join in on activities with your friends and family?	• How often would you say you have missed school or work because of your child's asthma in the past? What about now?
Perceptions of SAMS[b]	• What about the system helped you manage your child's asthma? What about the system made it difficult?	• N/A	• Tell me about how comfortable you were using the SAMS system. • If you could describe the SAMS system to someone, what would you say about it? Why? • Now with this system, there is very little human interaction because you receive text messages or alerts versus actually speaking to someone. What are your thoughts on this? • Which questions that were sent through the messages were helpful? Which ones were annoying? How could we make the messages better? • What was the best thing about the SAMS system? What about the worst thing?
Recommendations	• Some parents have suggested receiving feedback on how well their children are doing with their inhalers. What are your thoughts on this? What types of things do you think would be helpful to know? • If you were to pick an ideal scenario for support on managing your child's asthma, what would you look for? Would you want mostly to use technology or combine it with telephone calls? If combined, how frequently would be the right amount of time for telephone calls?	• Some people have suggested we add a type of game in the system to make it more fun. What do you think about that? When you play electronic games, what makes them the most fun? What types of rewards do you like to receive on the games when you reach certain levels?	• Tell me what would have made this easier to use or what you would change about it if you could.

[a]N/A: not applicable.

[b]SAMS: Smartphone Asthma Management System.

Results

Sample Characteristics

A total of 19 parent-child dyads (overall mean age of the children was 10.89 years, SD 2.73) participated in the postintervention interviews from the original 22 dyads who completed the study intervention. Furthermore, 3 additional children completed the study procedures (phase 1), which included a 60-day data capture, but were lost to follow-up and unable to be reached for inclusion in the postintervention interview; thus, they were not included in the qualitative analysis. Overall, 7 of the 19 who completed the study intervention were female. Participants (N=22) reported insurance coverage as Medicaid (18/22, 82%), self-paid (3/22, 14%), and private insurance (1/22, 5%). Most (n=16) of the participants were African American (84%), with 2 children identifying as mixed race and 1 child as white. Phase 1 included 11 participants, mean age of children was 10.9 years (SD 2.73),

followed by an additional 8 participants in phase 2, mean age of children was 12 years (SD 3.02). Phase-specific demographic data are presented in Table 2, and the child and parent survey tools used in phase 2 are highlighted in Table 3.

Table 2. Study population demographics by design phase.

Demographics	Phase 1 (n=14), n (%)	Phase 2 (n=8), n (%)
Age (years)		
8-10	9 (64)	4 (50)
11-13	3 (21)	2 (25)
14-16	2 (14)	2 (25)
Gender		
Male	5 (36)	6 (75)
Female	9 (64)	2 (25)
Race		
White	0 (0)	6 (75)
Black	13 (93)	1 (13)
Other	1 (7)	1 (13)
Ethnicity		
Hispanic	0 (0)	0 (0)
Non-Hispanic	14 (100)	8 (100)
Insurance		
Medicaid	11 (79)	7 (88)
Private	1 (7)	0 (0)
Self-paid	2 (14)	1 (13)
Clinical characteristics		
Asthma severity		
Intermittent	0 (0)	0 (0)
Mild persistent	1 (7)	0 (0)
Moderate persistent	8 (57)	7 (88)
Severe persistent	3 (21)	0 (0)
Asthma control		
Well controlled	0 (0)	0 (0)
Not well controlled	12 (86)	7 (88)
Very poorly controlled	1 (7)	1 (13)

Table 3. Child and parent survey tools used in development phase 2 (n=8).

Survey tool	Enrollment, mean (SD)	Final 2-month visit, mean (SD)	*P* value
Asthma control test[a]	17.4 (3.4)	21.6 (2.3)	.01
Child asthma belief scale	3.6 (0.8)	4.3 (0.6)	.06
Reported medication adherence survey	3.8 (0.7)	4.2 (0.7)	.30
Parental asthma self-efficacy scale	4.2 (0.5)	4.4 (0.3)	.19
Parental belief in treatment efficacy	4.7 (0.2)	4.5 (0.6)	.39

[a]*P* value less than .05 enrollment compared with final visit controlling for patient as a random effect.

Impact of Asthma on Life

Parents and children alike voiced concerns over the impact of asthma on their daily lives. Parental concerns ranged from financial burdens, loss of time from work, challenges in managing busy schedules, and an overwhelming sense of fear and uncertainty related to asthma exacerbations and what that means for their children's well-being. Children explained how asthma affects their ability to engage in activities, causes them to miss school days, and can even lead to embarrassment and potential stigma from being considered *different* from other children. Exemplar quotes are presented in Textbox 1.

Textbox 1. Exemplar quotes on the impact of asthma on daily life.

Activities

- "I can't always run as long as I want." [Child no. 6, phase 2]
- "There are certain things I can't do like strenuous exercises or certain sports." [Child no. 8, phase 1]
- "If it's something really active, I can't do it as much as everybody else does...I tried to join the dance team, but some of the stuff they do—they do a lot of stuff—so I'll get tired faster than everybody else." [Child no. 12, phase 1]
- "Like when you play basketball, you can only play for a certain amount of time." [Child no. 4, phase 2]
- "Actually, when I am running the bases my chest feels tight." [Child no. 9, phase 1]
- "It's kind of hard because I can't do certain stuff anymore that I used to do." [Child no. 7, phase 2]

School

- "Missed maybe four days within the last year or so." [Parent no. 12, phase 1]
- "Sometimes at PE [physical education], when we're like running around and I was in the corridor, my lungs started to hurt." [Child no. 2, phase 1]
- "Missed only one week." [Child no. 9, phase 1]
- "I think it was two or three days he was out." [Parent no. 11, phase 1]
- "I missed a little bit of school a couple of weeks ago because I had a flare-up." [Child no. 8, phase 2]
- "A month...maybe five days." [Parent no. 6, phase 2]

Family life

- "Asthma affects our life daily because it just impacts everything. When he has an asthma attack, I miss work, he misses school, my mom is responsible for getting my daughters to school because he is in the hospital. That is when it is not controlled." [Parent no. 1, phase 1]
- "I have two smaller children and that [hospitalizations] kind of pulls you away from the two that you have at home or stuff you have to do for the younger kids or other kids." [Parent no. 12, phase 1]
- "We don't turn the heat on overnight because it makes him cough. I was like 'girls, bundle up because either he can breathe and you're cold or he can't breathe'." [Parent no. 2, phase 2]

Work

- "If the school calls, I got to leave work and go pick her up." [Parent no. 4, phase 2]
- "While he was in the hospital [Dad] missed the first day and a half [of work] wanting to make sure everything was okay with him." [Parent no. 9, phase 1]
- "Sometimes when he is in distress and all and has to go to the hospital because of his asthma, we both [parents] sometimes miss work. I have to do what I have to do for my child." [Parent no. 8, phase 2]
- "At work, I'm a supervisor, I'm with patients 'cause I'm in the health care field so I can't just up and leave the patients, but I also have a child that I have to worry about. That comes first. I have to take a lot of days off." [Parent no. 6, phase 2]

Autonomy

As children age and mature, there is a natural tendency for them to want to self-manage their health and health care needs, which supports their intrinsic motivation to be more independent [33]. This desire is often juxtaposed with parental needs to ensure optimal health outcomes for their children and the difficulties they face in relinquishing control of day-to-day disease management. Parental concerns can stem from fear associated with asthma exacerbations and the safety and well-being of their children. One parent (phase 2) shared that she just recently allowed her 8-year-old son to start sleeping in his own room because she was afraid that he may struggle with breathing during the night and she might not be able to respond in a timely manner. The experience of watching their child go through an asthma crisis brought about significant fear for many parents, reducing their comfort level in allowing their child to self-manage their disease:

> *Just the thought of not being able to breathe—it's like drowning and you are not in water or something. It's just terrifying sound to me.* [Parent no. 2, phase 2]

> *Sometimes it can be a little scary, especially if she is having difficulty breathing and not verbalizing it...you have to look out for certain signs like if they are gasping for air or like if their breathing pattern has changed or increased or decreased.* [Parent no. 1, phase 2]

> *It's very scary...when [he] first came home from the hospital, I was nervous. I watched him all the time sleeping and if I hear a cough, I get so nervous...I have to monitor and make sure that he's taking his medicine and getting this right.* [Parent no. 11, phase 1]

> *When he had to be flown to the hospital for his asthma flare-up, that was pretty scary.* [Parent no. 8, phase 2]

As part of the study intervention, participants received reminder notifications on medication administration and survey completion. Participants conveyed strong support for the reminder alerts in facilitating autonomy among children and in transitioning from parent-managed to shared responsibility. Older adolescents with access to their own mobile devices expressed that they were able to take on a more active role and expanded autonomy through self-management of medication administration and symptom management. One parent shared how her 15-year-old daughter "knows more about what she is dealing with" now that she is older and that "she is self-managing" (Parent no. 12, phase 1). Another parent indicated that her 10-year-old daughter (phase 2), who was using the parent's phone, would take the initiative, ask for the phone, and would remind the mother that she had to record her medications and answer the questions. Even children as young as 8 years old demonstrated the ability and desire to self-manage their disease. One mother conveyed how her son would state "this is private" when related to the responses he was entering for the questions, as they pertained to his personal health (Parent no. 2, phase 2) and he wanted to take responsibility for the entries:

> *It helps me. Like it reminds me that I need, in case I forget to take my medicine or anything, it reminds me to stop whatever I'm doing and go take my medicine.* [Child no. 1, phase 2]

> *Now that she is older, I let her do it...it teaches them responsibility...That gives them a little feel good. I'm doing something. They're more in control like that...it makes them feel good about themselves, too.* [Parent no. 8, phase 1]

> *I try to allow him to do it more than myself, so when he wasn't around, then I didn't do it 'cause I wanted him to be responsible to do it.* [Parent no. 6, phase 2]

> *I prefer the app because you could actually do it by yourself.* [Child no. 8, phase 1]

Competence

Knowledge, understanding, and prior experience with asthma and asthma care management factored into participant responses:

> *Well, it's been a little bit of a learning curve for me just because I've never had to deal with anyone with asthma before.* [Parent no. 3, phase 2]

> *We may think that they're actually doing it correct, but if we're not taught, then we don't know how to tell them to correct stuff that they're doing.* [Parent no. 1, phase 2]

Participants expressed a strong comfort level with the SAMS system following the initial demonstration, with several sharing similar feedback that it was quite easy to use. All participating children shared the sentiment that it was easy to use and highlighted the use of technology as a normative component of their daily lives:

> *I'm a kid, so the phone is my life.* [Child no. 12, phase 1]

> *I think that it's good because like I like computers and stuff and I like to be on the phone a lot.* [Child no. 8, phase 1]

The features that contributed to the increased competence in managing asthmatic symptoms included educational content and the ability to tailor feedback following the video capture of the use of inhaler. For the video capture of the inhaler use technique, participants were prompted with a series of on-screen prescripted prompts to promote proper inhaler use techniques. Tailored feedback was provided by the respiratory therapist following a review of the video capture feature within the app. The tailored feedback included tips on proper inhalation techniques such as mouth placement on the inhaler, holding one's breath after inhalation, and shaking the inhaler before use. This was particularly true for 1 newly diagnosed child, as he initially ran out of medication early because of improper technique. Others describe the benefit of capturing inhaler use for technique refinement and parental ability to ensure proper medication administration adherence:

> *He didn't know how to really use it. He's new to this, so he just pumping away and thinking he-he pumped the first time, it's not going in, so we do it again and keep on trying till get it. I think that's why he ran out so fast.* [Parent no. 11, phase 1]

> *I was able to use the help option there within the app, and it gave me some directives. I was able to figure it out pretty quickly.* [Parent no. 9, phase 1]

> *I think that's good because you can learn how to take your medicine better so that it'll work better.* [Child no. 8, phase 1]

> *The tool to me is super easy for somebody that's new to having a child with asthma. It really did help me...it's walking you through—this is exactly how you are to do it.* [Parent no. 3, phase 2]

The participating children shared a sense of pride and accomplishment in self-managing their care, an essential factor in enhanced competence, which can facilitate increased engagement:

> *I felt really comfortable. It was an easy app to learn and I've been using it.* [Child no. 3, phase 2]

> *[I was] pretty comfortable 'cause it's like I knew I took my medicine and...it helps you remember to take your meds...I didn't wanna get sick.* [Child no. 6, phase 2]

Relatedness

Children reported the importance of relatedness, underscoring the importance of peer relationships and social acceptance during childhood and adolescent development. They further expressed how asthma has the potential to increase social isolation and limit the ability to engage with peers at the same level or for the same duration (Textbox 1). Several articulated concerns over being perceived as *different*, a strong desire to be considered *normal*, and as preferring to maintain privacy regarding the awareness of their disease with others. An example of this was highlighted by an 8-year-old boy who became very upset when using the inhaler and answering questions at school as his classmates "didn't even know I have asthma" and once realized, children noticed and began talking:

> *But once I did it at school because she came and had lunch with me and then people were looking at it and I was mad.* [Child no. 2, phase 2]

Feeling supported and cared for by others were strong attributes of relatedness throughout the intervention and, in general, participants appreciated concern over health and well-being conveyed by others while participating in the study. One mother commented that it was good "to get to interact with the person on the other side" when referencing the video interaction and personal connectedness using mobile technology (Parent no. 4, phase 2). Others expressed appreciation in knowing that providers were actively involved and concerned with their child's well-being and care management:

> *With the continual feedback or reminders, it actually made me aware that somebody was, in fact, watching.* [Parent no. 1, phase 2]

> *[The intervention] reassured me that there is genuine concern about how my son is doing and how he is recovering and whatever other treatments he may need.* [Parent no. 9, phase 1]

> *A lot of people in my class, when they heard that I was in the hospital because of my asthma, they said that they were, like, worried about me...it made me feel good.* [Child no. 3, phase 2]

Challenges of Smartphone Asthma Management System

Although there was unanimous support for SAMS among all interviewees, some expressed challenges. The time required for video uploads was problematic for some, especially when waiting to upload one video before administering a second dose. One mother described how it was somewhat awkward for her son to hold the phone while capturing the video and simultaneously self-administering medication. She also shared that her smartphone would "time-out" and necessitate touching the screen, which would interrupt the video.

Bluetooth connection and proximity to the smartphone yielded challenges for families when the child did not have his or her own phone. This was exemplified further when children would be away from home for a period and data capture would be delayed. The exemplar quotes are presented in Table 4.

Table 4. Exemplar quotes on the benefits and challenges of Smartphone Asthma Management System.

Role	Benefits	Challenges
Parent	• "Half the times I have so much going on, I forget. Whereas with the app, it it's not taken, it pops up on the phone...and this way, it makes it much easier as far as controlling how she takes her medicine." [Parent no. 12, phase 1] • "It was helping me to remind her to take her medication...'cause the alarm goes off so for her to take her medicine and for her to answer the questions, so it's like a double reminder in the evening when we are so busy." [Parent no. 7, phase 1] • "I think the text messages and alerts, not actually having to sit on the phone with a human and the hold times, the wait times—technology has taken over. I think it's just so much more simpler." [Parent no. 12, phase 1] • "I don't mind talking to people, but I don't want you to call me every day. You know, text me every day is fine. You send me a message every day it's fine, but I'm more of a texter in general than I am a talker." [Parent no. 2, phase 2] • "That was helpful because...she was kind of doing it wrong. So, when we talked to her, she kind of told us, you know, certain ways you're supposed to do it and count. That was real helpful." [Parent no. 4, phase 2] • "I would ask her...'oh, what type of feedback did you get in the video?' and she expressed that she was told to sit up straight and I think something to do with the spacer." [Parent no. 1, phase 2]	• "So, you are sitting there for like four or five minutes because you take your second puff to get, you know, to film that...it was just a little difficult." [Parent no. 3, phase 2] • "I would have to site my phone in a neutral location or she would have to get closer to the room in order for it to pick up. As far as I was concerned, that's the only flaw. Just like I said, she had went out for the weekend and wasn't home. So, the phone doesn't pick up if it's not her phone and it's not where she is. That for me is the only issue I have. Other than that, it was good." [Parent no. 12, phase 1]
Child	• "It helps you remember to take your meds." [Child no. 6, phase 2]	N/A[a]

[a]N/A: not applicable.

Benefits of Smartphone Asthma Management System

A recurring theme throughout the interviews was the benefit of reminders on medication adherence. Although intentions were strong to follow medical advice, the realities of balancing demanding schedules and daily responsibilities pose challenges with actual adherence. Numerous quotes supported the advantage of SAMS in facilitating improved medication adherence (Table 4).

Convenience was another benefit shared by the participants. Parents articulated how they often struggle to manage work and family responsibilities and how finding time to have a conversation during the work day with a provider can be challenging; however, they felt that SAMS facilitated communication in a convenient manner for all parties. Moreover, parents felt that SAMS improved day-to-day care management and reduced unplanned provider visits, indicative of fewer asthmatic exacerbations for the children and the ability for parents to reduce time lost from work (Table 4).

The use of technology to increase medication adherence and accuracy of medication administration was supported by children and parents. Technology was also embraced as a means to facilitate more immediate communication and assessment and as a mechanism for timely and targeted intervention. In addition, the use of technology, whether via messaging or through video interaction, was strongly preferred by most participants as the primary means of routine communication and interaction. None of the participants had concerns over any potential burden associated with the frequency of EMA. Although text messaging and video conferencing were valued means of interaction, none of the participants wanted to completely replace traditional face-to-face encounters.

Discussion

Principal Findings

The results of this study indicate that children with high-risk asthma and their parents prefer to use technology to facilitate medication administration and disease management. Although parents still want the ability to personally interact with clinicians on an as-needed basis, children and parents articulated an inclination toward the use of technology to manage routine maintenance and monitoring. Interestingly, parents did not have safety or privacy concerns with the use of video recordings of their children using inhalers. This may be related to the ready availability of access to clinicians and support provided with minimal impact on work and school schedules, which likely indicates that any potential privacy concerns were outweighed by the benefits of the intervention. Further exploration of this is warranted among more socioeconomically diverse populations.

In addition, the use of Bluetooth-enhanced mHealth apps may be a viable mechanism for optimizing youth-parent-clinician engagement in the management of children with high-risk asthma. As technology has become a normative part of life for children and parents in the United States, capitalizing on innovative strategies to improve self-management of chronic disease among high-risk populations and reduce disease burden is a natural step forward. Although our study focused specifically on SAMS, the knowledge gained from this study may be transferable to other apps and mobile platforms.

Our research was strengthened through the engagement of children and adolescents with asthma and their parents. We demonstrated that children, as the end users of mobile apps, are willing and able to participate in the self-management of their asthma and welcomed the use of technology to facilitate disease and symptom management, allowing for increased autonomy and responsibility of care. There were no concerns over the frequency of responding with EMA, although further study is needed to identify if this would remain consistent with longer intervention periods. Understanding the preferences of children and parents is crucial in translational research efforts, allowing providers to best understand what works and may not work for patients outside clinical settings and living in the community, and it also has relevance for interventional research design and clinical practice.

Although conducting dyadic interviews with parents and children may have influenced the children's responses, it was important to understand the preferences and experiences of both parents and children following the use of the SAMS app. All children actively contributed during the KIIs, and interview questions were not sensitive in nature, thus minimizing the likelihood of child participants withholding thoughts because of parental presence. Including children and parental input in an iterative mobile app design process can enhance the likelihood of adaptation and adherence in clinical trials and allows clinicians to more broadly understand the preferences and practices of managing chronic pediatric illnesses, which can be translated into practice recommendations to improve clinical outcomes.

Our study, through the assessment of the impact of asthma on life, builds upon recommendations to include quality of life assessments as an integral component of asthma interventions designed to enhance adherence [34]. Understanding how asthma affects day-to-day life for children and families provides insight into factors that may serve as motivational drivers for this population by tailoring mHealth interventions by using personalized priorities to promote behavior change. Incorporating input from end users enables a process of co-designing the app and intervention between clinical and academic researchers and children with asthma and their parents. In addition, the use of the SDT as a theoretical framework proved effective, allowing insight into motivational drivers of behavioral change, knowledge levels and skill acquisition, and factors related to one's need for interconnectedness in personal relationships. Both mobile apps and the use of SDT can facilitate the understanding and management of children with high-risk asthma residing in community settings. Furthermore, this study helps fill an identified gap in mobile app development as many have not used models of health behavior change in app design and content development.

Strengths and Limitations

This study has several limitations. Our study design was informed by a wide range of causal models; however, a more structured approach to causal factors and exploration of

mechanisms to support self-management through the app is recommended for future studies. In addition, in this iterative feasibility pilot study, we chose not to script out the roles within the dyads and focused on the natural process of the dialog between the parent and child. It may be beneficial to explore relational interactions when considering the interplay of interpersonal relationships and self-management. In future studies, we intend to include measures of motivation and other SDT constructs. Although our study included mechanisms to interact with providers and captured EMA data, the app currently lacks an automated interface with electronic medical records, a noted gap among apps designed to manage asthma [13]. Study recruitment via provider referrals may have enabled providers to refer patients with a higher likelihood of interest in mHealth or those who providers felt needed better asthma management. Furthermore, another limitation was the discontinuation of objective adherence monitoring via Bluetooth. Although this report focused on the postintervention qualitative perspectives of participating children and parents, it is worthwhile to note the limitation of not including passive adherence monitoring in both phases of the actual pilot intervention.

Acknowledgments

This study received funding support from the South Carolina Clinical and Translational Research Institute (UL 1 TR000062) and the Health Sciences South Carolina.

Conflicts of Interest

None declared.

References

1. Centers for Disease Control and Prevention. 2017. Most Recent Asthma Data URL: https://www.cdc.gov/asthma/most_recent_data.htm [accessed 2018-02-22]
2. Roemer M. The Agency for Healthcare Research and Quality. 2008. Health Care Expenditures for the Five Most Common Children's Conditions, 2008: Estimates for US Civilian Noninstitutionalized Children, Ages 0-17 URL: https://meps.ahrq.gov/data_files/publications/st349/stat349.shtml [accessed 2018-02-22]
3. Halterman JS, Auinger P, Conn KM, Lynch K, Yoos HL, Szilagyi PG. Inadequate therapy and poor symptom control among children with asthma: findings from a multistate sample. Ambul Pediatr 2007;7(2):153-159. [doi: 10.1016/j.ambp.2006.11.007] [Medline: 17368410]
4. Cui W, Zack MM, Zahran HS. Health-related quality of life and asthma among United States adolescents. J Pediatr 2015 Feb;166(2):358-364 [FREE Full text] [doi: 10.1016/j.jpeds.2014.10.005] [Medline: 25454936]
5. Akinbami LJ, Moorman JE, Liu X. Asthma prevalence, health care use, and mortality: United States, 2005-2009. Natl Health Stat Report 2011 Jan 12(32):1-14 [FREE Full text] [Medline: 21355352]
6. Dean BB, Calimlim BM, Kindermann SL, Khandker RK, Tinkelman D. The impact of uncontrolled asthma on absenteeism and health-related quality of life. J Asthma 2009 Nov;46(9):861-866. [doi: 10.3109/02770900903184237] [Medline: 19905909]
7. Orrell-Valente JK, Jarlsberg LG, Hill LG, Cabana MD. At what age do children start taking daily asthma medicines on their own? Pediatrics 2008 Dec;122(6):e1186-e1192. [doi: 10.1542/peds.2008-0292] [Medline: 19047221]
8. Huckvale K, Car M, Morrison C, Car J. Apps for asthma self-management: a systematic assessment of content and tools. BMC Med 2012 Nov 22;10:144 [FREE Full text] [doi: 10.1186/1741-7015-10-144] [Medline: 23171675]
9. Huckvale K, Morrison C, Ouyang J, Ghaghda A, Car J. The evolution of mobile apps for asthma: an updated systematic assessment of content and tools. BMC Med 2015 Mar 23;13:58 [FREE Full text] [doi: 10.1186/s12916-015-0303-x] [Medline: 25857569]
10. Alquran A, Lambert KA, Farouque A, Holland A, Davies J, Lampugnani ER, et al. Smartphone applications for encouraging asthma self-management in adolescents: a systematic review. Int J Environ Res Public Health 2018 Oct 29;15(11):E2403 [FREE Full text] [doi: 10.3390/ijerph15112403] [Medline: 30380692]
11. Holtz BE, Murray KM, Hershey DD, Dunneback JK, Cotten SR, Holmstrom AJ, et al. Developing a patient-centered mhealth app: a tool for adolescents with type 1 diabetes and their parents. JMIR Mhealth Uhealth 2017 Apr 19;5(4):e53 [FREE Full text] [doi: 10.2196/mhealth.6654] [Medline: 28428167]
12. Castensøe-Seidenfaden P, Husted GR, Teilmann G, Hommel E, Olsen BS, Kensing F. Designing a self-management app for young people with type 1 diabetes: methodological challenges, experiences, and recommendations. JMIR Mhealth Uhealth 2017 Oct 23;5(10):e124 [FREE Full text] [doi: 10.2196/mhealth.8137] [Medline: 29061552]
13. Ramsey RR, Caromody JK, Voorhees SE, Warning A, Cushing CC, Guilbert TW, et al. A systematic evaluation of asthma management apps examining behavior change techniques. J Allergy Clin Immunol Pract 2019;7(8):2583-2591. [doi: 10.1016/j.jaip.2019.03.041] [Medline: 30954644]
14. Glanz K, Rimer BK, Viswanath K, editors. Health Behavior: Theory, Research, and Practice. San Francisco, CA: Jossey-Bass; 2015.

15. West JH, Belvedere LM, Andreasen R, Frandsen C, Hall PC, Crookston BT. Controlling your 'app'etite: how diet and nutrition-related mobile apps lead to behavior change. JMIR Mhealth Uhealth 2017 Jul 10;5(7):e95 [FREE Full text] [doi: 10.2196/mhealth.7410] [Medline: 28694241]
16. Sieverdes JC, Gregoski M, Patel S, Williamson D, Brunner-Jackson B, Rundbaken J, et al. mHealth medication and blood pressure self-management program in Hispanic hypertensives: a proof of concept trial. Smart Homecare Technol Telehealth 2013 Oct;1:1. [doi: 10.2147/shtt.s49633]
17. Sieverdes JC, Price M, Ruggiero KJ, Baliga PK, Chavin KD, Brunner-Jackson B, et al. Design and approach of the living organ video educated donors (LOVED) program to promote living kidney donation in African Americans. Contemp Clin Trials 2017 Oct;61:55-62 [FREE Full text] [doi: 10.1016/j.cct.2017.07.007] [Medline: 28687348]
18. Geryk LL, Roberts CA, Sage AJ, Coyne-Beasley T, Sleath BL, Carpenter DM. Parent and clinician preferences for an asthma app to promote adolescent self-management: a formative study. JMIR Res Protoc 2016 Dec 6;5(4):e229 [FREE Full text] [doi: 10.2196/resprot.5932] [Medline: 27923777]
19. Sleath B, Carpenter DM, Davis SA, Watson CH, Lee C, Loughlin CE, et al. Improving youth question-asking and provider education during pediatric asthma visits. Patient Educ Couns 2018 Jun;101(6):1051-1057 [FREE Full text] [doi: 10.1016/j.pec.2018.01.013] [Medline: 29402570]
20. Lejuez CW, Hopko DR, Acierno R, Daughters SB, Pagoto SL. Ten year revision of the brief behavioral activation treatment for depression: revised treatment manual. Behav Modif 2011 Mar;35(2):111-161. [doi: 10.1177/0145445510390929] [Medline: 21324944]
21. Deci EL, Ryan RM. The 'what' and 'why' of goal pursuits: human needs and the self-determination of behavior. Psychol Inq 2000 Oct;11(4):227-268. [doi: 10.1207/s15327965pli1104_01]
22. Silva MN, Marques MM, Teixeira PJ. Testing theory in practice: the example of self-determination theory-based interventions. Eur Health Psyc 2014;16(5):171-180 [FREE Full text]
23. Teufel II RJ, Patel SK, Shuler AB, Andrews AL, Nichols M, Ebeling MD, et al. Smartphones for real-time assessment of adherence behavior and symptom exacerbation for high-risk youth with asthma: pilot study. JMIR Pediatr Parent 2018 Oct 5;1(2):e8 [FREE Full text] [doi: 10.2196/pediatrics.9796] [Medline: 31518299]
24. Moore RC, Depp CA, Wetherell JL, Lenze EJ. Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. J Psychiatr Res 2016 Apr;75:116-123 [FREE Full text] [doi: 10.1016/j.jpsychires.2016.01.011] [Medline: 26851494]
25. Burke LE, Shiffman S, Music E, Styn MA, Kriska A, Smailagic A, et al. Ecological momentary assessment in behavioral research: addressing technological and human participant challenges. J Med Internet Res 2017 Mar 15;19(3):e77 [FREE Full text] [doi: 10.2196/jmir.7138] [Medline: 28298264]
26. Guest G, Macqueen K, Namey E. Applied Thematic Analysis. Thousand Oaks, CA: Sage Publications; 2012.
27. Schreier M. Qualitative Content Analysis in Practice. Thousand Oaks, CA: Sage Publications; 2012.
28. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
29. Hesse-Biber S, Leavy P. The Practice of Qualitative Research. Second Edition. Thousand Oaks, CA: Sage Publications; 2011.
30. Morgan DL. Integrating Qualitative and Quantitative Methods: A Pragmatic Approach. Thousand Oaks, CA: Sage Publications; 2014.
31. Guba EG, Lincoln YS. Naturalistic Inquiry. Thousand Oaks, CA: SAGE Publications; 1985.
32. Denzin N, Lincoln Y. Handbook of Qualitative Research. Fourth Edition. Thousand Oaks, CA: Sage Publications; 2011.
33. Patrick H, Williams GC. Self-determination theory: its application to health behavior and complementarity with motivational interviewing. Int J Behav Nutr Phys Act 2012 Mar 2;9:18 [FREE Full text] [doi: 10.1186/1479-5868-9-18] [Medline: 22385676]
34. Tran N, Coffman JM, Sumino K, Cabana MD. Patient reminder systems and asthma medication adherence: a systematic review. J Asthma 2014 Jun;51(5):536-543. [doi: 10.3109/02770903.2014.888572] [Medline: 24506699]

Abbreviations

EMA: ecological momentary assessment
KII: key informant interview
NH: non-Hispanic
SAMS: Smartphone Asthma Management System
SDT: self-determination theory

Edited by E Borycki, G Eysenbach; submitted 28.06.19; peer-reviewed by A Dima, C Hall, R Ramsey; comments to author 09.10.19; revised version received 20.12.19; accepted 10.04.20; published 03.07.20.

Please cite as:
Nichols M, Miller S, Treiber F, Ruggiero K, Dawley E, Teufel II R
Patient and Parent Perspectives on Improving Pediatric Asthma Self-Management Through a Mobile Health Intervention: Pilot Study
JMIR Form Res 2020;4(7):e15295
URL: https://formative.jmir.org/2020/7/e15295
doi:10.2196/15295
PMID:32442127

Original Paper

Use of a Mobile App to Augment Psychotherapy in a Community Psychiatric Clinic: Feasibility and Fidelity Trial

Atif Adam[1], MPH, MD, PhD; Ameena Jain[2], CRNP, DNP; Alexandra Pletnikova[3], BA; Rishi Bagga[1], BA; Allison Vita[1], BSc; Lisa N Richey[3], BA; Neda Gould[3], PhD; Supriya Munshaw[4], PhD; Kavi Misrilall[1], BSc, MBA; Matthew E Peters[1,3], MD

[1]Rose: Smarter Mental Health, Washington, DC, United States

[2]Key Point Health Services, Inc, Baltimore, MD, United States

[3]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University, Baltimore, MD, United States

[4]Carey Business School, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:
Atif Adam, MPH, MD, PhD
Rose: Smarter Mental Health
3400 Prospect St NW
Washington, DC, 20007
United States
Phone: 1 4103363626
Email: dratif@askrose.com

Abstract

Background: Even though 1 in 5 Americans experience some form of mental illness each year, 80% have been shown to discontinue psychotherapy prematurely. The traditional psychotherapy service delivery model, consisting of isolated clinical sessions, lacks the ability to keep patients engaged outside clinical sessions. Newer digital mental health platforms can address the clinical need for a robust tool that tracks mental well-being and improves engagement in patients with depressive symptoms.

Objective: The primary goals of this feasibility study were to (1) assess compliance among providers and their patients with a digital mental health platform protocol, and (2) examine the usability and fidelity of a mobile app through structured participant feedback.

Methods: A sample of 30 participants was recruited for a 5-week study from a community-based mental health clinic in Baltimore, Maryland, USA. Inclusion criteria were: aged 18 years or older, having access to a smartphone, and having at least mild-to-moderate depression and/or anxiety as measured by the Patient Health Questionnaire-9 (PHQ-9) and Generalized Anxiety Disorder-7 (GAD-7) scales, respectively. Eligible participants were randomized into one of two study arms: (1) the intervention arm or (2) the waitlist control arm. Participants in the intervention arm were asked to download the Rose app and were prompted to complete clinical assessments (PHQ-9 and GAD-7) every other week, daily mood and anxiety Likert scales, and daily journal entries. The participants in the waitlist arm served as controls for the study and completed the clinical assessments only. Both arms engaged in weekly psychotherapy sessions, with participant in-app input informing the psychotherapy process of the intervention arm, while those in the waitlist control arm continued their standard care. Outcomes of interest included adherence to completion of in-app assessments and usability of the Rose mobile app assessed through the modified Mobile Application Rating Scale.

Results: Over the study period, a sample of 30 participants used the Rose app 2834 times to complete clinical assessments. On average, 70% (21; 95% CI 61.14%-77.41%) of participants completed mood and anxiety daily check-ins and journal entries 5 days per week. Nearly all participants (29/30, 97%) completed all PHQ-9 and GAD-7 in-app scales during the study. Subjective impressions showed that 73% (22/30) of participants found the mobile app to be engaging and in line with their needs, and approximately 70% (21/30) of participants reported the app functionality and quality of information to be excellent. Additionally, more than two-thirds of the participants felt that their knowledge and awareness of depression and anxiety management improved through using the app.

Conclusions: Steady compliance and high app ratings showcase the utility of the Rose mobile mental health app in augmenting the psychotherapy process for patients with mood disorders and improving mental health knowledge. Future studies are needed to further examine the impact of Rose on treatment outcomes.

Trial Registration: ClinicalTrials.gov NCT04200170; https://clinicaltrials.gov/ct2/show/NCT04200170

(JMIR Form Res 2020;4(7):e17722) doi:10.2196/17722

KEYWORDS

mobile app; mental health; depression; anxiety

Introduction

Each year, 1 in 5 Americans (approximately 46.6 million people) experience some form of mental illness [1-3]. The National Institute of Mental Health estimated that 17.3 million adults in the United States had at least one major depressive episode in 2017, representing 7.1% of the population. However, studies show that up to 60% of adults suffering from mental illness report an inability to receive appropriate treatment and close to 80% seeking care prematurely discontinue psychotherapy [4-6]. Reasons for discontinuation include the stigma associated with psychotherapy, difficulties in connecting with therapists outside of clinical sessions, and poor fit with therapists [7,8].

In recent years, feedback-informed care (FIC) models have been used to improve patient compliance and psychotherapy outcomes [9,10]. FIC models are designed to encourage patients to provide feedback on care progress, assess well-being outside clinical sessions, and utilize a data driven approach to bolster the therapeutic alliance and build individually tailored therapies [9]. A key gap in FIC models has been the availability of resources and tools that allow patients to stay connected to their providers and capturing mental well-being data that currently can only be collected during relatively infrequent in-person psychotherapy sessions [11,12].

The widespread use and availability of smartphones have made successful adoption of mobile technology increasingly crucial in developing FIC models [12-14] and identifying best practices. Unfortunately, of the 10,000 mental health apps currently available, less than 5% are backed by evidenced-based research [15]. Novel health technologies looking to revolutionize mental health care delivery and empower patients need to have an evidence-base that captures engagement of patients in their daily life with their therapists and the utility of collecting wellbeing data to address specific health concerns. The Rose digital health platform was developed in response to the clinical need for a robust tool to track mental wellbeing outside of clinical care. Rose's digital health platform integrates a patient-facing mobile phone app for real-time mood tracking, clinical surveys, daily journaling, and curated insights with a web-based clinical dashboard allowing clinicians to check on their patient's mental health status between clinical visits (Figure 1).

The primary goals of this feasibility study were to (1) assess the usage compliance with a mobile mental health platform among providers and their patients with at least mild-to-moderate depression and/or anxiety, and (2) examine the usability and fidelity of the mobile app through structured participant feedback. We hypothesized consistent usage adherence to the mobile platform over the study's duration. A secondary exploratory aim evaluated the short-term impact of mobile app usage on mood and anxiety symptomatology.

Figure 1. Rose mobile app screens.

Methods

Recruitment

Participants were recruited from the patient population at Key Point Health Services, a sizeable outpatient community psychiatric clinic in Baltimore, Maryland, USA. Inclusion criteria consisted of age 18 years or older, regularly making on-going appointments with a psychotherapist at the clinic, having a score of >5 on the Patient Health Questionnaire-9 (PHQ-9) and/or >5 on Generalized Anxiety Disorders-7 (GAD-7), having access to a smartphone, and being fluent in English. Exclusion criteria included a current diagnosis of decompensated psychosis (eg, hallucinations, delusions, thought disorder) and/or current suicidal or homicidal ideation. The cut-off of 5 on the PHQ-9 and GAD-7 represent mild depression and mild anxiety from clinical severity categorization [16,17]. Eligible patients were informed of the study via brochures provided by their established psychotherapists. Interested patients were directed to a secure online screener, which collected data on inclusion and exclusion criteria, demographics, and the PHQ-9 and GAD-7. This pilot was designed as an open-label trial with randomized allocation into study arms. A study member who was not involved in the day-to-day study operations used the STATA 14 (StataCorp) sequence generator feature for randomization. These were stored in sealed envelopes. At study registration and the online screening, the study coordinator opened the randomization envelopes to allocate the eligible participants to either the intervention arm or the control arm. Eligible, consented participants were then randomized via computer algorithm into one of two study arms: (1) the intervention arm or (2) the waitlist control arm, in a 2:1 ratio. All nine study therapists had appointments with at least one participant from each arm. Of note, no patient medical records were accessed as part of the study.

Intervention Arm Design

Eligible participants continued meeting with their therapists for weekly psychotherapy sessions. Therapists were asked to register for the Rose web-based clinical dashboard, which generated a unique therapist ID. Participants in the intervention arm were asked to download the Rose app to their mobile device and were given their therapist's unique ID for linking within the app. Participants were informed that the health metrics would be shared with their therapist to help the clinician better plan upcoming sessions. Participants were prompted to complete a prepilot battery of assessments (PHQ-9, GAD-7, mood and anxiety Likert scales) upon first logging into the app. This battery was also pushed to the users at 2- and 4-week (postpilot) time points. Those in the intervention arm were also asked to continue using the Rose app daily during the 4-week period. The therapist was asked to use the Rose clinical dashboard to review data in conjunction with each weekly in-clinic session for four sessions. All in-clinic sessions were scheduled in the same clinical space where participants previously received care. Each clinical session was approximately 45 minutes long, and therapist documentation was done using their standard of care electronic medical record system.

Control Arm Design

The participants in the waitlist control arm served as controls for the study. They completed the prepilot and postpilot assessments (PHQ-9, GAD-7, mood and anxiety Likert scales) only. These participants continued their weekly standard care with the psychotherapy process. Waitlist participants were offered entrance to the intervention arm at the end of the study period or earlier if participants in the intervention arm dropped out mid-study. During their time on the waitlist, participants could reach out to study personnel if they needed assistance with their psychiatric care.

Ethics

The study was reviewed and approved by the Advarra Commercial Institutional Review Board. For data security purposes, Rose utilizes cloud-based file-sharing and data storage services that are ISO 27001 certified and HIPAA (Health Insurance Portability and Accountability Act) compliant. If at any point during the study, participants denoted suicidal ideation or other acute psychiatric concerns, they were automatically provided with emergency contact numbers and the study clinician (AJ) reached out directly. Additionally, if a high-risk participant was identified during in-person psychotherapy sessions, the therapist contacted the study clinician immediately with any concerns. If it was deemed that a participant required a higher level of care (eg, inpatient hospitalization), they were removed from the study.

Mobile Mental Health App (Rose)

Rose is a mobile health platform that consists of two primary components: (1) a patient-facing mobile app and (2) a clinician-facing web-based dashboard. Patients and clinicians are linked within the Rose system using a unique identifier that the clinician provides to the patient. The Rose app allows patients to track their mood and anxiety levels in real-time, complete validated assessments, and keep an in-app daily journal. Rose utilizes machine learning algorithms that identify mental health status and provide curated, individualized in-app daily insights for self-care. The Rose clinician-facing dashboard is designed for use in conjunction with in-person appointments. It is meant to augment in-person care by summarizing the data entered by the patient into the mobile app. Providers are given a walkthrough on using the dashboard and can reach the Rose team with any questions directly through the dashboard.

Measures

The measures were as follows:

- Patient Health Questionnaire-9 (PHQ-9) [17] and Generalized Anxiety Disorder-7 (GAD-7) [18] Scale. The PHQ-9 and GAD-7 are valid and reliable screening tools for depression and anxiety, respectively. These assessments parallel the diagnostic symptom criteria that define DSM-IV major depression and GAD. The Rose app prompted users to complete these assessments every two weeks. The app follows the format and temporal framework corresponding to DSM-IV criteria.
- Modified Mobile Application Rating Scale (mMARS) (Multimedia Appendix 1) [19]. The MARS is a well-established framework for classifying and assessing

the objective and subjective quality of apps, as well as their perceived impact. It is designed to score apps on the criteria of engagement, functionality, aesthetics, information quality, and subjective app quality. While the MARS framework is extensive, sections within the scale were not pertinent for this study and the overall length was deemed too tedious for patients to complete. The mMARS is a shortened version of the MARS, keeping the pertinent sections and modifying verbiage to reflect the current app usage.

- Mood & Anxiety Likert Scales [20]. The Rose app sends timed notifications once a day to rate current mood and anxiety, each on a 5-point Likert scale.
- In-app Journaling. The in-app journaling feature allows the user to enter a free text description of how they are doing, like in the case of a diary. A built-in sentiment analysis analyzes the entered sentences.

Statistical Power Calculation

The power calculation for this study is based on a meta-analysis of internet-delivered treatments for adult depression and anxiety. It was found that internet interventions used in tandem with professional psychotherapy support showed an average between-group effect size of d=0.61 (large mean effect size of d=1.0). Based on the study's five-psychotherapy session design (first session for on-boarding and four follow-up sessions), and paired-sample design (multiple measurements from the same person), a sample size of 30 has sufficient power (90%) to detect a moderately large effect size (d=0.8) [21]. Additionally, we recruited 15 patients as a waitlist control group who would be included in the intervention arm if drop-out or removal occurred.

Primary Analyses

The primary analysis included checking distributional assumptions from the primary scores and assessing relationships among covariates. Continuous variables were described using mean, standard deviation, median, minimum and maximum, and 95% confidence interval for parameter estimation.

To determine whether any significant differences between groups existed at baseline, independent t tests were conducted on continuous baseline variables (eg, age, PHQ-9, GAD-7), and chi-square analyses were performed on categorical variables (gender, race, ethnicity).

Descriptive in-app metrics were calculated for compliance of assessments with PHQ-9 and GAD-7 every two weeks and daily mood/anxiety Likert scales and journaling. In-app use was also explored for time-of-day frequency and weekly variations.

- Adherence over time. Adherence was defined as the completion of in-app assessments at predefined intervals. For example, daily adherence to mood assessment was seen if a participant completed at least one of the mood or anxiety Likert scales each day. Non-adherence was recorded as days with no data. Each assessment measure was reviewed individually for the duration of the study (4 weeks) for the percentage of adherent days and was stratified for each study week in which the app was being used. Logistic regression models were carried out to examine whether adherence to in-app assessments would be associated with the severity of depressive symptoms at baseline.
- Fidelity of app use. Fidelity defines the degree to which programs are implemented as intended by the program developer [22]. In this study we looked at five main components of fidelity for the mobile app using the mMARS scale: Engagement, Functionality, Aesthetics, Quality of Information, and Subjective Quality. Summary proportions from the mMARS were reviewed in all five areas of app use.

Exploratory Analyses

Pre-post analysis (Wilcoxon matched-pairs signed-rank test) was used to evaluate the short-term impacts of mobile app usage in the intervention arm and the control arm within in-person psychotherapy sessions.

All analyses were conducted using STATA 14 (StataCorp). All statistical tests were judged for significance based upon a two-tailed alpha level of P<.05. All subjects were included in intention-to-treat (ITT) analyses.

Results

Participant Characteristics

Table 1 shows the demographic information and baseline scores on clinical variables for those with data from the entire sample (n=45). Each therapist saw on average 3 patients (CI: 1-5). Participants in the intervention arm (n=30) were an average of 36 years old. Close to 60.00% of the participants were female and 29.03% reported completing college or higher level of education. Participants were mostly non-Hispanic making up 92.58%. 82.14% were Caucasian, 10.71% African American, and 7.14% identified with more than one race. Just over half (51.61%) used Android devices as their daily smartphone and the rest used Apple iOS devices. Other than the proportion of female participants (P=.02), there were no significant demographic differences between arms.

Looking at baseline mental health status, 42.22% of the sample was in the moderately-severe or severe range of depression, as measured by the PHQ-9. GAD-7 ratings showed 9.60% of the participants were in the moderately-severe or critical range for anxiety. There were no statistical differences in the distribution of participants randomized to each arm based on baseline mental health status.

Table 1. Demographic and baseline characteristics of study participants.

Characteristic	Total (n=45)	Intervention arm (n=30)	Waitlist arm (n=15)	*P* value
Age, median (range)	31 (18-65)	29 (18-65)	35 (19-54)	.75
Female sex, n (%)	29 (70.73%)	16 (60.71)	12 (92.31)	.02
Education, n (%)				.32
High school or less	29 (64.44)	19 (63.33)	10 (76.92)	
College or higher	12 (26.67)	9 (30.00)	29.03%	
Race, n (%)				.74
White	33 (78.57%)	25 (83.33)	9 (69.23)	
Black	5 (11.90%)	3 (10.71)	2 (15.38)	
Hispanic	1 (2.38%)	0 (0)	1 (7.69)	
Mixed	3 (7.14%)	2 (7.14)	1 (7.69)	
Device, n (%)				.62
Android	19 (42.22%)	14 (46.67)	5 (38.46)	
Apple iOS	25 (55.56%)	16 (53.33)	8 (61.54)	
PHQ-9, mean (SD)	13.47 (5.90)	12.35 (5.57)	16.08 (5.32)	.89
GAD-7, mean (SD)	11.96 (6.02)	11.19 (5.77)	13.23 (6.50)	.74

Adherence Over Time

Participants in the intervention arm used the Rose mobile app over 4 weeks as part of care management. Table 2 details the descriptive statistics on participants' daily use of the mood Likert scale, anxiety Likert scale, and online journaling; overall and by study week. Over the study period, patients adhered to completing all three assessments at an average of 69% per week (5 days a week). As seen in Table 2, adherence across all three estimates was reasonably similar. Adherence was highest in the first week, averaging at 81.17% (95% CI: 75.31%-87.04%) across all three assessments, with a linear decrease at week 4 to 56.03% (95% CI: 43.92%-68.14%) adherence. Majority adherence of 80% (24/30 days) was seen in 14 of 30 participants completing the mood Likert scale, 15 of 30 participants completing the anxiety Likert scale, and 13 of 30 participants completing the journal.

Looking at the use of PHQ-9 and GAD-7 in-app scales (Table 3), participants adhered to filling in all three assessments at an average rate of 97% over the study period (every two weeks). Full adherence (3/3 assessments) was seen in 27/30 of the patients.

Table 2. Weekly adherence rates for in-app mood and anxiety Likert scales and journaling.

Week	Adherence % (SD)		
	Mood check-in	Anxiety check-in	Journal
1	82.86 (14.24)	81.90 (16.06)	76.19 (18.89)
2	70.95 (25.02)	70.48 (28.13)	60.95 (33.19)
3	64.29 (29.95)	61.90 (32.10)	52.86 (33.91)
4	58.10 (32.69)	62.86 (33.79)	57.14 (35.04)

Table 3. Biweekly adherence rates for in-app PHQ-9 and GAD-7 assessments.

Week	Adherence % (SD)	
	PHQ-9[a]	GAD-7[b]
0	100 (0)	100 (0)
2	100 (0)	100 (0)
4	90 (30.51)	90 (30.51)

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

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

Fidelity of App Use

Using the mMARS scale we evaluated the perceived usability of the mobile app as rated by participants (Figure 2). Overall, participants thought the app was very engaging, with approximately 73% finding Rose very interesting and appropriate for the target audience. Approximately 83% of the participants reported the app to be very accurate, easy to use, and easy to navigate. Similarly, 75% reported the layout and graphics to be very pleasing and 71% felt the overall look was only adequate. Furthermore, 54% reported that the information provided through the app was of good or excellent quality, and 77% of the participants felt it covered a comprehensive range of topics. Overall, 67% said they would recommend the app outside of the study.

We further asked participants about the perceived impact of the Rose app on their knowledge, attitudes, intentions to change, as well as the likelihood of actual change in depression and/or anxiety symptoms (Figure 3). More than two-thirds of the participants felt that their knowledge of the topic and awareness of the importance of depression and anxiety management improved through using the app. Similarly, 67% reported they were more than likely to take progressive steps to address depression and more than half stated that the app made them likely to seek help for depressive symptoms when needed.

Figure 2. mMARS – Perceived Usability of Rose.

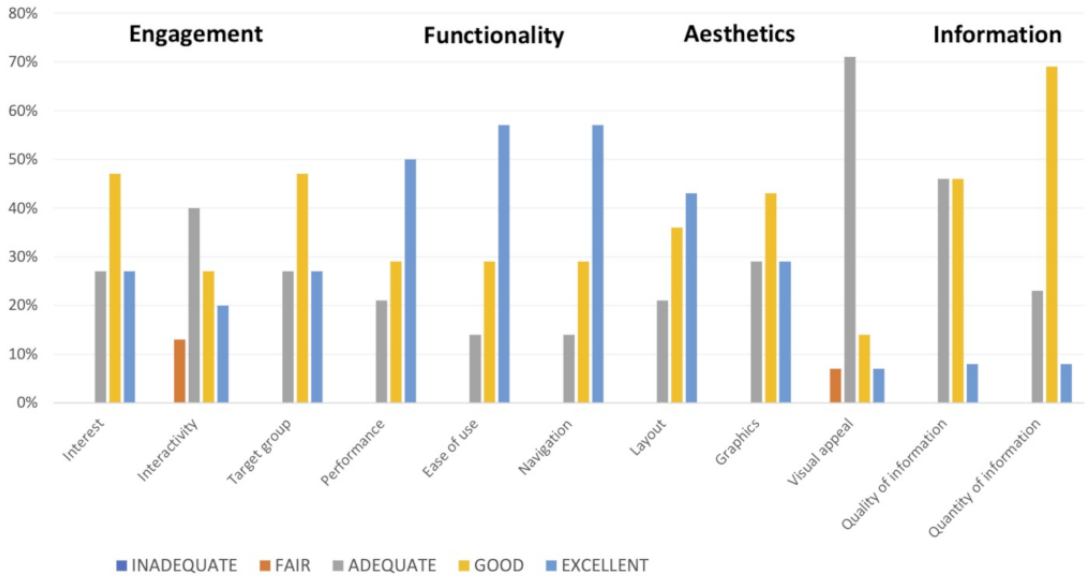
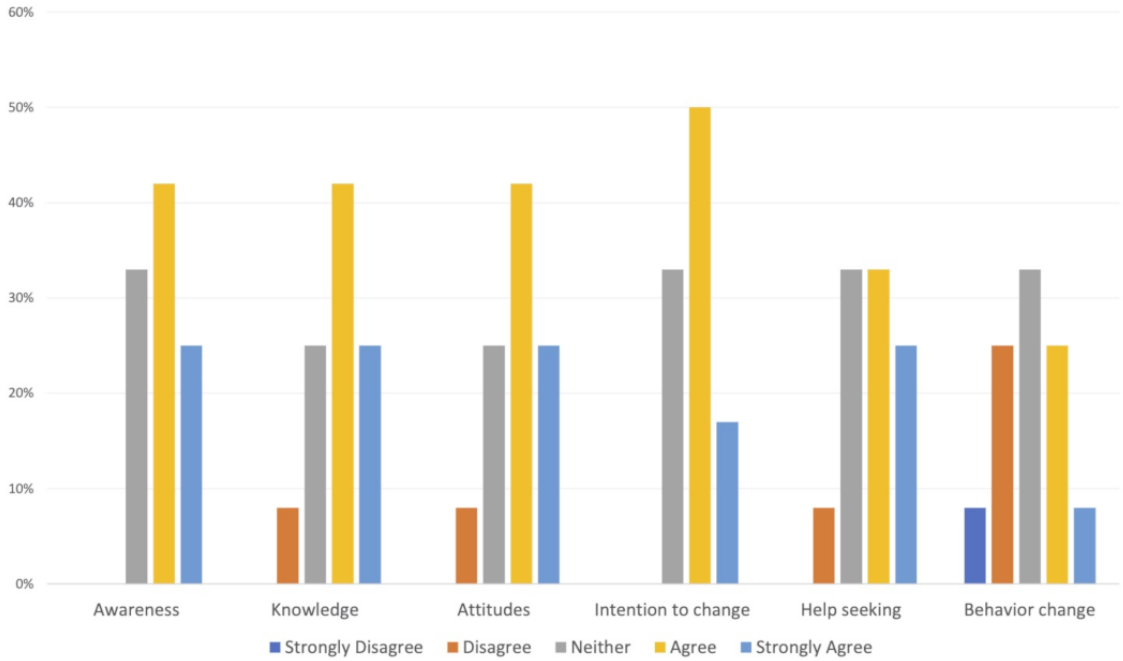


Figure 3. Impact of App Usage on Mental Health Knowledge and Attitudes.



Exploratory Analyses

We looked for changes in depression and anxiety symptoms (measured through PHQ-9 and GAD-7) before and after the study in both the intervention and control arms (Table 4).

Participants in the intervention arm showed statistically significant improvements in measurements of both depression and anxiety symptoms. On the PHQ-9, there was an average 4-point improvement in participants over four weeks (prepilot average score: 15 vs postpilot average score: 11; *P*=.01). On the GAD-7, there was an average 3-point improvement in participants over four weeks (prepilot average score: 11 vs postpilot average score: 8; *P*=.005). Overall, 47% of patients with moderate-to-severe depression improved to mild depression and 56% of patients with moderate-to-severe anxiety improved to mild anxiety by the end of the study. Additionally, Kruskal-Wallis H test revealed statistically significant improvements in both PHQ-9 (*P*=.02) and GAD-7 (*P*=.001) scores over the course of the study in the intervention group.

Among the six controls who completed the pre- and postpilot assessments, neither PHQ-9 nor GAD-7 scores showed any significant changes (Table 5). Even though there was a difference in mean scores on both PHQ-9 and GAD-7 scales, statistical comparison on PHQ-9 and GAD-7 scores between intervention and control arms did not show statistical significance.

Table 4. Pre-post change in depression and anxiety symptoms: intervention (n=30) vs control (n=15).

Variable	Prepilot assessment, mean (SD)	Postpilot assessment, mean (SD)	*P* value
Intervention arm (n=30)			
PHQ-9[a]	15.33 (8.94)	11.04 (8.35)	.01
GAD-7[b]	11.43 (6.46)	8.17 (6.06)	.005
Control arm (n=15)			
PHQ-9	13.83 (6.74)	13.17 (7.68)	.53
GAD-7	10.17 (7.86)	10.17 (8.75)	1.0

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

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

Table 5. Pre-post analysis of changes in depression and anxiety symptoms.

Variable	Intervention arm, mean (SD)	Control arm, mean (SD)	*P* value
PHQ-9[a] change	–4.30 (1.54)	0.67 (0.99)	.11
GAD-7[b] change	–3.11 (1.23)	0 (1.29)	.41

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

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

Discussion

Principal Findings

This study demonstrated the fidelity of using a digital mental health platform (Rose) for monitoring daily mental health and wellbeing outside of psychotherapy sessions and maintaining engagement in patients with depression and anxiety symptoms. Exploratory analyses examined changes in depressive and anxiety symptoms at baseline versus after study interventions. Improved PHQ-9 scores were observed in 76% (23/30) of intervention arm participants and 70% (21/30) showed improvements in their GAD-7 scores over the study period. Furthermore, patients rated the Rose mobile app as having high engagement and usability as well as content curation relevant to their needs.

The study paralleled current psychotherapy care practices for moderate depression. These results showed using Rose is no worse than not using an app within psychotherapy sessions. Consistent compliance by patients between and during clinical psychotherapy sessions indicates that high levels of adherence can be achieved and retained. A larger multisite randomized controlled trial is being planned to look at the impact of the Rose digital health platform with an active (alternative mobile app) control.

Comparison With Prior Work

The economic impact of mental illness is marked, with the United States alone seeing $210 billion in annual medical expenditures tied to mental health disorders [23,24]. The added societal burdens of mental illness include productivity losses, absenteeism, and reduced economic growth [25]. On average, depressive symptoms lead to approximately 27 lost workdays and 18 days of reduced productivity per year in people diagnosed with depression [26]. Research attention to address these needs has exponentially increased in recent years, especially around the area of mobile mental health apps [14,27]. Research groups working in this space have found, through both observational and clinical trials, positive impacts on the utility of smartphones with onboard sensors to diagnose psychiatric disorders [12]. However, studies of adherence to use of mobile mental health apps have been limited. A meta-analysis by Tourous et al looked at 18 clinical trials that investigated the

efficacy of smartphone interventions targeting depressive symptoms. The authors note that lower rates of engagement over time have been found in numerous mental health app studies [28], with higher rates being more common in shorter intervention interactions. Rates of adherence were notably higher with apps that had added components of human support, with real world engagement rates close 17% for peer support apps, but adding human interactions may reduce scalability, especially in low resource settings [28]. This suggests that intervention designs have to be adaptable and customizable to the ways in which people use smartphones. Results from the current study demonstrate the utility of smartphone-driven interventions to improve on FIC and in maintaining nonhuman-supported adherence. Furthermore, there are few studies that have looked at design, fidelity, and usability of mental health apps tied to intervention outcome and patient engagement [29,30]. This study is also one of the first to look at mobile mental health app outcome metrics and interaction principles in a longitudinal investigational format.

Limitations

Participants in this study were selected from a single study site and were not randomized by therapists. Thus, there may be a sampling bias that limits generalizability of the results. Furthermore, the sample consisted of 60%-90% women, which is not representative of the population of individuals with depression/anxiety in clinical settings or the community. While the pilot study had a waitlist control arm, it did not include an active comparison group (ie, using an alternate mobile app). Furthermore, only 40% of the control arm completed the postpilot follow-up. This limited the pre-post analytical comparison. A follow-up randomized control trial is needed to evaluate inferences regarding the efficacy and contrast of effect sizes between the Rose digital health platform and other interventions.

Conclusions

Steady compliance and high app ratings showcase the usability of Rose in practice and the utility of digital FIC models in improving patient compliance to therapy. Although exploratory, reduction in clinical depression and anxiety symptoms over the four-week study period highlights the potential utility of the Rose app for psychotherapy augmentation. Future studies are being developed to (1) further evaluate the clinical effectiveness of the Rose app through a multisite micro-randomized clinical trial, and (2) examine the impact of the Rose digital health platform on other diagnoses, including substance use and pain.

Acknowledgments

All authors have contributed substantially to the conceptualization, design, analysis, interpretation, and drafting of the manuscript. AJ provided expert guidance and leadership throughout the project – she was involved throughout study management, overseeing participant recruitment, staff training and writing the manuscript. MP and AA provided leadership in helping guide the app review and training of staff, technical support during the study for the health center, study design, parameterizing the models, data analysis and writing the manuscript. AP focused on participant recruitment and data collection. KM provided oversight and technical input on the mobile app integration into clinical practice and working with patient app issues. RB was involved in staff training and implementation of the clinical dashboard, gathering user feedback, and manuscript preparation. LR and NG were involved in manuscript drafting. The authors are grateful to Mr. Raghav Sambasivan, Mr. Shravan Venkatesan, Mr. Darsh Patel, Mr. Anthony Jairam, and Ms. Allison Vita for their tremendous support helping run the Rose study and for helping build the mobile app platform. The authors also appreciate the supportive partnership of Key Point Health Services, Inc.

Conflicts of Interest

AA, KM, and MP are on the executive team of Ask Rose Inc, the company that developed the mobile app used in this study. AV is an employee with Ask Rose, Inc. NG is a clinical consultant to Ask Rose, Inc. This study was funded by Ask Rose, Inc.

Multimedia Appendix 1
Modified Mobile Application Rating Scale (mMARS).
[PDF File (Adobe PDF File), 1682 KB - formative_v4i7e17722_app1.pdf]

References

1. Indu PS, Anilkumar TV, Pisharody R, Russell PSS, Raju D, Sarma PS, et al. Prevalence of depression and past suicide attempt in primary care. Asian J Psychiatr 2017 Jun;27:48-52. [doi: 10.1016/j.ajp.2017.02.008] [Medline: 28558895]
2. Pratt LA, Brody DJ. Depression in the U.S. household population, 2009-2012. NCHS Data Brief 2014 Dec(172):1-8 [FREE Full text] [Medline: 25470183]
3. Pratt L, Brody DJ. Depression in the United States household population, 2005-2006. NCHS Data Brief 2008 Sep(7):1-8 [FREE Full text] [Medline: 19389321]
4. Sayers J. Mental health: new understanding, new hope. In: The world health report 2001. Geneva: World Health Organization; 2001.

5. Wang PS, Berglund P, Olfson M, Pincus HA, Wells KB, Kessler RC. Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun 01;62(6):603-613. [doi: 10.1001/archpsyc.62.6.603] [Medline: 15939838]
6. 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 01;62(6):629-640. [doi: 10.1001/archpsyc.62.6.629] [Medline: 15939840]
7. Swift JK, Greenberg RP. Premature discontinuation in adult psychotherapy: a meta-analysis. J Consult Clin Psychol 2012 Aug;80(4):547-559. [doi: 10.1037/a0028226] [Medline: 22506792]
8. Werbart A, Andersson H, Sandell R. Dropout revisited: patient- and therapist-initiated discontinuation of psychotherapy as a function of organizational instability. Psychother Res 2014 Feb 13;24(6):724-737. [doi: 10.1080/10503307.2014.883087] [Medline: 24524334]
9. Miller SD, Hubble MA, Chow D, Seidel J. Beyond measures and monitoring: Realizing the potential of feedback-informed treatment. Psychotherapy (Chic) 2015 Dec;52(4):449-457. [doi: 10.1037/pst0000031] [Medline: 26641375]
10. Delgadillo J, de Jong K, Lucock M, Lutz W, Rubel J, Gilbody S, et al. Feedback-informed treatment versus usual psychological treatment for depression and anxiety: a multisite, open-label, cluster randomised controlled trial. The Lancet Psychiatry 2018 Jul;5(7):564-572. [doi: 10.1016/s2215-0366(18)30162-7]
11. Delgadillo J, Overend K, Lucock M, Groom M, Kirby N, McMillan D, et al. Improving the efficiency of psychological treatment using outcome feedback technology. Behav Res Ther 2017 Dec;99:89-97. [doi: 10.1016/j.brat.2017.09.011] [Medline: 29024821]
12. Lindhiem O, Bennett CB, Rosen D, Silk J. Mobile technology boosts the effectiveness of psychotherapy and behavioral interventions: a meta-analysis. Behav Modif 2015 Nov 17;39(6):785-804 [FREE Full text] [doi: 10.1177/0145445515595198] [Medline: 26187164]
13. Lui JHL, Marcus DK, Barry CT. Evidence-based apps? A review of mental health mobile applications in a psychotherapy context. Professional Psychology: Research and Practice 2017 Jun;48(3):199-210. [doi: 10.1037/pro0000122]
14. Menon V, Rajan T, Sarkar S. Psychotherapeutic Applications of Mobile Phone-based Technologies: A Systematic Review of Current Research and Trends. Indian J Psychol Med 2017;39(1):4-11 [FREE Full text] [doi: 10.4103/0253-7176.198956] [Medline: 28250552]
15. Torous J, Roberts LW. Needed Innovation in Digital Health and Smartphone Applications for Mental Health: Transparency and Trust. JAMA Psychiatry 2017 May 01;74(5):437-438. [doi: 10.1001/jamapsychiatry.2017.0262] [Medline: 28384700]
16. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
17. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
18. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and Standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the General Population. Medical Care 2008;46(3):266-274. [doi: 10.1097/mlr.0b013e318160d093]
19. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
20. Ahearn EP. The use of visual analog scales in mood disorders: A critical review. Journal of Psychiatric Research 1997 Sep;31(5):569-579. [doi: 10.1016/s0022-3956(97)00029-0]
21. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009 Dec 15;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
22. Dusenbury L, Brannigan R, Falco M, Hansen WB. A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. Health Educ Res 2003 Apr;18(2):237-256. [doi: 10.1093/her/18.2.237] [Medline: 12729182]
23. Roehrig C. Mental Disorders Top The List Of The Most Costly Conditions In The United States: $201 Billion. Health Aff (Millwood) 2016 Jun 01;35(6):1130-1135. [doi: 10.1377/hlthaff.2015.1659] [Medline: 27193027]
24. Davlasheridze M, Goetz SJ, Han Y. The Effect of Mental Health on US County Economic Growth. Review of Regional Studies 2018;48(2):155-171.
25. Trautmann S, Rehm J, Wittchen H. The economic costs of mental disorders: Do our societies react appropriately to the burden of mental disorders? EMBO Rep 2016 Sep 04;17(9):1245-1249 [FREE Full text] [doi: 10.15252/embr.201642951] [Medline: 27491723]
26. Brohan E, Thornicroft G. Stigma and discrimination of mental health problems: workplace implications. Occupational Medicine 2010 Aug 18;60(6):414-415. [doi: 10.1093/occmed/kqq048]
27. Boschen MJ, Casey LM. The use of mobile telephones as adjuncts to cognitive behavioral psychotherapy. Professional Psychology: Research and Practice 2008 Oct;39(5):546-552. [doi: 10.1037/0735-7028.39.5.546]

28. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: A systematic review and meta-analysis. J Affect Disord 2020 Feb 15;263:413-419. [doi: 10.1016/j.jad.2019.11.167] [Medline: 31969272]
29. Hays R, Henson P, Wisniewski H, Hendel V, Vaidyam A, Torous J. Assessing Cognition Outside of the Clinic: Smartphones and Sensors for Cognitive Assessment Across Diverse Psychiatric Disorders. Psychiatr Clin North Am 2019 Dec;42(4):611-625. [doi: 10.1016/j.psc.2019.08.003] [Medline: 31672211]
30. Hilty DM, Chan S, Torous J, Luo J, Boland RJ. Mobile Health, Smartphone/Device, and Apps for Psychiatry and Medicine: Competencies, Training, and Faculty Development Issues. Psychiatr Clin North Am 2019 Sep;42(3):513-534. [doi: 10.1016/j.psc.2019.05.007] [Medline: 31358129]

Abbreviations

FIC: feedback-informed care
GAD-7: Generalized Anxiety Disorder-7
HIPPA: Health Insurance Portability and Accountability Act
ITT: intention-to-treat (ITT)
mMARS: modified Mobile Application Rating Scale
PHQ-9: Patient Health Questionnaire-9

Edited by G Eysenbach; submitted 10.01.20; peer-reviewed by A Walsh, S Jahed, J D'Arcey; comments to author 09.03.20; revised version received 05.04.20; accepted 09.04.20; published 03.07.20.

Please cite as:
Adam A, Jain A, Pletnikova A, Bagga R, Vita A, N Richey L, Gould N, Munshaw S, Misrilall K, Peters ME
Use of a Mobile App to Augment Psychotherapy in a Community Psychiatric Clinic: Feasibility and Fidelity Trial
JMIR Form Res 2020;4(7):e17722
URL: https://formative.jmir.org/2020/7/e17722
doi:10.2196/17722
PMID:32618572

Original Paper

Usability and Acceptability of a Smartphone App to Assess Partner Communication, Closeness, Mood, and Relationship Satisfaction: Mixed Methods Study

Shelby L Langer[1], PhD; Neeta Ghosh[2], MPH, MA; Michael Todd[3], PhD; Ashley K Randall[4], PhD; Joan M Romano[5], PhD; Jonathan B Bricker[6,7], PhD; Niall Bolger[8], PhD; John W Burns[9], PhD; Rachel C Hagan[1], BS; Laura S Porter[10], PhD

[1]Center for Health Promotion and Disease Prevention, Edson College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

[2]Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, United States

[3]Edson College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

[4]Counseling and Counseling Psychology, College of Integrative Sciences and Arts, Arizona State University, Phoenix, AZ, United States

[5]Psychiatry and Behavioral Sciences, School of Medicine, University of Washington, Seattle, WA, United States

[6]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, United States

[7]Department of Psychology, University of Washington, Seattle, WA, United States

[8]Department of Psychology, Columbia University, New York, NY, United States

[9]Psychiatry and Behavioral Sciences, Rush Medical College, Rush University, Chicago, IL, United States

[10]Psychiatry and Behavioral Sciences, School of Medicine, Duke University, Durham, NC, United States

Corresponding Author:
Shelby L Langer, PhD
Center for Health Promotion and Disease Prevention
Edson College of Nursing and Health Innovation
Arizona State University
500 North Third Street
Phoenix, AZ, 85004
United States
Phone: 1 6024960823
Fax: 1 6024961128
Email: shelby.langer@asu.edu

Abstract

Background: Interpersonal communication is critical for a healthy romantic relationship. Emotional disclosure, coupled with perceived partner responsiveness, fosters closeness and adjustment (better mood and relationship satisfaction). On the contrary, holding back from disclosure is associated with increased distress and decreased relationship satisfaction. Prior studies assessing these constructs have been cross-sectional and have utilized global retrospective reports of communication. In addition, studies assessing holding back or perceived partner responsiveness have not taken advantage of smartphone ownership for data collection and have instead required website access or use of a study-provided device.

Objective: This study aimed to examine the (1) usability and acceptability of a smartphone app designed to assess partner communication, closeness, mood, and relationship satisfaction over 14 days and (2) between-person versus within-person variability of key constructs to inform the utility of their capture via ecological momentary assessment using the participants' own handheld devices.

Methods: Adult community volunteers in a married or cohabiting partnered relationship received 2 smartphone prompts per day, one in the afternoon and one in the evening, for 14 days. In each prompt, participants were asked whether they had conversed with their partner either since awakening (afternoon prompt) or since the last assessment (evening prompt). If yes, a series of items assessed enacted communication, perceived partner communication, closeness, mood, and relationship satisfaction (evening only). Participants were interviewed by phone, 1 week after the end of the 14-day phase, to assess perceptions of the app. Content analysis was employed to identify key themes.

Results: Participants (N=27; mean age 36, SD 12 years; 24/27, 89% female; 25/27, 93% white and 2/27, 7% Hispanic) responded to 79.2% (555/701) of the total prompts sent and completed 553 (78.9%) of those assessments. Of the responded prompts, 79.3% (440/555) were characterized by a report of having conversed with one's partner. The app was seen as highly convenient (mean 4.15, SD 0.78, scale: 1-5) and easy to use (mean 4.39, SD 0.70, scale: 1-5). Qualitative analyses indicated that participants found the app generally easy to navigate, but the response window too short (45 min) and the random nature of receiving notifications vexing. With regard to the variability of the app-delivered items, intraclass correlation coefficients were generally <0.40, indicating that the majority of the variability in each measure was at the within-person level. Notable exceptions were enacted disclosure and relationship satisfaction.

Conclusions: The findings of this study support the usability and acceptability of the app, with valuable user input to modify timing windows in future work. The findings also underscore the utility of an intensive repeated-measures approach, given the meaningful day-to-day variation (greater within-person vs between-person variability) in communication and mood.

(JMIR Form Res 2020;4(7):e14161) doi:10.2196/14161

KEYWORDS

ecological momentary assessment; smartphone; mobile phone; communication; disclosure; affect

Introduction

Interpersonal communication is critical to the development and maintenance of romantic relationships [1]. The manner in which partners convey verbal and nonverbal information to each other plays a major role in the psychological functioning of both individuals and their relationship as a whole [2]. Much attention has been paid to disclosure, the act of revealing inner experiences to an interaction partner [3]. Disclosure is part of a larger interactional process that can lead to intimacy if it is met with responsive listening, due to perceptions of being cared for and understood [4]. Indeed, open discussions, the expression of thoughts and feelings, and responsiveness have all been found to be associated with increased intimacy and, in turn, increased relationship satisfaction and decreased psychological distress [5,6]. On the contrary, avoidance behaviors such as holding back have been found to be associated with lower intimacy, lower relationship satisfaction, and greater distress [5-13].

Much of the research informing the understanding of the links between communication and mood, as well as communication and relationship quality, has been based on cross-sectional survey studies in which communication is measured via a global retrospective report. Although informative, these measures are subject to recall biases and may be colored by a respondent's current state [14]. Ecological momentary approaches afford assessment of experiences and behaviors in naturalistic contexts and in real time, which allow for the assessment of temporal processes [15-19]. Owing to the advances in technology, smartphone apps provide a platform to gather momentary data by assessing thoughts, feelings, and behaviors via notifications to respond to self-report surveys with language-based data [20]. Wheeler and Reis [21] referred to these assessments as "small events" and noted 3 different approaches to the timing of assessment: interval-contingent recording, signal-contingent recording, and event-contingent recording. The focus here is on the signal-contingent recording, a method by which subjects are prompted to report on a recent experience or event when a signal (ie, a smartphone notification) is received, in this case, on a random schedule within a fixed time interval. The advantages of the signal-contingent recording approach are that the report is close in time to the event, thus reducing recall errors and the likelihood of reappraisal. The disadvantages are that the notifications may be intrusive, and rarer events are less likely to be captured [21].

The primary purpose of this pilot project was to assess the usability and acceptability of a smartphone app designed to gather twice-daily reports of communication with a romantic partner as well as mood and closeness and daily reports of relationship satisfaction. The project was distinct from past research in multiple aspects. First, although the assessment of disclosure and responsiveness is not novel [22,23], to the authors' knowledge, no study has assessed holding back from disclosure using a smartphone-based ecological momentary assessment (EMA). Second, a variety of communicative behaviors (eg, enacted and perceived disclosure, holding back, support provision) were captured with regard to the general conversation and not tied to specific concerns as is done in much of the medical literature, for example, in the work designed to capture partner responses to patient pain [24]. Third, most of the previous studies assessing partner communication using EMAs utilized paper-and-pencil diary methods, web-based methods wherein participants were instructed to log in and provide reports at certain times of the day, or electronic devices provided in the study [15,25,26]. In this study, we utilized the advantage of the ubiquity of smartphone ownership in the United States [27] to prompt responding on a device that participants are likely to have with them or close at hand. Thus, participants do not need to learn the mechanics of an unfamiliar device nor do they have to carry a separate device that could be bothersome or unwieldy. Project costs were also reduced.

To capture the constructs of interest, we used the LifeData platform, which is a template-based website that affords easy and economical creation of a smartphone app downloadable on iOS and Android platforms. In this study, the approach to the examination of usability and acceptability was both quantitative and qualitative. App-derived user data were used to examine the percentage of notifications that were responded to and completed. We also assessed the frequency of respondents reporting an interaction with their partner and whether this varied by the time of day. A qualitative analysis of semistructured interview questions (posed 1 week after completion of the 14-day EMA) identified participants'

perceptions of the ease of navigation and convenience of using the app.

Secondarily, we sought to examine the between-person vs within-person variability of key constructs (eg, disclosure, holding back, closeness, mood, and relationship satisfaction). We expected relationship satisfaction to differ between persons and be more stable over time within persons and hence only assessed that construct once per day, in the evening. Daily experience methods are ideal for the examination of within-person processes [15-17]. We assumed that mood would vary within persons based on past research [28,29]. The examination of variability of communication items was exploratory. Greater within-person variability relative to between-person variability would provide support for the utility of examining these variables repeatedly and in real time.

Methods

Participants

All procedures were approved by the Institutional Review Board (IRB) of Arizona State University. Participants were recruited from ResearchMatch, a free and secure registry that matches scientific studies to willing volunteers. Volunteers provide basic demographic and health information and agree to be contacted if they are a match for specific studies. At the time of writing this paper (April 14, 2020), ResearchMatch had 768 active studies, 8298 researchers across 169 institutions, and 146,987 volunteers.

Screening occurred in 2 stages. On the basis of the available demographic data, volunteers who were aged 18 years and above and residing in either North Carolina (NC) or Washington (WA) state were identified, mirroring the recruitment sites for a larger study to follow. The 1172 volunteers who met these criteria were sent an approach message conveying this study's title and purpose, an overview of procedures, and the full inclusion criteria (18 years and above, residing in NC or WA, married or in a committed and cohabiting relationship of at least one year, ability to speak and understand English, and ownership of an iOS or Android smartphone).

Over a 2-week recruitment period, 149 of the 1172 matches responded to the approach message, with 104 conveying a willingness to be contacted and 45 declining further contact. Reasons for the decline were self-perceived ineligibility (n=29), lack of interest (n=8), lack of time (n=4), and no reason (n=4). The remaining 1023 matches did not respond to the approach message within the 2-week recruitment time frame and therefore were not pursued further. Of the willing 104 matches, 2 matches invited their spouse/partner to participate through IRB-approved snowball sampling. Although it was informative to know that these partners were willing to participate given the plans to recruit couples for a larger study, only 1 member of each of these 2 dyads was included in this analysis sample to ensure data independence.

Eligible volunteers responding affirmatively to the ResearchMatch contact message (and the two partner referrals) were contacted by phone to verify eligibility and confirm willingness to download the smartphone app: 2 declined participation, 7 were deemed ineligible via telephone screen, 13 did not have the correct contact information (unreachable), 54 did not respond to phone contact, and 30 were enrolled; 3 participants were excluded from the analyses: 2 as described earlier (one randomly from each of the 2 enrolled couples) and 1 who provided no data. This resulted in an analysis sample of 27 individuals.

Procedures

Following consent, participants were instructed to download a free smartphone app called *RealLife Exp*, designed specifically for the study using LifeData, a web-based app development system. The project manager (second author) guided participants through the download process over phone. Upon download and registration, participants began receiving notifications to complete assessments twice daily for 14 days: once in the afternoon, between 1:00 PM and 2:00 PM, and once in the evening, between 7:30 PM and 8:30 PM (local time). Notifications were set to arrive randomly within these time windows. These time frames were chosen because we assumed that conversations with partners would be less likely to occur in the early morning. The evening time point was seen as not too late but sufficiently late to capture evening/dinner conversations. The time windows to begin each assessment were 45 min in length. Specific items and the constructs they were designed to assess are described in the following sections.

Participants could earn up to US $50 for completing all parts of the study: US $42 for completion of the smartphone-based assessments and US $8 for the follow-up phone interview. If they responded to 80% of the notifications or more, they received the full amount of US $42. If they responded to less than 80% of the notifications, they received US $1.50 per completed notification. The payment was in the form of an Amazon gift card sent via email.

Demographics

An initial assessment included questions to gather demographic characteristics such as age, sex, race, ethnicity, and length of the relationship with the partner.

Communication With Partner

At each assessment, participants were asked whether they had talked to their partner since waking up (afternoon assessment) or since the last set of questions (evening assessment). Those responding "yes" were asked a series of follow-up questions about the conversation to assess their own communicative behavior and perceptions of their partner's communicative behavior. Disclosure and holding back were adapted from the Emotional Disclosure Scale [30]. Disclosure was assessed via a single item, "To what extent did you express your feelings during this conversation?" A parallel item assessed perceived partner disclosure, "To what extent did you feel that your partner expressed his/her feelings?" Holding back was also assessed with a single item, "To what extent did you hold back from expressing your feelings?" Additional items assessed facets of responsiveness: "To what extent did you support your partner?" "To what extent did you understand your partner?" "To what extent did you feel that your partner supported you?" and "To

what extent did you feel that your partner understood you?" All of these items were rated on a 1 (*not at all*) to 5 (*a lot*) scale.

Closeness

Closeness was assessed with a single item, "How close do you feel to your partner right now?" Ratings were made on a 1 (*not at all*) to 5 (*extremely*) scale.

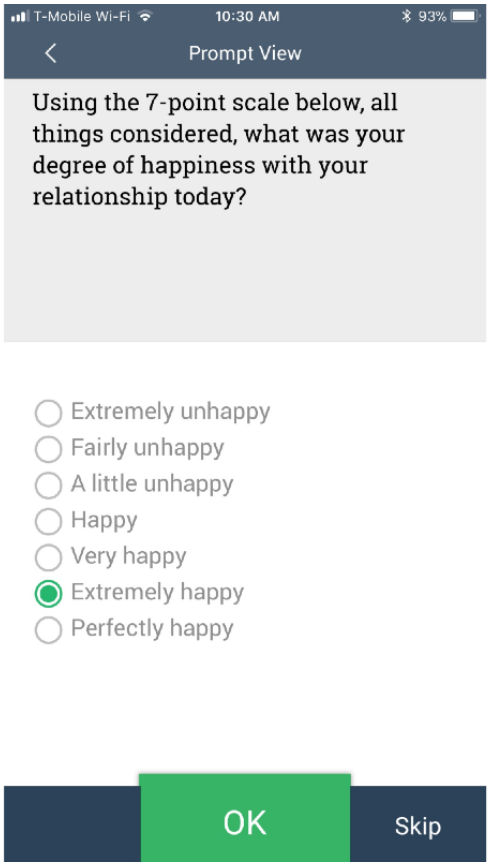
Mood

Mood was measured using an abbreviated version of the Profile of Mood States [31], following Cranford et al [28]. Three items formed each of the 4 subscales: anxious mood (anxious, on edge, and uneasy), depressed mood (sad, hopeless, and discouraged), anger (angry, resentful, and annoyed), and vigor (vigorous, cheerful, and lively). Ratings were made on a 1 (*not at all*) to 5 (*extremely*) scale, and the time referent was "right now." For example, "How on edge do you feel right now?"

Relationship Satisfaction

Relationship satisfaction was assessed with a single item from the Dyadic Adjustment Scale,[32], specifically item 31, following Auger et al [33]. This item was posed only at the evening assessment. Participants were asked, "All things considered, what was your degree of happiness with your relationship today?" Following the standard scale, options ranged from *extremely unhappy* to *perfectly happy*, coded from 1 to 7. See Figure 1 for a screenshot of this question.

Figure 1. Screenshot of the app.



Follow-Up Interview

One week after the end of the 14-day EMA phase, participants were contacted by phone by the second author for a follow-up interview to assess perceptions of the app. Two questions were closed-ended, one to assess ease of use and the other to assess convenience, both indicators of acceptability. The other questions were open-ended, one to assess the convenience of notification timings (another indicator of acceptability) and the other to assess the ease of navigation (an indicator of usability or how well the app functions; Table 1). In posing the open-ended questions, the interviewer probed for clarification as necessary and took detailed notes, including verbatim speech.

Table 1. Measures of usability and acceptability.

Measures	Source	Type of data
Usability (completion and navigation)		
Among notifications sent, the number responded to	App-derived user data	Objective, quantitative
Among notifications sent, the number completed	App-derived user data	Objective, quantitative
How easy was it to navigate within the app? (open-ended)	Follow-up interview	Qualitative
Acceptability (ease and convenience)		
On a scale of 1-5, how easy was it for you to use the app?	Follow-up interview	Subjective, quantitative
On a scale of 1-5, how convenient was it for you to use the app?	Follow-up interview	Subjective, quantitative
How convenient were the notification timings? (open-ended)	Follow-up interview	Qualitative

Analyses

Quantitative

Univariate descriptive statistics were used to summarize the sample's demographic characteristics, EMA response and completion rates, and ratings of the app using SPSS 24.0. Descriptive statistics were also used to characterize the sample with respect to communication, closeness, mood, and relationship satisfaction items. To capture the proportion of the total variance in each item attributable to between-person differences vs within-person (ie, day-to-day) variability, we computed intraclass correlation coefficients (ICCs) from minimum norm quadratic unbiased estimation (MINQUE) estimates of variance components (between-person variance and within-person variance) using the minque package [34] in R. ICCs were computed for afternoon and evening assessments separately. Using Poisson regression models, the association of each background characteristic (measured at baseline) with the response rate (count of responses to prompts) and completion rate (count of completed assessments) was examined separately for afternoon and evening assessments. To test for afternoon vs evening differences in response and completion rates and in the frequency of speaking to one's partner, we estimated single-predictor logistic regression models with bootstrap standard errors adjusted for a within-person clustering using the rms package [35] in R.

Qualitative Review

The second author conducted content analysis of the raw qualitative interview data, which included interviewer notes of responses to the open-ended items listed in Table 1 and direct participant quotations. The practical nature of the topic, relatively small sample size, and ease of capturing and interpreting participant responses to items and probes did not warrant an audio-recording or multiple coders. Methodological rigor was maintained by reviewing all of the detailed notes from each participant multiple times before generating preliminary codes, coding and categorizing identified issues by type and frequency of occurrence, identifying themes and refining codes in an iterative process, and continuing the analysis until no further themes were emerging from the data [36].

Results

Sample Characteristics

Table 2 displays the demographic characteristics of the analysis sample. The average age of participants was 36 (SD 12) years. Most participants identified as female (24/27, 89%), white (25/27, 93%), and non-Hispanic (25/27, 93%). The length of participants' current marriage or partnered relationship varied greatly, with 22% (6/27) reporting relationships of 1 to 2 years and 15% (4/27) reporting being in their current relationships for 16 or more years.

Table 2. Demographic characteristics of the sample (N=27).

Variable[a]	Values
Age (years)	
Mean (SD)	36.41 (11.99)
Range	22-64
Gender, n (%)	
Male	3 (11)
Female	24 (89)
Race, n (%)	
Black or African American	1 (4)
White	25 (93)
Multiracial	1 (4)
Ethnicity, n (%)	
Hispanic or Latino/Latina	2 (7)
Non-Hispanic	25 (93)
Length of relationship in years, n (%)	
1-2	6 (22)
3-5	4 (15)
6-10	7 (26)
11-15	6 (22)
>16	4 (15)

[a]The length of the relationship was assessed categorically.

App-Derived Usability Metrics

Table 3 presents app-derived usability metrics. Among 701 total notifications sent across participants and both afternoon and evening assessments, 555 (79.2%) were responded to and 553 (78.9%) were completed. These values did not differ as a function of the assessment time point (P>.27). In addition, counts of responses to afternoon and evening EMA prompts and counts of afternoon and evening EMA completion rates were unrelated to age, gender, race (dichotomized as white vs black or multiracial), ethnicity, or relationship length (P>.64).

Among the prompts responded to, 79.3% (440/555) were characterized by a report of having conversed with one's partner (either since waking up for the afternoon prompt or since the last assessment for the evening prompt). This rate was higher for the evening vs afternoon time point (86% vs 72%); Wald z from Poisson regression was 2.79 (P=.005).

Table 3. App-derived usability metrics.

Usability metrics	Total	Afternoon	Evening	*P* value
Number of notifications sent, n	701	345	356	N/A[a]
Notifications responded to, n (%)	555 (79.2)	268 (77.7)	287 (80.6)	.28
Assessments completed, n (%)	553 (78.9)	268 (77.7)	285 (80.1)	.37
Conversed with partner since waking up or last notification (prompts responded to), n (%)	440 (79.3)	193 (72.0)	247 (86.1)	.005

[a]N/A: not applicable.

Self-Report Ratings of Acceptability

Of the 27 participants, 26 completed the follow-up interview. Mean ratings of ease of use and convenience of the app fell well above the midpoint of the 1-5 (*not at all* to *extremely*) scale: mean 4.39 (SD 0.70) for "How easy was it for you to use the app?" and mean 4.15 (SD 0.78) for "How convenient was it for you to use the app?"

Content Analysis of Responses to the Open-Ended Interview Questions

In what follows, we describe themes derived from content analysis of the open-ended interview items listed in Table 1. Representative participant quotations are included to illustrate salient findings. Two broad categories emerged: (1) technical functioning and navigation and (2) response convenience, notification timings, and session active window.

Technical Functioning and Navigation

Overall, participants found the app to be very streamlined and user-friendly, perceiving it as a useful tool that functioned without technical difficulty. Notifications arrived as planned and were visible on the home screen of the phone and as an alert on the app icon. Navigation was reported to be simple for the most part, with an intuitive process to move forward:

> *App was very simple...could not be made easier.*
>
> *Just hit OK and move forward, really easy.*
>
> *Very basic and simple, nice to have it on the home screen.*

Specific actions within the navigation process elicited comments. Some found the skip option useful, particularly if the answer was unknown or if a respondent felt uncomfortable sharing the information in question. The "go back" function provoked some frustration. Participants who wanted to review their previous answer but then decided not to change it after all had to reselect the same answer to move forward:

> *If I went back, I had to re-click my answer even if I didn't want to change it.*

There were several indications that the download process (which the project manager instructed the participant to do step-by-step over the telephone) was quite complicated and time-consuming. There was a notable difference between the perceived complexity of the download and the reported simplicity of using the app, suggesting that direct guidance and a user brochure could greatly facilitate this process:

> *Very easy...the setup that you walked me through, that was harder … then it all worked fine.*

Response Convenience, Notification Timings, and Session Active Window

In general, participants indicated that delivery of the assessment via the app was very convenient, insofar as they typically had their smartphone available and usually saw or heard the notification or looked for it within the expected timeframe.

Participants reported it to be much easier if they responded from the phone's home screen, rather than going into the app itself and seeing the alert there and then beginning the survey (if the notification was missed):

> *I liked that you could just swipe the notification and go right into it, much easier than if you missed it and the notification went away, then had to go to the app and see the alert.*

Participants also said that being active in another app could impede response to the notification because some apps do not allow new notifications while open.

Perceptions regarding session timings were largely driven by differences in personal schedules. There was a marked preference for the evening session as most respondents were more able to interrupt their activities at that time. One primary recommendation was to change the timing of the afternoon session:

> *It was really hard for me in the afternoon. I would rather it was coming either mid-morning or at a more traditional lunchtime.*
>
> *I tended to remember the evening one more, so I would check the phone more periodically.*

The 45-min window during which the notification remained active (it expired and was no longer accessible after 45 min) was too short for most of the participants. The primary recommendation was to increase the session active window to at least one hour and preferably to 1.5 hours to accommodate events that last an hour:

> *Biggest issue, expired too quickly!*
>
> *The window was way too short and many times I found it difficult to answer in the time.*

Randomization of the session timings within an hour window was frustrating to many participants. Some reported resorting to setting alarms and "waiting" for the notification to arrive. Irritation with randomization tended to increase over the 14-day activity. The primary issue was not knowing when the notification would arrive and anxiety over "missing" it:

> *It was a little vexing...I set an alarm so I would be ready, but since it always changed the time it was really a little crazy.*

Descriptive Statistics for Key Variables

Table 4 displays means, standard deviations, and ICCs for key constructs, as a function of the time point (afternoon or evening). Levels of enacted and perceived disclosure were relatively high, as were levels of support provision, understanding, and perceived partner support and understanding. Levels of holding back were low. To reiterate, all of these items were rated on a 1-5 (not at all to a lot) scale. Relationship satisfaction, rated on a 1-7 (extremely unhappy to perfectly happy) scale, was moderately high on average (mean 5.20, SD 1.28). Levels of vigor were below the scale midpoint of 3 on average. Levels of anxiety, anger, and depressed affect were below a score of 2 on average.

The ratio of between-person variance to the total variance (where total variance=between-person variance + within-person variance) is reflected by the ICC values listed in the fifth column of Table 4. As shown in Table 4, ICCs for EMA measures were generally <0.40, indicating that the majority of the variability in each measure was at the within-person level, rather than at the between-person level, suggesting that there was meaningful day-to-day variation in these variables. Notable exceptions were enacted disclosure (ICCs=0.46 and 0.42 for afternoon and evening assessments, respectively), closeness (ICCs=0.41 and 0.40, respectively), and relationship satisfaction (ICC=0.59), indicating that between-person differences in these variables were relatively more stable across the 14-day period.

Table 4. Means, standard deviations, and intraclass correlation coefficients for smartphone-assessed constructs.

Variable and time of day[a]	N	Value, mean (SD)	ICC[b]
To what extent did you express your feelings?			
Afternoon	193	3.67 (1.22)	0.46
Evening	245	3.64 (1.17)	0.42
To what extent did you feel that your partner expressed his/her feelings?			
Afternoon	189	3.74 (1.16)	0.18
Evening	242	3.87 (1.08)	0.24
To what extent did you hold back from expressing your feelings?			
Afternoon	193	1.62 (1.04)	0.23
Evening	246	1.77 (1.06)	0.37
To what extent did you support your partner?			
Afternoon	189	3.78 (1.25)	0.33
Evening	243	3.91 (1.09)	0.38
To what extent did you understand your partner?			
Afternoon	189	3.90 (1.09)	0.25
Evening	241	3.91 (1.02)	0.29
To what extent did you feel that your partner supported you?			
Afternoon	190	3.85 (1.22)	0.19
Evening	242	3.81 (1.08)	0.31
To what extent did you feel that your partner understood you?			
Afternoon	190	3.70 (1.20)	0.14
Evening	241	3.76 (1.06)	0.17
How close do you feel to your partner right now?			
Afternoon	267	4.04 (0.93)	0.41
Evening	283	4.04 (0.91)	0.40
POMS[c] vigor subscale			
Afternoon	268	2.82 (0.85)	0.22
Evening	284	2.59 (0.81)	0.25
POMS anxiety subscale			
Afternoon	268	1.76 (0.90)	0.30
Evening	284	1.66 (0.81)	0.32
POMS anger subscale			
Afternoon	268	1.48 (0.81)	0.27
Evening	284	1.50 (0.75)	0.26
POMS depressed affect subscale			
Afternoon	268	1.60 (0.84)	0.32
Evening	284	1.52 (0.74)	0.33
Relationship satisfaction			
Evening	285	5.20 (1.28)	0.59

[a]All items were rated on a 1-5 scale except for relationship satisfaction which was rated on a 1-7 scale.

[b]ICC: intraclass correlation coefficient.

[c]POMS: Profile of Mood States.

Discussion

Principal Findings

The primary goal of this pilot project was to examine the usability and acceptability of a smartphone app designed to assess communication with a romantic partner, closeness, mood, and relationship satisfaction repeatedly over the course of 14 days. The app was rated as easy to navigate, and the response rate was quite good. Of the 701 total notifications sent, 555 (79.2%) were responded to and 553 (78.9%) were completed. Comparing these rates with those reported in the literature is challenging, given the wide variability in the frequency of prompts, number and content of items posed, and sample characteristics. Incentive structures also likely vary. However, in general, the completion rates fell within the ranges reported by other research teams [37], in some cases higher by 8% to 14% [38-40] and in other cases lower by 4% to 7% [41]. These differences may, in part, be explained by differences in numbers of items, for example, the battery was somewhat longer than that described by Perndorfer et al [41].

With regard to acceptability, the app was rated as convenient to use on average (mean 4.15 on a 1-5 scale). However, qualitative analyses provided a more nuanced understanding of the perceived acceptability of the app. Participants expressed difficulty with 3 aspects related to timing: (1) The afternoon prompt came between 1:00 PM and 2:00 PM, which may have been difficult for employed participants due to work-related demands. (2) Notification times (signals to respond) were randomized within a 1-hour period. Inability to anticipate the notification's arrival was frustrating. (3) The active time window to begin each assessment was 45 min, which participants felt was too short. On the basis of this feedback, modifications were made to the app for a larger ongoing study of couples coping with cancer. In the ongoing study, notifications are delivered at fixed times (at noon and 8:00 PM), and the time window to complete assessments is 2 hours. We also added reminders (a LifeData feature not available at the time of the pilot) that arrive every 20 min within the open window. Further modifications were made in response to the technical navigation issues raised. For example, the app user guide was refined to clearly describe how and when to use the skip and go back functions.

One possible drawback of signal-contingent recording is that infrequent behaviors might not be captured [21]. Indeed, we did not know before conducting this study whether conversations with a spouse or partner would occur during the periods in question. Findings suggest that partner conversations are sufficiently frequent to warrant an assessment of the occurrence and nature of those conversations using EMA methods. Among the prompts responded to, 79% were characterized by a report of having conversed with one's partner (either since waking up for the afternoon prompt or since the last assessment for the evening prompt). This value was significantly higher for evening vs afternoon reports. Most participants were likely away from home at work during the day time, or if at home, may have been engaged in activities apart from their partner, making daytime conversations less likely. However, this is a conjecture, as we did not formally assess the employment status (though it was mentioned by some participants in the interview).

When conversations did occur, they were characterized, on average, by moderately high levels of disclosure, both enacted and perceived. Participants also saw themselves, in general, as being supportive and understanding, and in turn as receiving support and understanding. Holding back was less likely to occur. While exact comparisons to other reports in the literature are difficult to draw given inconsistencies in the communicative behaviors measured and, in some cases, the use of different rating scales, the relative frequency of the behaviors is generally in line with reports derived from traditional questionnaire measures of communication. For example, Porter et al [42] observed moderately high levels of disclosure and low levels of holding back among patients with gastrointestinal cancer and their spouses.

A secondary goal of this study was to determine between-person variability vs within-person variability of the study variables. ICCs underscore the utility of the EMA approach for the measurement of the communication items, all indicating greater within-person variability than between-person variability. Disclosure showed lower within-person variability than other communication measures, perhaps reflecting a dispositional tendency to express emotion across time and situations. Similar to the majority of communication behaviors, measures of anger, anxiety, depressed mood, and vigor showed considerable day-to-day variability within persons. Closeness and relationship satisfaction showed greater stability over the 14-day period. The relationship satisfaction finding is consistent with that reported by Gadassi et al [26] who administered the same single item. These findings are also in line with this study's expectation that this construct would vary less over time within persons than between persons.

Limitations

Limitations of this pilot study must be considered. By design, the number of participants was small. The sample was comprised largely of non-Hispanic white women, limiting the generalizability of the results. As women tend to be more emotionally expressive than men, [43], this could explain the fairly high levels of enacted disclosure and low levels of holding back. The recruitment source, ResearchMatch, also limits generalizability. This website matches researchers to willing volunteers, that is, persons open to the idea of research and perhaps motivated to earn incentives for participation. A published analysis of the ResearchMatch volunteer database (N=15,871) indicated that 81% of volunteers identified as white and 95% as non-Hispanic [44]. The average age was 38 years, and most volunteers (73%) were female. The demographic composition of this small sample mirrors this larger pool.

It is also important to note that this study, while focused on partner communication, was not dyadic in nature. This approach was chosen to hasten recruitment and based on the assumption that usability and acceptability data from one member of a dyad would be sufficient to inform the next steps for a larger study with dyads. Relatedly, partner characteristics were not assessed nor were participants asked to report on partner characteristics including demographic characteristics. Therefore, it is not known

whether members of each couple were of the same or different sex/gender. We also relied entirely on participant self-reports of their own behavior and reports of their partner's behavior. Thus, it is not possible to examine concordance between self- and partner reports of communicative behavior.

As described in the Introduction, much of the research on couple communication has been designed to assess communication in reference to a specific concern or topic. This is often the case in laboratory-based studies wherein couples are asked to discuss either a relevant shared stressor or a conflictual topic. It has also been the case in numerous questionnaire-based assessments of holding back in which couples are asked to rate the extent to which they (1) disclosed and (2) held back from disclosing a number of different illness-related concerns [45-49]. In this study, the participants were not recruited based on a common stressor or illness. Conversations were not constrained to a specific topic nor were the participants asked to report on the topic. Therefore, we do not know what was discussed nor do we have a sense of the valence of each conversation. These contextual variables may be important to examine as moderators in future research. Perceived lack of responsiveness from one's partner, for example, may be more deleterious in the context of a highly stressful topic, such as serious illness or relationship distress, vs in the context of daily hassles.

Despite the study's limitations, findings from this pilot project lend support for the use of smartphone apps to assess communication in real time and in naturalistic settings. They also underscore the advantages of using a web-based template for app creation, a highly affordable option as opposed to hiring a programmer or developer. On the basis of the usability data and feedback from participants, this smartphone app has been since adapted for use with a larger sample of patients with cancer and their cohabiting partners/spouses. The larger study is still in process but initial results suggest strong completion rates and acceptability of the app. Future interventions designed to train couples in adaptive communication could potentially make use of EMA data such as these to inform targeted approaches and to monitor the response to the intervention.

Acknowledgments

This study was funded by R01 CA201179 (multiple principal investigators: SLL and LSP).

Conflicts of Interest

None declared.

References

1. Vangelisti AL. Interpersonal processes in romantic relationships. In: Knapp ML, Daly JA, editors. The SAGE Handbook of Interpersonal Communication. Thousand Oaks, CA: SAGE Publications; 2011:597-631.
2. Miller R. Intimate Relationships. Eighth Edition. New York, USA: McGraw Hill; 2018.
3. Reis H, Shaver P. Intimacy as an interpersonal process. In: Duck S, editor. Handbook of Personal Relationships: Theory, Research and Interventions. Chichester, England: Wiley; 1988:367-389.
4. Laurenceau J, Barrett LF, Rovine MJ. The interpersonal process model of intimacy in marriage: a daily-diary and multilevel modeling approach. J Fam Psychol 2005 Jun;19(2):314-323. [doi: 10.1037/0893-3200.19.2.314] [Medline: 15982109]
5. Manne S, Badr H. Intimacy and relationship processes in couples' psychosocial adaptation to cancer. Cancer 2008 Jun 1;112(11 Suppl):2541-2555 [FREE Full text] [doi: 10.1002/cncr.23450] [Medline: 18428202]
6. Manne SL, Ostroff JS, Norton TR, Fox K, Goldstein L, Grana G. Cancer-related relationship communication in couples coping with early stage breast cancer. Psychooncology 2006 Mar;15(3):234-247. [doi: 10.1002/pon.941] [Medline: 15926198]
7. Hagedoorn M, Dagan M, Puterman E, Hoff C, Meijerink WJ, Delongis A, et al. Relationship satisfaction in couples confronted with colorectal cancer: the interplay of past and current spousal support. J Behav Med 2011 Aug;34(4):288-297 [FREE Full text] [doi: 10.1007/s10865-010-9311-7] [Medline: 21222025]
8. Langer SL, Brown JD, Syrjala KL. Intrapersonal and interpersonal consequences of protective buffering among cancer patients and caregivers. Cancer 2009 Sep 15;115(18 Suppl):4311-4325 [FREE Full text] [doi: 10.1002/cncr.24586] [Medline: 19731352]
9. Langer SL, Romano JM, Todd M, Strauman TJ, Keefe FJ, Syrjala KL, et al. Links between communication and relationship satisfaction among patients with cancer and their spouses: results of a fourteen-day smartphone-based ecological momentary assessment study. Front Psychol 2018;9:1843 [FREE Full text] [doi: 10.3389/fpsyg.2018.01843] [Medline: 30364167]
10. Manne S, Badr H, Zaider T, Nelson C, Kissane D. Cancer-related communication, relationship intimacy, and psychological distress among couples coping with localized prostate cancer. J Cancer Surviv 2010 Mar;4(1):74-85 [FREE Full text] [doi: 10.1007/s11764-009-0109-y] [Medline: 19967408]
11. Manne S, Dougherty J, Veach S, Kless R. Hiding worries from one's spouse: protective buffering among cancer patients and their spouses. Cancer Res Ther Cont 1999;8(1-2):175-188 [FREE Full text]
12. Traa MJ, de Vries J, Bodenmann G, Den Oudsten BL. Dyadic coping and relationship functioning in couples coping with cancer: a systematic review. Br J Health Psychol 2015 Feb;20(1):85-114. [doi: 10.1111/bjhp.12094] [Medline: 24628822]
13. Winterheld HA. Hiding feelings for whose sake? Attachment avoidance, relationship connectedness, and protective buffering intentions. Emotion 2017 Sep;17(6):965-980. [doi: 10.1037/emo0000291] [Medline: 28277711]

14. Shiffman S, Stone AA. Introduction to the special section: ecological momentary assessment in health psychology. Health Psychol 1998 Jan;17(1):3-5 [FREE Full text] [doi: 10.1037/h0092706]
15. Bolger N, Davis A, Rafaeli E. Diary methods: capturing life as it is lived. Annu Rev Psychol 2003;54:579-616. [doi: 10.1146/annurev.psych.54.101601.145030] [Medline: 12499517]
16. Bolger N, Laurenceau J. Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research. New York, USA: The Guilford Press; 2013.
17. Gable S, Gosnell C, Prok T. Close relationships. In: Mehl M, Conner T, editors. Handbook of Research Methods for Studying Daily Life. New York, USA: The Guilford Press; 2011:511-524.
18. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
19. Smyth JM, Stone AA. Ecological momentary assessment research in behavioral medicine. J Happiness Stud 2003;4(1):35-52 [FREE Full text] [doi: 10.1023/a:1023657221954]
20. Harari GM, Müller SR, Aung MS, Rentfrow PJ. Smartphone sensing methods for studying behavior in everyday life. Curr Opin Behav Sci 2017 Dec;18:83-90 [FREE Full text] [doi: 10.1016/j.cobeha.2017.07.018]
21. Wheeler L, Reis H. Self-recording of everyday life events: origins, types, and uses. J Pers 1991;59(3):339-354 [FREE Full text] [doi: 10.1111/j.1467-6494.1991.tb00252.x]
22. Pagani AF, Donato S, Parise M, Bertoni A, Iafrate R, Schoebi D. Explicit stress communication facilitates perceived responsiveness in dyadic coping. Front Psychol 2019;10:401. [doi: 10.3389/fpsyg.2019.00401] [Medline: 30873090]
23. Visserman ML, Righetti F, Impett EA, Keltner D, van Lange PA. It's the motive that counts: perceived sacrifice motives and gratitude in romantic relationships. Emotion 2018 Aug;18(5):625-637. [doi: 10.1037/emo0000344] [Medline: 28604038]
24. Wilson SJ, Martire LM, Sliwinski MJ. Daily spousal responsiveness predicts longer-term trajectories of patients' physical function. Psychol Sci 2017 Jun;28(6):786-797 [FREE Full text] [doi: 10.1177/0956797617697444] [Medline: 28459650]
25. Badr H, Laurenceau J, Schart L, Basen-Engquist K, Turk D. The daily impact of pain from metastatic breast cancer on spousal relationships: a dyadic electronic diary study. Pain 2010 Dec;151(3):644-654. [doi: 10.1016/j.pain.2010.08.022] [Medline: 20833474]
26. Gadassi R, Bar-Nahum LE, Newhouse S, Anderson R, Heiman JR, Rafaeli E, et al. Perceived partner responsiveness mediates the association between sexual and marital satisfaction: a daily diary study in newlywed couples. Arch Sex Behav 2016 Jan;45(1):109-120. [doi: 10.1007/s10508-014-0448-2] [Medline: 25680818]
27. Mobile Fact Sheet. Pew Research Center. 2019. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2019-03-21] [WebCite Cache ID 772sLkKHs]
28. Cranford JA, Shrout PE, Iida M, Rafaeli E, Yip T, Bolger N. A procedure for evaluating sensitivity to within-person change: can mood measures in diary studies detect change reliably? Pers Soc Psychol Bull 2006 Jul;32(7):917-929 [FREE Full text] [doi: 10.1177/0146167206287721] [Medline: 16738025]
29. Iida M, Shrout P, Laurenceau J, Bolger N. Using diary methods in psychological research. In: Cooper H, Camic P, Long D, Panter A, Rindskopf D, Sher K, editors. APA Handbook of Research Methods in Psychology, Volume 1: Foundations, Planning, Measures, and Psychometrics. Washington, DC: American Psychological Association; 2012:277-305.
30. Pistrang N, Barker C. The partner relationship in psychological response to breast cancer. Soc Sci Med 1995 Mar;40(6):789-797. [doi: 10.1016/0277-9536(94)00136-h] [Medline: 7747213]
31. McNair DM, Lorr M, Droppleman LF. Edits Manual for the Profile of Mood States (POMS). San Diego, CA: Educational and Industrial Testing Services; 1992.
32. Spanier GB. Measuring dyadic adjustment: new scales for assessing the quality of marriage and similar dyads. J Marriage Fam 1976 Feb;38(1):15-28 [FREE Full text] [doi: 10.2307/350547]
33. Auger E, Menzies-Toman D, Lydon JE. Daily experiences and relationship well-being: the paradoxical effects of relationship identification. J Pers 2017 Oct;85(5):741-752. [doi: 10.1111/jopy.12283] [Medline: 27589212]
34. Wu J. minque: Various Linear Mixed Model Analyses. R Documentation. 2019. URL: https://rdrr.io/cran/minque/ [accessed 2020-06-08]
35. Harrell F. rms: Regression Modeling Strategies. The Comprehensive R Archive Network. 2019. URL: https://CRAN.R-project.org/package=rms [accessed 2020-06-08]
36. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2006;5(1):80-92 [FREE Full text] [doi: 10.1177/160940690600500107]
37. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, et al. Utilizing a personal smartphone custom app to assess the patient health questionnaire-9 (PHQ-9) depressive symptoms in patients with major depressive disorder. JMIR Ment Health 2015;2(1):e8 [FREE Full text] [doi: 10.2196/mental.3889] [Medline: 26543914]
38. Badr H, Pasipanodya EC, Laurenceau J. An electronic diary study of the effects of patient avoidance and partner social constraints on patient momentary affect in metastatic breast cancer. Ann Behav Med 2013 Apr;45(2):192-202. [doi: 10.1007/s12160-012-9436-8] [Medline: 23150234]
39. Gidlow CJ, Randall J, Gillman J, Silk S, Jones MV. Hair cortisol and self-reported stress in healthy, working adults. Psychoneuroendocrinology 2016 Jan;63:163-169. [doi: 10.1016/j.psyneuen.2015.09.022] [Medline: 26447679]

40. Intille S, Haynes C, Maniar D, Ponnada A, Manjourides J. MEMA: microinteraction-based ecological momentary assessment (EMA) using a smartwatch. Proc ACM Int Conf Ubiquitous Comput 2016 Sep;2016:1124-1128 [FREE Full text] [doi: 10.1145/2971648.2971717] [Medline: 30238088]
41. Perndorfer C, Soriano EC, Siegel SD, Laurenceau J. Everyday protective buffering predicts intimacy and fear of cancer recurrence in couples coping with early-stage breast cancer. Psychooncology 2019 Feb;28(2):317-323 [FREE Full text] [doi: 10.1002/pon.4942] [Medline: 30426612]
42. Porter LS, Keefe FJ, Hurwitz H, Faber M. Disclosure between patients with gastrointestinal cancer and their spouses. Psychooncology 2005 Dec;14(12):1030-1042. [doi: 10.1002/pon.915] [Medline: 15712247]
43. Fischer A, LaFrance M. What drives the smile and the tear: why women are more emotionally expressive than men. Emotion Rev 2014 Dec 5;7(1):22-29 [FREE Full text] [doi: 10.1177/1754073914544406]
44. Harris PA, Scott KW, Lebo L, Hassan N, Lightner C, Pulley J. ResearchMatch: a national registry to recruit volunteers for clinical research. Acad Med 2012 Jan;87(1):66-73 [FREE Full text] [doi: 10.1097/ACM.0b013e31823ab7d2] [Medline: 22104055]
45. Manne S, Kashy D, Kissane D, Ozga M, Virtue S, Heckman C. The course and predictors of perceived unsupportive responses by family and friends among women newly diagnosed with gynecological cancers. Transl Behav Med 2019 Jul 16;9(4):682-692 [FREE Full text] [doi: 10.1093/tbm/iby087] [Medline: 30189025]
46. Manne SL, Kashy DA, Kissane DW, Ozga M, Virtue SM, Heckman CJ. Longitudinal course and predictors of communication and affect management self-efficacy among women newly diagnosed with gynecological cancers. Support Care Cancer 2020 Apr;28(4):1929-1939. [doi: 10.1007/s00520-019-04989-6] [Medline: 31367918]
47. Manne S, Kashy DA, Siegel S, Virtue SM, Heckman C, Ryan D. Unsupportive partner behaviors, social-cognitive processing, and psychological outcomes in couples coping with early stage breast cancer. J Fam Psychol 2014 Apr;28(2):214-224 [FREE Full text] [doi: 10.1037/a0036053] [Medline: 24611691]
48. Oh S, Ryu E. Does holding back cancer-related concern affect couples' marital relationship and quality of life of patients with lung cancer? An actor-partner interdependence mediation modeling approach. Asian Nurs Res (Korean Soc Nurs Sci) 2019 Oct;13(4):277-285 [FREE Full text] [doi: 10.1016/j.anr.2019.10.001] [Medline: 31605768]
49. Zhaoyang R, Martire LM, Stanford AM. Disclosure and holding back: communication, psychological adjustment, and marital satisfaction among couples coping with osteoarthritis. J Fam Psychol 2018 Apr;32(3):412-418 [FREE Full text] [doi: 10.1037/fam0000390] [Medline: 29698013]

Abbreviations

EMA: ecological momentary assessment
ICC: intraclass correlation coefficient
IRB: Institutional Review Board

Edited by G Eysenbach; submitted 28.03.19; peer-reviewed by R Zhaoyang, S Berrouiguet, D Fulford; comments to author 16.01.20; accepted 14.05.20; published 06.07.20.

Please cite as:
Langer SL, Ghosh N, Todd M, Randall AK, Romano JM, Bricker JB, Bolger N, Burns JW, Hagan RC, Porter LS
Usability and Acceptability of a Smartphone App to Assess Partner Communication, Closeness, Mood, and Relationship Satisfaction: Mixed Methods Study
JMIR Form Res 2020;4(7):e14161
URL: https://formative.jmir.org/2020/7/e14161
doi:10.2196/14161
PMID:32628614

Original Paper

Self-Administered Skills-Based Virtual Reality Intervention for Chronic Pain: Randomized Controlled Pilot Study

Beth D Darnall[1], PhD; Parthasarathy Krishnamurthy[2], MBA, PhD; Jeannette Tsuei[3], MA; Jorge D Minor[4], MD

[1]Department of Anesthesiology, Perioperative and Pain Medicine, Stanford University School of Medicine, Palo Alto, CA, United States

[2]C.T. Bauer College of Business, University of Houston, Houston, TX, United States

[3]AppliedVR, Inc, Los Angeles, CA, United States

[4]L.A. Pain & Wellness Institute, Los Angeles, CA, United States

Corresponding Author:
Beth D Darnall, PhD
Department of Anesthesiology, Perioperative and Pain Medicine
Stanford University School of Medicine
1070 Arastradero Road
Suite 200, MC 5596
Palo Alto, CA, 94304
United States
Phone: 1 (650) 725 9642
Email: bdarnall@stanford.edu

Abstract

Background: Patients with chronic pain often have limited access to comprehensive care that includes behavioral pain management strategies. Virtual reality (VR) is an immersive technology and emerging digital behavioral pain therapy with analgesic efficacy for acute pain. We found no scientific literature on skills-based VR behavioral programs for chronic pain populations.

Objective: The primary aim of this study is to evaluate the feasibility of a self-administered VR program that included content and skills informed by evidence-based behavioral treatment for chronic pain. The secondary aim is to determine the preliminary efficacy of the VR program in terms of average pain intensity and pain-related interference with activity, stress, mood, and sleep, and its impact on pain-related cognition and self-efficacy. The tertiary aim was to conduct a randomized controlled trial (RCT) and compare the VR treatment with an audio-only treatment. This comparison isolated the immersive effects of the VR program, thereby informing potential mechanisms of effect.

Methods: We conducted an RCT involving a web-based convenience sample of adults (N=97) aged 18-75 years with self-reported chronic nonmalignant low back pain or fibromyalgia, with an average pain intensity >4 over the past month and chronic pain duration >6 months. Enrolled participants were randomly assigned to 1 of 2 unblinded treatments: (1) VR: a 21-day, skills-based VR program for chronic pain; and (2) audio: an audio-only version of the 21-day VR program. The analytic data set included participants who completed at least 1 of 8 surveys administered during the intervention period: VR (n=39) and audio (n=35).

Results: The VR and audio groups launched a total of 1067 and 1048 sessions, respectively. The majority of VR participants (n=19/25, 76%) reported no nausea or motion sickness. High satisfaction ratings were reported for VR (n=24/29, 83%) and audio (n=26/33, 72%). For VR efficacy, symptom improvement over time was found for each pain variable (all P<.001), with results strengthening after 2 weeks. Importantly, significant time×group effects were found in favor of the VR group for average pain intensity (P=.04), pain-related inference with activity (P=.005), sleep (P<.001), mood (P<.001), and stress (P=.003). For pain catastrophizing and pain self-efficacy, we found a significant declining trend for both treatment groups.

Conclusions: High engagement and satisfaction combined with low levels of adverse effects support the feasibility and acceptability of at-home skills-based VR for chronic pain. A significant reduction in pain outcomes over the course of the 21-day treatment both within the VR group and compared with an audio-only version suggests that VR has the potential to provide enhanced treatment and greater improvement across a range of pain outcomes. These findings provide a foundation for future research on VR behavioral interventions for chronic pain.

(JMIR Form Res 2020;4(7):e17293) doi:10.2196/17293

KEYWORDS

chronic pain; virtual reality; behavioral medicine; self-management; mobile phone; randomized controlled trial

Introduction

Background

The US Department of Health and Human Services [1,2], the Institute of Medicine [3], the National Institutes of Health, and the Centers for Disease Control and Prevention have called for the integration of evidence-based behavioral medicine strategies into the treatment of acute and chronic pain. These unified recommendations reflect a common understanding that pain is a biopsychosocial experience that requires a comprehensive treatment approach to equip individuals with skills to actively self-manage pain and symptoms [4].

Evidence-based behavioral pain treatments include cognitive behavioral therapy (CBT) for chronic pain [5-7], acceptance and commitment therapy [8], and mindfulness-based stress reduction [5,6]. Treatments typically include didactic content, goal setting, and experiential skills practice within sessions and skills practice between sessions. Effective chronic pain management techniques include biofeedback and relaxation training, with the latter being the mainstay of every effective behavioral treatment for chronic pain [9]. Evidence-based behavioral medicine treatments are commonly delivered individually or in group format, during treatment classes or sessions that are 1 to 2 hours in duration, with a course of treatment lasting 8-11 weeks [7]. Behavioral treatments for chronic pain have been shown to be effective for reducing pain bothersomeness [5,6], symptoms of depression [7,8], and pain-specific cognitive and emotional distress (eg, pain catastrophizing) [5-7], although these treatments have not been effective in reducing pain intensity. Behavioral treatments have also been shown to be effective for improving pain-related self-efficacy or confidence in one's ability to self-manage pain and engage in meaningful activities [5,6].

Despite the availability of efficacy data for behavioral treatments, multiple barriers may prevent patients from broadly accessing low-risk, integrative, nonpharmacologic pain self-management tools that address the psychosocial aspects of chronic pain [10]. Such barriers may include few skilled local therapists, poor insurance coverage, copayments associated with clinic visits, travel costs, and treatment time [10]. Even when delivered to participants at no cost, in-person behavioral medicine treatments can have poor patient engagement [11], thereby suggesting that new methods of treatment delivery are required to meet the needs of a broad range of patients. Accordingly, research has demonstrated preliminary efficacy for an ultrabrief, single-session, skills-based behavioral treatment class for chronic pain [12] as well as for mobile health teleconference-delivered multisession behavioral pain treatment [13]. Although both options offer new conveniences and possibly expanded access to care, particularly for patients with mobility limitations, the patient remains dependent on a therapist for the delivery of the treatment.

Digital therapeutics offer independent, home-based, on-demand access to behavioral treatment for acute and chronic pain. For instance, a brief, web-based, 90-min, skilled-based pain treatment was shown to reduce postsurgical opioid use in women who underwent surgery for breast cancer [14]. For chronic pain, digital multisession interventions have been shown to be effective [15-17]. Despite these successes, any one modality will not meet the needs of everyone; indeed, even web-based pain treatments that are offered at no cost have <60% engagement rates [14], thus underscoring the need to offer a diverse range of accessible treatment options for chronic pain and, in particular, to identify treatments that may yield superior outcomes.

Virtual reality (VR) has emerged as an effective digital treatment for acute pain. With VR, the user wears a headset that fully restricts the vision field to the content displayed inside the headset screen, and auditory perception is directed to the audio delivered through the device (although not fully restricted; Figure 1). VR provides a multisensory, 3D immersive environment that stimulates the visual, auditory, and proprioception senses, thus engendering the perceptual experience that one is physically located inside and engaged with the virtual environment they are viewing [18,19], such as swimming with dolphins (Figure 2). VR has been used in numerous clinical settings and health conditions to treat anxiety and mental disorders [20-22], aids physical rehabilitation [23,24], and supports postsurgical recovery. Evidence suggests that VR is effective for managing acute pain, including pain elicited during medical procedures [25-29], and burn wound care [30,31] and in hospitalized patients [32,33].

Although several studies have investigated VR for chronic pain, the literature to date is limited. Promising studies have used VR to reduce pain and improve outcomes in complex regional pain syndrome [34], chronic headache/migraine pain [35], fibromyalgia [36,37], and chronic musculoskeletal pain [38,39]. Two recent reviews and meta-analyses reported the efficacy of VR for physical rehabilitation from painful spinal conditions [24] and for orthopedic rehabilitation in terms of reduced pain and disability [40]. In such studies, the user may engage with interactive VR alone or within the context of kinematic training. The literature to date is limited by studies that are conducted in experimental or clinical settings, do not compare to a non-VR group, or include very small samples. Most importantly, the VR studies to date are focused primarily on distraction or physical rehabilitation and do not include pain management education or cultivation of behavioral pain self-management skills (eg, diaphragmatic breathing or cognition and emotion regulation). Crucially, if found to be effectively delivered by VR, such content could serve as either a replacement or treatment extender for in-person behavioral medicine clinic visits. Studies are needed to determine whether behavioral pain management skills-based VR is effective and can facilitate sustained pain relief through skills mastery and increased self-efficacy for pain self-management. The VR platform could transcend many current barriers to care and potentially provide a scalable way to deliver on-demand home-based behavioral treatment for chronic pain.

Figure 1. Oculus Go virtual reality headset.

Figure 2. Image of Pain Care virtual reality content (swimming with dolphins).

Objectives

Building on the nascent literature on VR for chronic pain, we aim to evaluate a skills-based behavioral medicine VR program for chronic low back pain and fibromyalgia to provide preliminary data on its utility and efficacy as a stand-alone home-based treatment for people in the community. To accomplish this goal, we conducted an exploratory randomized controlled trial (RCT) with 3 main aims: (1) to determine the feasibility and satisfaction of a self-administered, at-home, skills-based 21-day VR intervention (pain care VR) for chronic pain; (2) to evaluate the preliminary efficacy of VR intervention for reducing average pain intensity and pain-related outcomes; and (3) to isolate the immersive effects of VR by comparing it with an audio-only treatment group. We hypothesized good feasibility and satisfaction for VR as well as the superiority of VR over audio treatment for pre-post improvement across the pain indicator variables.

Methods

Design and Setting

This was a parallel-group, randomized controlled clinical trial involving 2 home-based behavioral interventions applied in a community-based, web-based convenience sample of people with chronic pain conducted between March and May 2019.

Procedures

Participants were recruited remotely through web-based advertisements on Facebook and The Mighty, a digital health community. Internet and computer literacy were implicit de facto eligibility criteria. Eligibility screening involved a brief telephone assessment for enrollment criteria outlined in the Inclusion and Exclusion Criteria section.

Inclusion and Exclusion Criteria

Following electronic informed consent (see Multimedia Appendix 1 for the study consent form), study participants were randomized one-to-one using a Research Electronic Data Capture Cloud random number generator and allocated to the treatment group. All study procedures were completed remotely, and no in-person visits were required. Study participants were not blinded to the treatment group assignment because of the obvious nature of the mode of delivery of their assigned treatment. Participants assigned to the VR group were called by a study staff to ensure receipt of the mailed VR kit and for a brief orientation to materials. A study staff member was available by phone at the request of study participants in both groups. Participant compensation was prorated based on the number of surveys completed; they received up to US $30 in the form of an Amazon electronic gift card following completion of their final study survey. The study was approved by the Western institutional review board (Puyallup, WA).

Treatment Groups

Both treatment groups received the same didactic content delivered in distinct formats (VR versus audio; see Multimedia Appendix 2 for program content). Treatment consisted of a variety of sessions to support participants in learning self-management skills based on evidence-based CBT principles as well as biofeedback and mindfulness strategies used in pain management. The program was designed to improve self-regulation of cognitive, emotional, and physiological responses to stress and pain and comprised 3 main content categories:

- Skills rooted in pain CBT: brief didactics describe how thoughts and emotions can impact pain and techniques on thought restructuring and adaptive regulation of pain-related cognition and emotion (eg, addressing pain catastrophizing tendencies).
- Relaxation training: participants engage in diaphragmatic breathing exercises to enhance parasympathetic nervous system function and decrease physiological hyperarousal. Importantly, relaxation training was optimized in the VR group with visual biofeedback that displayed the environment responding to the users' physiologic behavior during the exercises (Figure 3). The software uses a patent-pending algorithm to detect the user's exhalations with the VR headset's built-in microphone, along with a hardware breath amplifier that directs the user's breath toward the microphone. The user's breath is then visualized in the VR environment as either breath particles or waves expanding outward from the user, with the synchronization of physical exhalation and the visual effects deepening the immersion of the experience and helping the user slow and deepen their respiration.
- Mindfulness: mindfulness content encouraged awareness of the mind and body (somatic cues) as well as thought release (nonattachment).

Figure 3. Image of Pain Care virtual reality content that visualizes the user's breath.

The 21-day program consisted of 4-8 treatment sessions from each content category with the duration of session length ranging from 1 to 15 min. Each treatment session was indexed as complete if the participants initiated the experience. Participants could replay the completed sessions.

Virtual Reality Group

VR group participants were mailed an Oculus Go VR headset preloaded with VR software developed by AppliedVR. The Oculus Go is an easy-to-use, stand-alone VR headset with an accompanying orientation-tracked controller, a head-mounted display with a high-resolution screen, built-in spatial audio speakers, and an integrated microphone. A VR user manual sent via email included an instructional video on the proper use of the VR headset. Participants were instructed to telephone VR support staff if they had any questions or difficulty in using the headset. Participants were instructed to complete 1 VR treatment session daily for 21 days (Multimedia Appendix 2).

Audio Group

The audio program consisted of the majority of the same narrative content contained in the VR program, with changes made to the descriptive titles for each session. Owing to VR having a visual media form and audio that specifically references the changing images in the VR environment, approximately one-third of the VR program could not be included verbatim in the audio. Rather, the audio session topical content was closely matched to the corresponding VR session for that day and adapted to eliminate any references to visual content that would be confusing to the listener. Participants received an electronic link to the audio recordings on SoundCloud (a music streaming platform) where they could choose to stream or download the audio recordings on their smartphones, laptops, or desktop computers. Participants were instructed to complete 1 audio treatment session daily for 21 days (Multimedia Appendix 2).

Data Collection and Variable Measurement

Data collection consisted of electronically collected patient-reported measures and objective use data collected from the VR devices and audio access logs.

In accordance with the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) recommendations, we included multiple methods to evaluate the importance of change in outcome measures across 4 recommended domains: pain intensity, health-related quality of life as defined by physical functioning, emotional functioning, and ratings of overall improvement [41,42].

Data Collection Time Points

Data were collected across 4 phases of the study: screening, pretreatment baseline, treatment, and posttreatment (day 22). The pretreatment baseline assessment period involved 3 survey time points: days −9, −6, and −3 (averaged to create a single pretreatment baseline value). Surveys were distributed 7 times during the active treatment period (days 1, 3, 6, 9, 12, 15, and 18) and posttreatment on days 21 and 22.

Measures

The Defense and Veterans Pain Rating Scale (DVPRS) [43] was used to measure average pain intensity, and the DVPRS interference scale was used to measure pain interference on activity, sleep, mood, and stress over the past 24 hours [44].

Average Pain Intensity

The average pain intensity was rated for the previous 24 hours using an 11-point numeric rating scale (0=no pain and 10=as bad as it could be; nothing else matters). The average pain intensity was assessed pretreatment (baseline), during treatment, and at the end of treatment on day 21.

Pain Interference on Activity, Mood, Sleep, and Stress

Participants were asked to rate the extent to which their pain interfered with their activity, mood, sleep, and stress over the past 24 hours (0=does not interfere and 10=completely interferes). Pain interference was assessed pretreatment (baseline), during treatment, and at the end of treatment on day 21.

Pain Catastrophizing

The 13-item Pain Catastrophizing Scale (PCS) [45] is a validated instrument widely used clinically and in pain research to assess patterns of negative cognition and emotion in the context of actual or anticipated pain. Despite having discrete subscales for rumination, magnification, and feelings of helplessness related to pain, prior work has shown that the PCS operates unidimensionally [46]. For this study and the purpose of brevity, the following 4 items were used: “It’s terrible and I think it’s never going to get any better,” “I become afraid that the pain will get worse,” “I can’t seem to keep it out of my mind,” and “I keep thinking about how badly I want the pain to stop.” Respondents rate the frequency in which they experience such thoughts on a scale ranging from 0 (not at all) to 4 (all the time). The items are summed to create a total score and index for pain catastrophizing. The four-item PCS was administered on day 0 (baseline) and on day 22.

Pain Self-Efficacy

The two-item Pain Self-Efficacy Questionnaire (PSEQ-2) is a validated instrument used to assess patients’ confidence in their ability to carry out their daily activities [47]. The scale includes the following 2 items: “I can still accomplish most of my goals in life, despite the pain” and “I can live a normal lifestyle, despite the pain.” Respondents rate their responses to the items using a 7-point scale, ranging from 0 (not at all confident) to 4 (completely confident). The two-item scores are summed to create a total score. PSEQ-2 was administered on day 0 (baseline) and on day 22.

Patient Global Impression of Change

Aligning with IMMPACT recommendations for pain research [48], Patient Global Impression of Change was assessed on day 22 (posttreatment survey) using the question, “Since the beginning of the study, how has your pain changed?” on a 7-point scale, ranging from *much worse* to *much better*.

Satisfaction With Treatment

Satisfaction with treatment was assessed on day 22 (posttreatment survey) using the question, "How satisfied or dissatisfied are you with the ability of this VR (audio) program to relieve your symptoms?" on a 5-point scale (1=extremely dissatisfied to 5=extremely satisfied).

Motion Sickness and Nausea

Adverse experiences with using VR were assessed on day 22 (posttreatment survey) using the question, "Did you experience any motion sickness or nausea while using VR?" on a 4-point scale, with 0=never, 1=sometimes, 2=often, and 3=always. This single-item evaluation of cybersickness was adapted from related work on visual images and motion sickness.

Study Sample and Analytic Data Set

A convenience sample of 97 participants who met all study criteria was enrolled, randomized, and allocated to the VR or audio treatment groups (VR=47 and audio=50). Of 97 participants, 88 (VR=42 and audio=46) provided pretreatment baseline values for the 5 pain indicators and moved to the intervention phase of the study. Of the 88 participants, 74 (VR=35 and audio=39) completed at least one survey over the 21-day treatment phase. The analytic sample comprised 74 participants who completed the baseline measures and at least one survey during the intervention phase.

Statistical Analysis

The feasibility of the VR treatment was indexed using 3 aspects of the participants' experience: engagement, satisfaction, and adverse effects nausea/discomfort. For engagement, individual-level session data exist only for the VR group; as such, we presented a descriptive statistic of the total number of sessions launched by the VR and audio groups. Group differences in posttreatment satisfaction (day 22) were assessed using a *t* test. For nausea and discomfort, we provided the proportion of participants who did not experience any nausea/discomfort.

With regard to the secondary outcome of the efficacy of the VR treatment on the 5 pain-related indicators (average pain intensity and pain interference with activity, mood, sleep, and stress), we specified a repeated measures model with time of measurement (henceforth *time*) as the sole predictor. The effect of interest was whether there was a significant improvement in the 5 pain indicators from baseline through the end of treatment, assessed through the significance of the main effect for time. To test the immersive effects of VR relative to audio treatment, we analyzed each of the 5 pain indicators in a linear mixed model framework in which treatment (VR versus audio) was specified as a between-participants factor and time was specified as a within-participants factor with participant-specific intercept and time specified as random effects. The effect of interest was whether the improvement in the pain indicators was different for the VR versus the audio group over time, which was assessed through the significance of the group×time interaction. Posttreatment effect sizes (baseline to treatment completion at day 21) were computed by treatment group using an adaptation of Cohen *d* to suit the repeated measures design [49].

The four-item PCS and two-item PSEQ scales were analyzed in a mixed modeling framework, except that there were only 2 time points (baseline day 0 and day 22). We aimed to test 3 questions: (1) whether pain catastrophizing reduced over time and pain self-efficacy increased over time, both assessed through the significance of the time main effect; (2) whether the VR versus audio group had a differential effect, assessed through the group main effect; and (3) whether the immersive aspect of VR produced a differential impact compared with audio over time, assessed through a time×group interaction.

Group equivalence was assessed through univariate tests of association between treatment groups (VR/audio) on demographics (age, gender, race, household income, number of children ≥17 in the home, employment status, and relationship status), anhedonia, depression, and baseline levels of pain intensity and pain-related interference to sleep, activity, stress, and mood.

Results

Sample Characteristics

No significant differences were found between the VR and audio groups for baseline demographic, anhedonia, and depression variables (Table 1). Treatment groups differed in duration of pain since onset, with the VR group having greater pain duration as indexed by the following. Pain duration of 1 year to <5 years represented 40% (14/35) of the VR group and 21% (8/39) of the audio group, whereas pain duration of <1 year represented 3% (1/35) of the VR group and 21% (8/39) of the audio group (P=.03).

Furthermore, the baseline values for average pain intensity and pain-related interference with activity, mood, sleep, and stress were found to be equivalent between the study groups (all P values>.29; Multimedia Appendix 3).

Table 1. Participant characteristics by treatment group (N=74).

Variable	Audio (n=39)	Virtual reality (n=35)	*P* value[a]
Gender, n (%)			**.47**
Male	26 (67)	26 (74)	
Female	13 (33)	9 (26)	
Age group (years), n (%)			**.63**
25-34	3 (8)	3 (9)	
35-44	8 (21)	5 (14)	
45-54	12 (31)	11 (31)	
55-64	7 (18)	11 (31)	
65-74	9 (23)	5 (14)	
Race, n (%)			**.63**
Missing	2 (5)	6 (17)	
African American	3 (8)	4 (14)	
Asian	2 (5)	0 (0)	
White	28 (76)	21 (72)	
Hispanic/Latino	2 (5)	3 (10)	
Multiracial/other	1 (3)	0 (0)	
Native American/Pacific Islander	1 (3)	1 (3)	
Education, n (%)			**.17**
Missing	2 (5)	6 (17)	
Some high school	3 (8)	0 (0)	
High school graduate	11 (30)	14 (48)	
Some college	2 (5)	4 (14)	
Bachelor degree	13 (35)	6 (21)	
Postgraduate	8 (22)	5 (17)	
Employment, n (%)			**.75**
Missing	2 (5)	6 (17)	
Full time	16 (43)	11 (38)	
Part time	7 (19)	9 (31)	
Not working	3 (8)	3 (10)	
Retired	1 (3)	1 (3)	
Unable to work	10 (27)	5 (17)	
Marital status, n (%)			**.16**
Missing	2 (5)	6 (17)	
Married/civil union	16 (43)	15 (52)	
Widowed	1 (3)	1 (3)	
Divorced/separated	5 (14)	5 (17)	
Single/cohabitating	1 (3)	4 (14)	
Single	14 (38)	4 (14)	
Pain onset, n (%)			**.03**
6 months to <1 year	8 (21)	1 (3)	
1 year to <5 years	8 (21)	14 (40)	
5 years to <10 years	10 (26)	13 (37)	

Variable	Audio (n=39)	Virtual reality (n=35)	*P* value[a]
>10 years	13 (33)	7 (20)	
Anhedonia			**.20**
Mean (SD)	1.4 (0.8)	1.1 (0.8)	
Minimum to maximum	0.0-3.0	0.0-3.0	
Median (IQR)	1.0 (1.0-2)	1.0 (1.0-2)	
Depression			**.15**
Mean (SD)	1.2 (0.7)	0.9 (0.7)	
Minimum to maximum	0.0-3.0	0.0-3.0	
Median (IQR)	1.0 (1.0-2)	1.0 (0.0-1)	

[a]The *P* values represent the parametric test of significant differences between the virtual reality and audio conditions [50].

Feasibility of At-Home Virtual Reality Interventions for Chronic Pain

This study was designed to provide indicators for feasibility (indexed by participant engagement, satisfaction, and adverse effects) for an at-home self-administered 3-week VR chronic pain treatment program. In the following section, we reported the results for these 3 feasibility indicators.

Participant Engagement (Number of Treatment Sessions Completed)

Participants were encouraged to follow the 21-day treatment schedule and told they could repeat any completed experiences at any time. In contrast to the VR group, individual-level session launch data do not exist for the audio group, and engagement data exist only at the group level. As such, we limited the description to the total number of sessions launched by the treatment group for the study period. We observed a total of 1048 and 1067 sessions launched in the audio group and the VR group, respectively (these group-level data do not account for discrepancy in group size). VR participants completed an average of 34.4 sessions (SD 20.30), which exceeded the minimum number (21 sessions) they were asked to complete. Overall, 54 participants responded to the day 22 survey (audio=29 and VR=25), which included items on satisfaction with treatment and motion sickness and nausea (for the VR group).

Treatment Satisfaction

Among participants who completed the postintervention survey (n=54), 84% (n=21/25) of participants in the VR group and 72% (n=21/29) of participants in the audio group were either *extremely satisfied* or *very satisfied*.

Motion Sickness and Nausea

Of the 25 VR respondents who completed the day 22 survey, 76% (n=19/25) reported never experiencing nausea or motion sickness. Of the 6 participants who reported nausea or motion sickness, 5 reported experiencing the lowest possible symptom frequency (*sometimes*) and 1 participant reported experiencing it *often*. We tested whether motion sickness and nausea impacted the engagement with the VR treatment by examining the number of VR sessions launched for those with *sometimes* (n=5) and *often* (n=1) nausea/motion sickness and the remainder of the VR group (n=29). We found that experiencing nausea/motion sickness at the lowest level did not negatively impact the use of VR, as indexed by the number of sessions launched (36 sessions versus 33 sessions for the remainder of the VR group). However, the single individual who reported nausea/motion sickness *often* launched only one-third of the number of sessions compared with the remainder of the VR group (11 sessions versus 33 sessions).

Survey Completion

There was no group difference for completion of the 8 treatment surveys administered during and immediately following the 21-day treatment period (VR=5.5 and audio=5.6; *P*=.89).

Preliminary Efficacy on Patient-Reported Outcomes (Virtual Reality Group Only)

For the VR group, we found significant effects on each of the 5 pain indicators using the repeated measures model in which time was the sole predictor (all *P*<.001). For brevity and clarity, we included only the baseline and 21-day mean values in the text but provided the mean values for all time points in the figures. The average reductions from baseline through day 21 were as follows: pain intensity was reduced by 30% (mean reduced from 4.6 to 3.2; Cohen *d* effect size of 0.71), pain-related activity interference reduced by 37% (mean reduced from 4.9 to 3.1; Cohen *d* effect size of 0.83), pain-related mood interference reduced by 50.0% (mean reduced from 5.4 to 2.7; Cohen *d* effect size of 0.94), pain-related sleep interference reduced by 40.4% (mean reduced from 5.2 to 3.1; Cohen *d* effect size of 0.87), and pain-related stress interference reduced by 49.1% (mean reduced from 5.3 to 2.7; Cohen *d* effect size of 0.89). All the aforementioned effect sizes met or exceeded the >30% threshold for clinically important changes, and improvement in pain-related mood interference met the substantial clinical importance threshold of >50% [48,51].

Comparison Between the Virtual Reality and Audio Treatment Groups

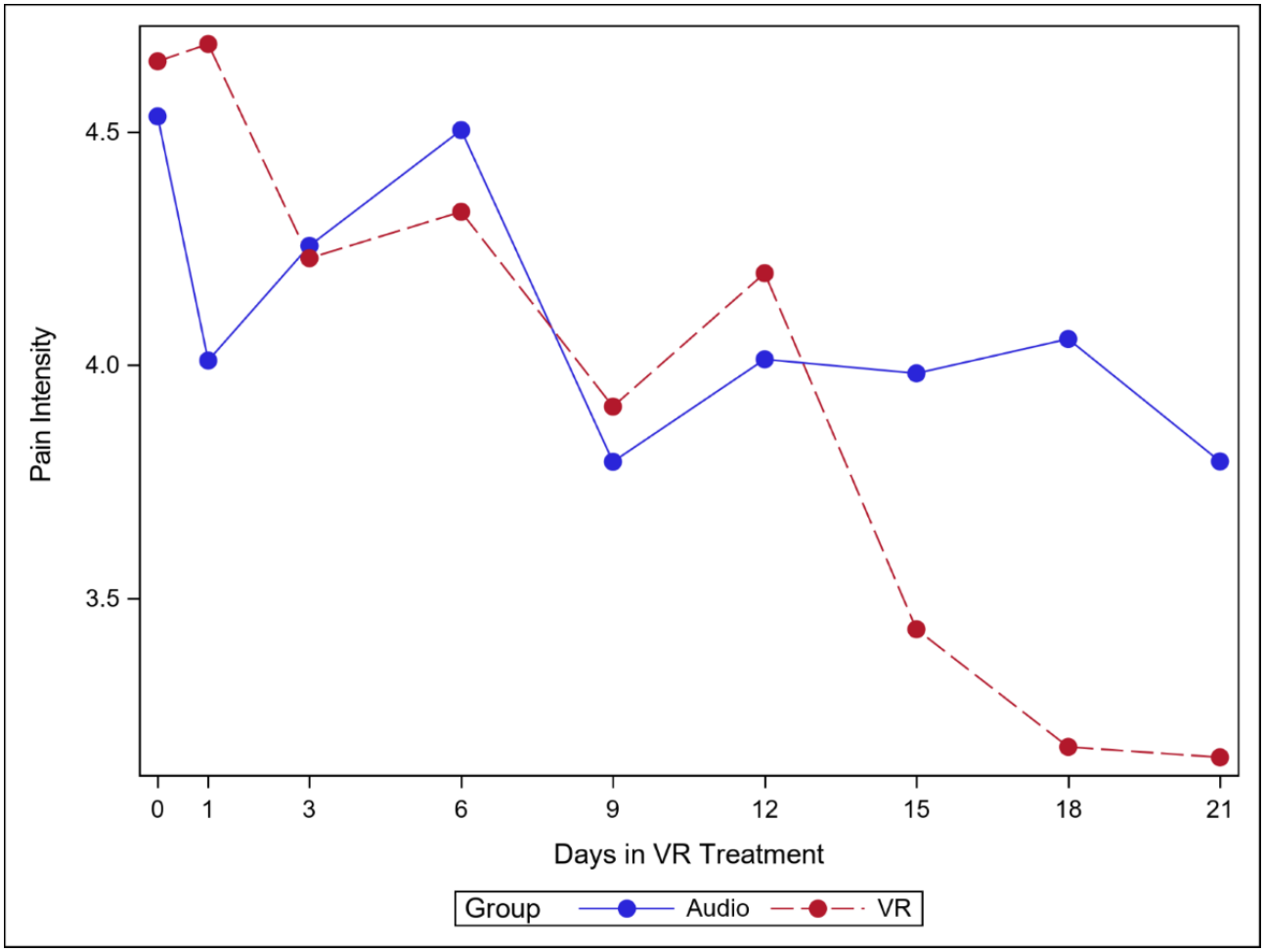
The 5 figures below (Figures 4-8) compared group effects over time for the 5 pain indicators. For all figures, the trend of the pain-related variable was displayed over time for participants in the VR and audio groups. Values in the x-axis refer to the number of days in the study, where *0* represents the baseline

computed as the average of the 3 preexposure measures on −9, −6, and −3 days before the start of the treatment.

Average Pain Intensity

Pain intensity decreased over time in both groups, and the decline was steeper for the VR and audio groups (P=.04), with differences becoming more pronounced from day 15 onward. It should be noted that none of the simple effects (ie, effect of group within time slice) are significant (Figure 4). As seen in Multimedia Appendix 4, from baseline to day 21, consistent with the time×group interaction, the Cohen d was smaller in the audio group than in the VR group (0.42 and 0.71, respectively).

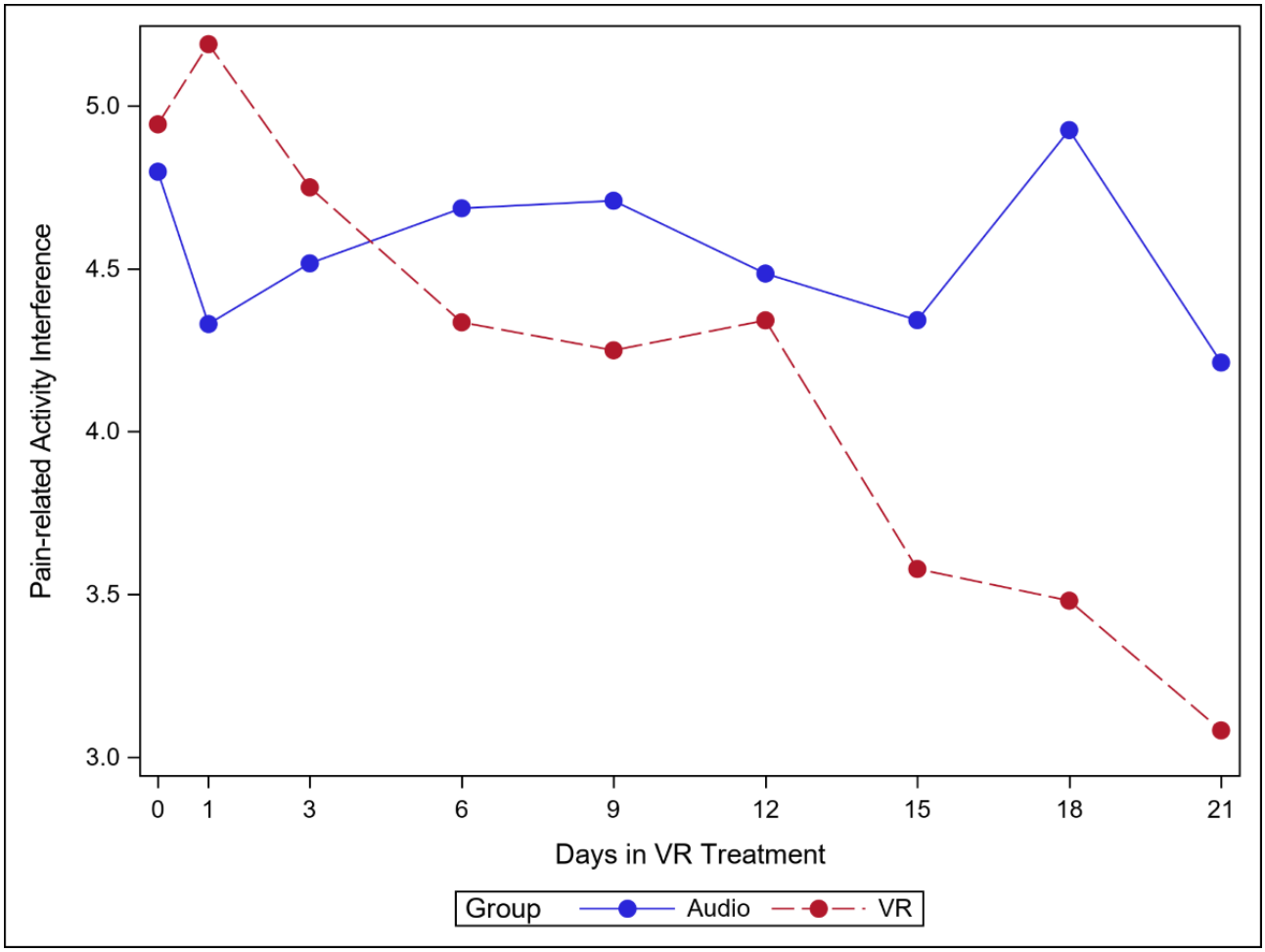
Figure 4. Effect of virtual reality vs audio on pain intensity over time.



Pain-Related Activity Interference

Activity interference decreased more steeply in the VR group than in the audio group (P=.005), with differences becoming more pronounced from day 15 onward (Figure 5). The simple effect (ie, effect of group within time slice) was significant at day 18 (VR<audio; P=.02). As seen in Multimedia Appendix 4, Cohen d was smaller in the audio group than in the VR group (0.26 and 0.83, respectively).

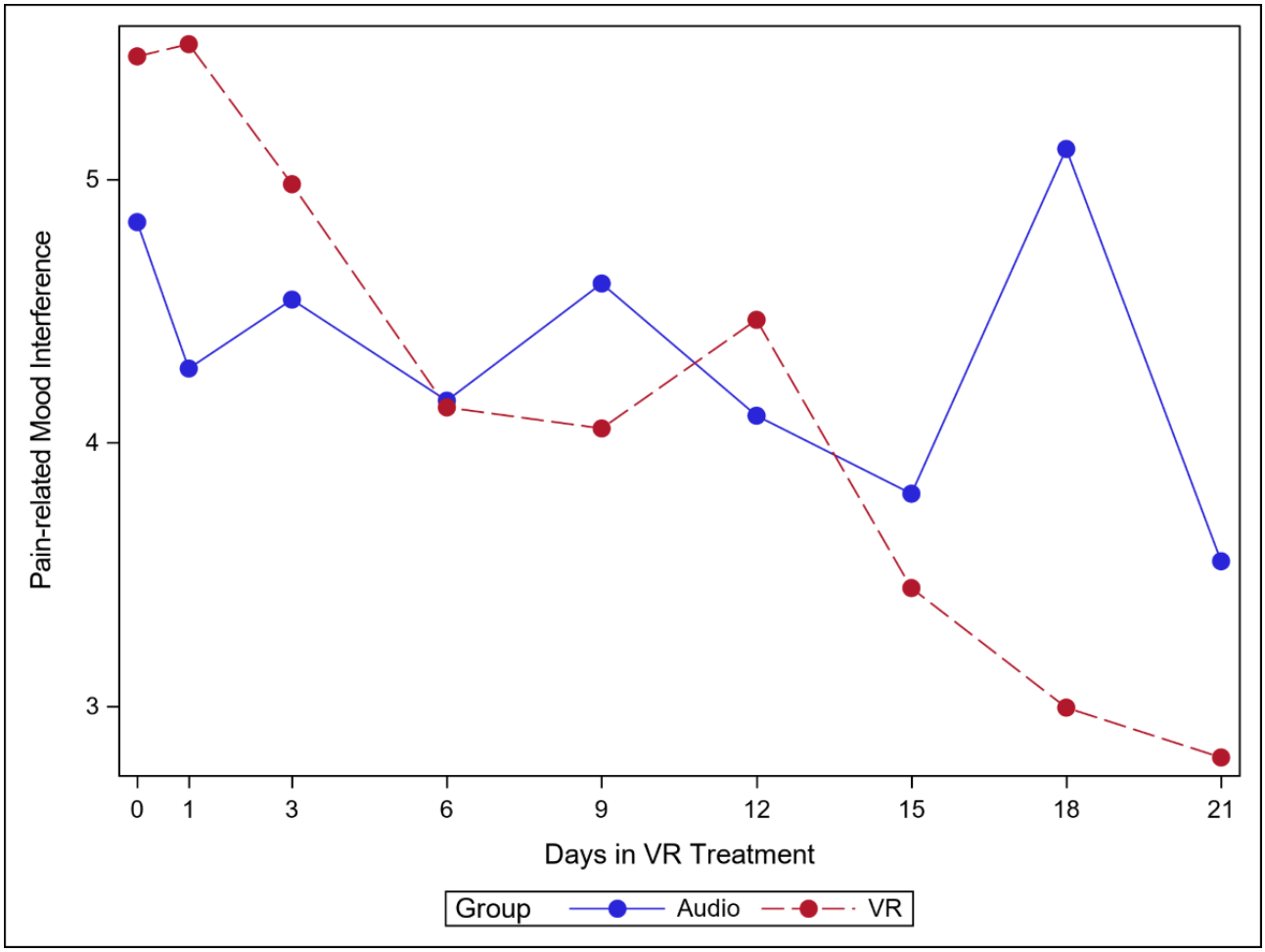
Figure 5. Effect of virtual reality vs audio on pain-related activity interference over time.



Pain-Related Mood Interference

As shown in Figure 6, mood interference appeared to decrease more steeply in the VR group than in the audio group. The difference between the VR and audio groups became more pronounced from day 15 onward. The simple effect (ie, effect of group within time slice) was significant at day 18 (VR<audio; $P<.001$). As seen in Multimedia Appendix 4, Cohen d was smaller in the audio group than in the VR group (0.76 and 0.94, respectively).

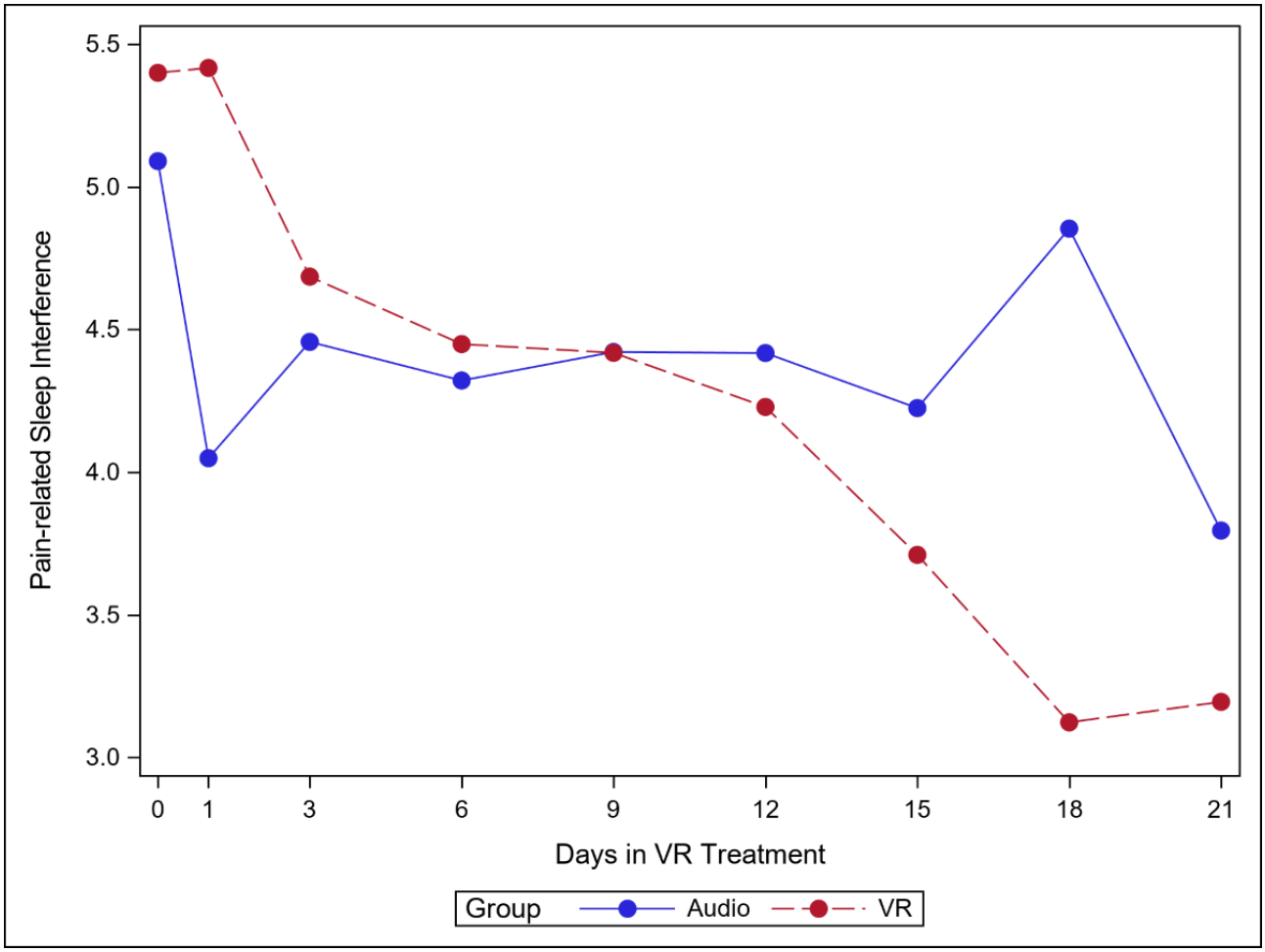
Figure 6. Effect of virtual reality vs audio on pain-related mood interference over time.



Pain-Related Sleep Interference

Reductions in sleep interference were greater in the VR group, with the simple effects (ie, effect of group within time slice) reaching significance at day 1 (audio<VR; *P*=.02) and on day 18 (VR<audio, Figure 7; *P*=.002).

Figure 7. Effect of virtual reality vs audio on pain-related sleep interference over time.

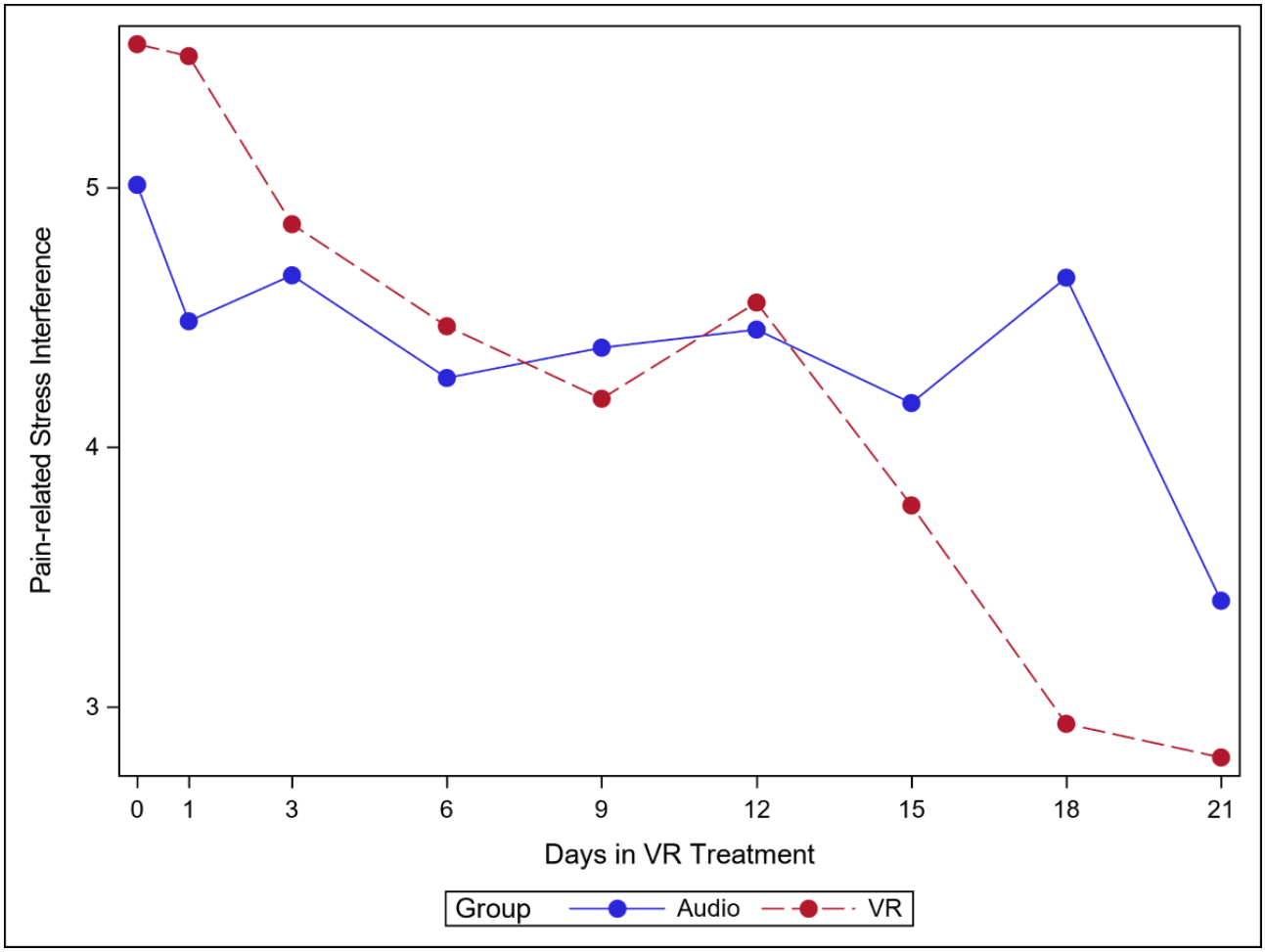


As seen in Multimedia Appendix 4, Cohen *d* was smaller in the audio group than in the VR group (0.64 and 0.87, respectively).

Pain-Related Stress Interference

As shown in Figure 8, stress interference appeared to decrease more steeply in the VR group compared with the audio group, and this difference was more pronounced from day 15 onward. The simple effect (ie, effect of group within time slice) was significant on day 18 (P=.01). As seen in Multimedia Appendix 4, Cohen *d* was smaller in the audio group than in the VR group (0.87 and 0.89, respectively).

Figure 8. Effect of virtual reality vs audio on pain-related stress interference over time.



Pain Catastrophizing

We observed a significant main effect of time, with pain catastrophizing decreasing for both groups over time ($P<.001$). The main effect for group was not significant ($P=.61$). Finally, we did not perform a time×group interaction ($P=.52$).

Pain Self-Efficacy

We observed a significant main effect of time, with pain self-efficacy increasing in both groups over time ($P<.047$). The main effect for group was not significant ($P=.68$). Finally, we did not perform a time×group interaction ($P=.46$).

Global Impression of Change

At day 22, among survey responders (n=54), 84% (n=21/25) of participants in the VR group reported that their pain was improved, 16% (n=4/25) reported no change, and 0% (n=0/25) reported worsening pain. In the audio group, 62% (n=18/29) of participants reported improvement, 34% (n=10/29) reported no change, and 3% (n=1/29) reported worsening pain.

Finally, because of the observed group difference in the duration of chronic pain (indexed as pain onset in Table 1), we conducted additional analyses with pain onset specified as a covariate in the model with time, group, and time×group as predictors of the 5 pain variables. The significance of the time main effect and time×group interaction effects was fully preserved. When intention-to-treat (ITT) analyses were applied all of the results were preserved with one exception; the time×treatment interaction was no longer significant.

Discussion

Principal Findings

We conducted an unblinded randomized controlled study in a web-based convenience sample of community-based participants with nonmalignant chronic low back pain and/or fibromyalgia by comparing at-home self-administered VR treatment with an audio-only treatment group (same audio content as VR with minor modifications made for one-third of the modules). The primary goal of the study was to evaluate the feasibility of a self-administered home-based VR program for chronic pain that included skills-based content informed by evidence-based CBT for chronic pain. The secondary goal was to conduct an RCT of the VR treatment to an audio-only treatment. We aimed to determine the preliminary efficacy of VR for reducing average pain intensity and pain-related interference with activity, mood, sleep, and stress over the 21-day treatment program. Our tertiary goal was to isolate the immersive effects of the skills-based VR program by comparing the effects between treatment groups.

VR demonstrated good feasibility as indexed by excellent participant engagement (average of 34 sessions completed over the 21-day treatment program), high ratings for satisfaction with the treatment (84%), and relatively low reporting for motion sickness and nausea (n=6/25, 24% of VR participants who

completed the day 22 survey). Of these 6 participants, 5 reported the lowest level of symptom frequency possible (*sometimes*). We found that this symptom level did not interfere with VR treatment engagement, as indexed by the number of sessions launched compared with participants in the remainder of the VR group. The single individual who experienced nausea and motion sickness *often* showed markedly decreased use of VR (11 sessions launched versus 34 for people with low nausea and the remainder of the VR group).

With respect to preliminary efficacy, VR demonstrated the significant reduction in average pain intensity and pain-related interference in activity, mood, sleep, and stress over the course of the 21-day treatment program. Treatment effect sizes suggested that a home-based stand-alone, digital, skills-based treatment program may affect clinically meaningful changes in patient-reported pain and pain correlates. The durability of the treatment effects reported here remains a topic for future research.

Although a significant body of research exists on the use of VR for acute pain management [26,28,30,52-54] and for physical rehabilitation [24,40], the use of VR as a platform to deliver behavioral medicine for chronic pain remains novel and understudied. Research on VR for acute pain is based on the premise that distraction is a primary mechanism of VR analgesia [33,55]. Therefore, the effects of distraction on pain are typically measured within a rapid time frame using study designs that align with drug trials. Investigations of VR for chronic pain require a different approach to align with the goal of sustained pain management. Although VR for physical rehabilitation shows sustained results, content is typically constrained to rehabilitative movement, exercise, and kinematic training [24,40] and is devoid of didactics and skills training contained in evidence-based behavioral medicine treatments [7].

Unlike distraction alone, behavioral health therapies rarely produce instantaneous results. Rather, skills acquisition and mastery require time, in part because of the multisession delivery of content, and become effective over the course of weeks, as the content is delivered more comprehensively, and patients practice skills during and between sessions and is correlated with patient engagement [56]. Our prior work suggests that an ultrabrief skill-based intervention for chronic pain evidenced clinically meaningful improvements in pain-related symptoms at 2 weeks, with even greater improvements evidenced at 4 weeks [12]. The results presented here dovetail with this literature and our fundamental understanding of how didactic and skills-based behavioral medicine treatment results accrue over time as participants receiving increasing amounts of knowledge and skills practice during active treatment [6,56,57]. In this study, the emergence of a more pronounced improvement in pain outcomes at day 15 supports the hypothesis that the didactic and skills-based elements of immersive behavioral medicine VR are operating as expected, although confirmatory studies are needed.

In evaluating the method of delivery of a skills-based treatment program, this study demonstrated superior reduction of most pain indicators in the VR group relative to the audio group after 2 weeks. Treatment group differences remained minimal until day 12, with VR superiority strengthening from day 15 onward and peaking on day 18. For the pain interference outcomes, the VR group showed greater improvement compared with the audio group. The exceptions were pain catastrophizing and pain self-efficacy, which improved for both treatment groups. We have no evidence that VR's superior treatment effects were explained by user engagement, as we observed similar rates of engagement for both groups. Our study design allowed us to isolate the immersive effect of VR relative to active treatment delivered via audio format only. On balance, our findings suggest that the treatment effect sizes for VR are both statistically significant and clinically meaningful for pain intensity and across the pain interference variables and are superior to the same treatment delivered by audio alone. A key aspect of the immersive VR experience involves dynamic interaction between the user's breath and the environment, wherein voice-over coaching directs the user to slow the breath to engage a parasympathetic response. The environment responds to the breath, provides visual feedback to the user, and possibly affords users enhanced acquisition of this skill relative to the audio-only content.

Strengths and Limitations

The strengths and limitations of our study bear careful attention. The interpretation of study findings is limited by the 2 pain conditions studied and the selection bias inherent in the web-based convenience sample. In addition, the analytic data set included only those participants who completed at least one study survey; accordingly, larger studies are needed to confirm the findings reported here and to determine generalizability. Medication use was not assessed and may have been a confounding variable. Chronic pain type and duration were self-reported, and there was no review of medical records to confirm diagnoses. Our ability to assess VR satisfaction and nausea/motion sickness was limited by only 25 of 35 participants completing the day 22 posttreatment survey. Although only 17% of the full VR sample reported experiencing cybersickness to any degree, we cannot rule out the possibility that early attrition may be partially attributable to these adverse effects, although notably, we did not find disparate attrition rates between the VR and audio groups. Commercially available VR programs typically offer a money-back guarantee trial period to allow customers to return their VR device for refund in cases of cybersickness. Although assignment to treatment group was random, differences in duration of pain since onset between the 2 treatment groups may have influenced pain outcomes during the course of the study, and we note that this difference favored the audio group. In addition, the analysis did not focus on the correlation between individual-level variations in the use of the intervention (VR or audio) and the pain indicators, and this remains a topic for future study.

Our study design merits consideration within the context of our findings. The audio treatment group was an active comparator with two-thirds of the audio content identical to the audio of the VR group, and one-third of the audio content closely matched the VR content for topical content, skills, and experience (minus the visual and interactive elements). This study was rigorously designed to isolate the immersive effects of VR rather than simply comparing VR with placebo or with

a weaker control group such as *usual care*. In studying a treatment delivered 2 ways, our results suggest that the immersive (3D visual elements, dynamic scenery, and 360-degree vision capability) and user interactivity with the environment are pleasing, engaging, and generally effective for those who do not experience cybersickness (many commercial VR companies offer consumers a trial period or money-back guarantee for cases of cybersickness).

Finally, the cost-effectiveness of VR for chronic pain relative to in-person behavioral medicine visits or other digital treatment options merits investigation. In-person behavioral medicine treatments require multiple clinic visits, travel costs, time from work and other obligations, and treatment copayment costs that may total several hundreds of dollars over a standard 8-session treatment package. Future research should compare home-based VR to in-person multisession CBT for chronic pain in terms of efficacy and cost-effectiveness. Although some 2D digital treatment options, such as web-based pain CBT or self-management programs, and may be sourced at no cost, engagement rates remain relatively low. Future research may explore participants' perceived comparative value of VR to audio-only treatment. VR provides patients and clinicians with a new home-based treatment option that may be preferred by some patients and also provide more effective pain management to a subset of individuals. Additional studies of longer duration may investigate the durability of treatment effects reported here.

Conclusions

This study is one of the first to explore how a self-administered home-based VR program rooted in behavioral medicine skills and techniques impacts chronic pain. The findings broadly suggest that VR holds promise as effective, stand-alone, home-based digital behavioral medicine for chronic pain. Additional studies are needed, including larger sample sizes, diverse chronic pain conditions, longer duration of study to best characterize the efficacy of VR in chronic pain management, and the impact of VR on other factors such as pain medication use and medical care utilization. Future studies may further elucidate the VR mechanisms of action and VR's role in expanding access to multimodal behavioral medicine for chronic pain.

Acknowledgments

This study was financially supported by AppliedVR.

Conflicts of Interest

BD is employed as a chief scientific adviser to AppliedVR. PK serves as a consultant for AppliedVR.

Multimedia Appendix 1
Participant consent form.
[DOC File , 84 KB - formative_v4i7e17293_app1.doc]

Multimedia Appendix 2
Schedule of 21-day virtual reality and audio programs.
[DOCX File , 14 KB - formative_v4i7e17293_app2.docx]

Multimedia Appendix 3
Baseline pain-related variables by treatment group.
[DOCX File , 17 KB - formative_v4i7e17293_app3.docx]

Multimedia Appendix 4
Effect sizes for change from baseline to day 21.
[DOCX File , 18 KB - formative_v4i7e17293_app4.docx]

References

1. Interagency Pain Research Coordinating Committee - NIH. 2016. National Pain Strategy Overview URL: https://www.iprcc.nih.gov/National-Pain-Strategy/Overview [accessed 2019-12-10]
2. The US Department of Health and Human Services. 2019. Pain Management Best Practices: Inter-Agency Task Force Report URL: https://www.hhs.gov/sites/default/files/pmtf-final-report-2019-05-23.pdf [accessed 2019-11-10]
3. Board on Health Sciences Policy, Committee on Advancing Pain Research Care and Education, Institute of Medicine. Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research. New York, USA: National Academies Press; 2011.
4. International Association for the Study of Pain, Subcommittee on Taxonomy. Classification of chronic pain. Descriptions of chronic pain syndromes and definitions of pain terms. Pain Suppl 1986;3:S1-226. [Medline: 3461421]
5. Turner JA, Anderson ML, Balderson BH, Cook AJ, Sherman KJ, Cherkin DC. Mindfulness-based stress reduction and cognitive behavioral therapy for chronic low back pain: similar effects on mindfulness, catastrophizing, self-efficacy, and

acceptance in a randomized controlled trial. Pain 2016 Nov;157(11):2434-2444 [FREE Full text] [doi: 10.1097/j.pain.0000000000000635] [Medline: 27257859]
6. Cherkin DC, Sherman KJ, Balderson BH, Cook AJ, Anderson ML, Hawkes RJ, et al. Effect of mindfulness-based stress reduction vs cognitive behavioral therapy or usual care on back pain and functional limitations in adults with chronic low back pain: a randomized clinical trial. J Am Med Assoc 2016;315(12):1240-1249 [FREE Full text] [doi: 10.1001/jama.2016.2323] [Medline: 27002445]
7. Williams AC, Eccleston C, Morley S. Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database Syst Rev 2012 Nov 14;11:CD007407 [FREE Full text] [doi: 10.1002/14651858.CD007407.pub3] [Medline: 23152245]
8. Hughes LS, Clark J, Colclough JA, Dale E, McMillan D. Acceptance and commitment therapy (ACT) for chronic pain: a systematic review and meta-analyses. Clin J Pain 2017 Jun;33(6):552-568. [doi: 10.1097/AJP.0000000000000425] [Medline: 27479642]
9. Darnall B. Psychological Treatment for Patients With Chronic Pain. Washington, DC: American Psychological Association Press; 2018.
10. Darnall BD, Scheman J, Davin S, Burns JW, Murphy JL, Wilson AC, et al. Pain psychology: a global needs assessment and national call to action. Pain Med 2016 Feb;17(2):250-263 [FREE Full text] [doi: 10.1093/pm/pnv095] [Medline: 26803844]
11. Thorn BE, Pence LB, Ward LC, Kilgo G, Clements KL, Cross TH, et al. A randomized clinical trial of targeted cognitive behavioral treatment to reduce catastrophizing in chronic headache sufferers. J Pain 2007 Dec;8(12):938-949. [doi: 10.1016/j.jpain.2007.06.010] [Medline: 17690017]
12. Darnall BD, Sturgeon JA, Kao M, Hah JM, Mackey SC. From catastrophizing to recovery: a pilot study of a single-session treatment for pain catastrophizing. J Pain Res 2014;7:219-226 [FREE Full text] [doi: 10.2147/JPR.S62329] [Medline: 24851056]
13. Kelleher SA, Winger JG, Dorfman CS, Ingle KK, Moskovich AA, Abernethy AP, et al. A behavioral cancer pain intervention: a randomized noninferiority trial comparing in-person with videoconference delivery. Psychooncology 2019 Aug;28(8):1671-1678. [doi: 10.1002/pon.5141] [Medline: 31162756]
14. Darnall B, Ziadni MS, Krishnamurthy P, Flood P, Heathcote LC, Mackey IG, et al. 'My surgical success': effect of a digital behavioral pain medicine intervention on time to opioid cessation after breast cancer surgery-a pilot randomized controlled clinical trial. Pain Med 2019 Nov 1;20(11):2228-2237 [FREE Full text] [doi: 10.1093/pm/pnz094] [Medline: 31087093]
15. Palermo TM, Murray C, Aalfs H, Abu-El-Haija M, Barth B, Bellin MD, Consortium for the Study of Chronic Pancreatitis, Diabetes and Pancreatic Cancer. Web-based cognitive-behavioral intervention for pain in pediatric acute recurrent and chronic pancreatitis: Protocol of a multicenter randomized controlled trial from the study of chronic pancreatitis, diabetes and pancreatic cancer (CPDPC). Contemp Clin Trials 2020 Jan;88:105898. [doi: 10.1016/j.cct.2019.105898] [Medline: 31756383]
16. Fisher E, Law E, Palermo TM, Eccleston C. Psychological therapies (remotely delivered) for the management of chronic and recurrent pain in children and adolescents. Cochrane Database Syst Rev 2014 May 14;2014:1 [FREE Full text] [doi: 10.1002/14651858.CD011118] [Medline: 25221436]
17. Palermo TM, Wilson AC, Peters M, Lewandowski A, Somhegyi H. Randomized controlled trial of an Internet-delivered family cognitive–behavioral therapy intervention for children and adolescents with chronic pain. Pain 2009 Nov;146(1):205-213 [FREE Full text] [doi: 10.1016/j.pain.2009.07.034] [Medline: 19695776]
18. Hoffman HG, Richards T, Coda B, Richards A, Sharar SR. The illusion of presence in immersive virtual reality during an fMRI brain scan. Cyberpsychol Behav 2003 Apr;6(2):127-131. [doi: 10.1089/109493103321640310] [Medline: 12804024]
19. Gold JI, Belmont KA, Thomas DA. The neurobiology of virtual reality pain attenuation. Cyberpsychol Behav 2007 Aug;10(4):536-544. [doi: 10.1089/cpb.2007.9993] [Medline: 17711362]
20. Geenen R, Bijlsma JW. Deviations in the endocrine system and brain of patients with fibromyalgia: cause or consequence of pain and associated features? Ann N Y Acad Sci 2010 Apr;1193:98-110. [doi: 10.1111/j.1749-6632.2009.05290.x] [Medline: 20398014]
21. Nararro-Haro MV, Hoffman HG, Garcia-Palacios A, Sampaio M, Alhalabi W, Hall K, et al. The use of virtual reality to facilitate mindfulness skills training in dialectical behavioral therapy for borderline personality disorder: a case study. Front Psychol 2016;7:1573 [FREE Full text] [doi: 10.3389/fpsyg.2016.01573] [Medline: 27853437]
22. Parsons TD, Rizzo AA. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis. J Behav Ther Exp Psychiatry 2008 Sep;39(3):250-261. [doi: 10.1016/j.jbtep.2007.07.007] [Medline: 17720136]
23. Parsons TD, Rizzo AA, Rogers S, York P. Virtual reality in paediatric rehabilitation: a review. Dev Neurorehabil 2009 Aug;12(4):224-238. [doi: 10.1080/17518420902991719] [Medline: 19842822]
24. Ahern MM, Dean LV, Stoddard CC, Agrawal A, Kim K, Cook CE, et al. The effectiveness of virtual reality in patients with spinal pain: a systematic review and meta-analysis. Pain Pract 2020 Mar 20:- epub ahead of print. [doi: 10.1111/papr.12885] [Medline: 32196892]
25. Gold JI, Kim SH, Kant AJ, Joseph MH, Rizzo AS. Effectiveness of virtual reality for pediatric pain distraction during iv placement. Cyberpsychol Behav 2006 Apr;9(2):207-212. [doi: 10.1089/cpb.2006.9.207] [Medline: 16640481]

26. Atzori B, Grotto RL, Giugni A, Calabrò M, Alhalabi W, Hoffman HG. Virtual reality analgesia for pediatric dental patients. Front Psychol 2018;9:2265 [FREE Full text] [doi: 10.3389/fpsyg.2018.02265] [Medline: 30532720]
27. Gold JI, Mahrer NE. Is virtual reality ready for prime time in the medical space? A randomized control trial of pediatric virtual reality for acute procedural pain management. J Pediatr Psychol 2018 Apr 1;43(3):266-275. [doi: 10.1093/jpepsy/jsx129] [Medline: 29053848]
28. Haisley KR, Straw OJ, Müller DT, Antiporda MA, Zihni AM, Reavis KM, et al. Feasibility of implementing a virtual reality program as an adjuvant tool for peri-operative pain control; results of a randomized controlled trial in minimally invasive foregut surgery. Complement Ther Med 2020 Mar;49:102356. [doi: 10.1016/j.ctim.2020.102356] [Medline: 32147069]
29. Özkan TK, Polat F. The effect of virtual reality and kaleidoscope on pain and anxiety levels during venipuncture in children. J Perianesth Nurs 2020 Apr;35(2):206-211. [doi: 10.1016/j.jopan.2019.08.010] [Medline: 31759833]
30. Hoffman HG, Chambers GT, Meyer WJ, Arceneaux LL, Russell WJ, Seibel EJ, et al. Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures. Ann Behav Med 2011 Apr;41(2):183-191 [FREE Full text] [doi: 10.1007/s12160-010-9248-7] [Medline: 21264690]
31. Khadra C, Ballard A, Déry J, Paquin D, Fortin J, Perreault I, et al. Projector-based virtual reality dome environment for procedural pain and anxiety in young children with burn injuries: a pilot study. J Pain Res 2018;11:343-353 [FREE Full text] [doi: 10.2147/JPR.S151084] [Medline: 29491717]
32. Birckhead B, Khalil C, Liu X, Conovitz S, Rizzo A, Danovitch I, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973 [FREE Full text] [doi: 10.2196/11973] [Medline: 30702436]
33. Dascal J, Reid M, IsHak WW, Spiegel B, Recacho J, Rosen B, et al. Virtual reality and medical inpatients: a systematic review of randomized, controlled trials. Innov Clin Neurosci 2017;14(1-2):14-21 [FREE Full text] [Medline: 28386517]
34. Sato K, Fukumori S, Matsusaki T, Maruo T, Ishikawa S, Nishie H, et al. Nonimmersive virtual reality mirror visual feedback therapy and its application for the treatment of complex regional pain syndrome: an open-label pilot study. Pain Med 2010 Apr;11(4):622-629. [doi: 10.1111/j.1526-4637.2010.00819.x] [Medline: 20202141]
35. Shiri S, Feintuch U, Weiss N, Pustilnik A, Geffen T, Kay B, et al. A virtual reality system combined with biofeedback for treating pediatric chronic headache--a pilot study. Pain Med 2013 May;14(5):621-627. [doi: 10.1111/pme.12083] [Medline: 23659372]
36. Garcia-Palacios S, Herrero R, Vizcaíno Y, Belmonte MA, Castilla D, Molinari G, et al. Integrating virtual reality with activity management for the treatment of fibromyalgia: acceptability and preliminary efficacy. Clin J Pain 2015 Jun;31(6):564-572. [doi: 10.1097/AJP.0000000000000196] [Medline: 25551475]
37. Botella C, Garcia-Palacios A, Vizcaíno Y, Herrero R, Baños RM, Belmonte MA. Virtual reality in the treatment of fibromyalgia: a pilot study. Cyberpsychol Behav Soc Netw 2013 Mar;16(3):215-223. [doi: 10.1089/cyber.2012.1572] [Medline: 23496678]
38. Wiederhold BK, Gao K, Sulea C, Wiederhold MD. Virtual reality as a distraction technique in chronic pain patients. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):346-352 [FREE Full text] [doi: 10.1089/cyber.2014.0207] [Medline: 24892196]
39. Sulea C, Soomro A, Wiederhold BK, Wiederhold MD. Quantifying the effectiveness of virtual reality pain management: a pilot study. Stud Health Technol Inform 2014;199:94-97. [Medline: 24875698]
40. Gumaa M, Youssef AR. Is virtual reality effective in orthopedic rehabilitation? A systematic review and meta-analysis. Phys Ther 2019 Oct 28;99(10):1304-1325. [doi: 10.1093/ptj/pzz093] [Medline: 31343702]
41. Deyo RA, Dworkin SF, Amtmann D, Andersson G, Borenstein D, Carragee E, et al. Report of the NIH task force on research standards for chronic low back pain. Phys Ther 2015 Feb;95(2):e1-18 [FREE Full text] [doi: 10.2522/ptj.2015.95.2.e1] [Medline: 25639530]
42. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, IMMPACT. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005 Jan;113(1-2):9-19. [doi: 10.1016/j.pain.2004.09.012] [Medline: 15621359]
43. Turk DC, Dworkin RH, Burke LB, Gershon R, Rothman M, Scott J, Initiative on Methods, Measurement and Pain Assessment in Clinical Trials. Developing patient-reported outcome measures for pain clinical trials: IMMPACT recommendations. Pain 2006 Dec 5;125(3):208-215. [doi: 10.1016/j.pain.2006.09.028] [Medline: 17069973]
44. Buckenmaier CC, Galloway KT, Polomano RC, McDuffie M, Kwon N, Gallagher RM. Preliminary validation of the defense and veterans pain rating scale (DVPRS) in a military population. Pain Med 2013 Jan;14(1):110-123. [doi: 10.1111/j.1526-4637.2012.01516.x] [Medline: 23137169]
45. Nassif TH, Hull A, Holliday SB, Sullivan P, Sandbrink F. Concurrent validity of the defense and veterans pain rating scale in VA outpatients. Pain Med 2015 Nov 1;16(11):2152-2161. [doi: 10.1111/pme.12866] [Medline: 26257151]
46. Sullivan MJ, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. Psychol Assess 1995;7(4):524-532 [FREE Full text] [doi: 10.1037/1040-3590.7.4.524]
47. Sun EC, Dixit A, Humphreys K, Darnall BD, Baker LC, Mackey S. Association between concurrent use of prescription opioids and benzodiazepines and overdose: retrospective analysis. Br Med J 2017 Mar 14;356:j760 [FREE Full text] [doi: 10.1136/bmj.j760] [Medline: 28292769]

48. Nicholas MK, McGuire BE, Asghari A. A 2-item short form of the pain self-efficacy questionnaire: development and psychometric evaluation of PSEQ-2. J Pain 2015 Feb;16(2):153-163. [doi: 10.1016/j.jpain.2014.11.002] [Medline: 25463701]
49. Dworkin RH, Turk DC, McDermott MP, Peirce-Sandner S, Burke LB, Cowan P, et al. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. Pain 2009 Dec;146(3):238-244. [doi: 10.1016/j.pain.2009.08.019] [Medline: 19836888]
50. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol 2013 Nov 26;4:863 [FREE Full text] [doi: 10.3389/fpsyg.2013.00863] [Medline: 24324449]
51. Gu H, Li D, Liu C, Rao Z. %ggBaseline: a SAS macro for analyzing and reporting baseline characteristics automatically in medical research. Ann Transl Med 2018 Aug;6(16):326 [FREE Full text] [doi: 10.21037/atm.2018.08.13] [Medline: 30364007]
52. Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole MR. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001 Nov;94(2):149-158. [doi: 10.1016/s0304-3959(01)00349-9] [Medline: 11690728]
53. Spiegel B, Fuller G, Lopez M, Dupuy T, Noah B, Howard A, et al. Virtual reality for management of pain in hospitalized patients: a randomized comparative effectiveness trial. PLoS One 2019;14(8):e0219115 [FREE Full text] [doi: 10.1371/journal.pone.0219115] [Medline: 31412029]
54. Hoffman HG, Rodriguez RA, Gonzalez M, Bernardy M, Peña R, Beck W, et al. Immersive virtual reality as an adjunctive non-opioid analgesic for pre-dominantly Latin American children with large severe burn wounds during burn wound cleaning in the intensive care unit: a pilot study. Front Hum Neurosci 2019;13:262 [FREE Full text] [doi: 10.3389/fnhum.2019.00262] [Medline: 31440148]
55. Tashjian VC, Mosadeghi S, Howard AR, Lopez M, Dupuy T, Reid M, et al. Virtual reality for management of pain in hospitalized patients: results of a controlled trial. JMIR Ment Health 2017 Mar 29;4(1):e9 [FREE Full text] [doi: 10.2196/mental.7387] [Medline: 28356241]
56. Spiegel BM. Virtual medicine: how virtual reality is easing pain, calming nerves and improving health. Med J Aust 2018 Sep 17;209(6):245-247. [Medline: 29764343]
57. Winger JG, Nunez C, Kelleher SA, Ingle KK, Gandhi V, Keefe FJ, et al. Predictors of intervention session completion in a randomized clinical trial of a behavioral cancer pain intervention. J Pain Symptom Manage 2020 Jun;59(6):1268-1277. [doi: 10.1016/j.jpainsymman.2020.01.020] [Medline: 32045676]

Abbreviations

CBT: cognitive behavioral therapy
DVPRS: Defense and Veterans Pain Rating Scale
IMMPACT: Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials
PCS: Pain Catastrophizing Scale
PSEQ-2: two-item Pain Self-Efficacy Questionnaire
RCT: randomized controlled trial
VR: virtual reality

Edited by P Santana-Mancilla; submitted 03.12.19; peer-reviewed by M Linnett, A Parks, D Ahern PhD; comments to author 15.01.20; revised version received 12.04.20; accepted 05.05.20; published 07.07.20.

Please cite as:
Darnall BD, Krishnamurthy P, Tsuei J, Minor JD
Self-Administered Skills-Based Virtual Reality Intervention for Chronic Pain: Randomized Controlled Pilot Study
JMIR Form Res 2020;4(7):e17293
URL: https://formative.jmir.org/2020/7/e17293
doi:10.2196/17293
PMID:32374272

Original Paper

What College Students Post About Depression on Facebook and the Support They Perceive: Content Analysis

Scottye Cash[1], MSSW, PhD; Laura Marie Schwab-Reese[2], MA, PhD; Erin Zipfel[1], MSW; Megan Wilt[3], BA; Megan Moreno[3], MD, MPH, MSED

[1]College of Social Work, The Ohio State University, Columbus, OH, United States

[2]Department of Health & Kinesiology, Purdue University, West Lafayette, IN, United States

[3]Department of Pediatrics, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Scottye Cash, MSSW, PhD
College of Social Work
The Ohio State University
1947 College Rd
Columbus, OH, 43210
United States
Phone: 1 614 292 3690
Email: cash.33@osu.edu

Abstract

Background: College students frequently use social media sites to connect with friends. Increasingly, research suggests college students and other young adults seek mental health-related support on social media, which may present a unique venue for intervention.

Objective: The purpose of this study was to examine college students' perceptions about displaying feelings of depression on Facebook and, in turn, how their social media friends responded.

Methods: A primarily quantitative online survey with open response questions was distributed to students at four US universities. Qualitative responses were analyzed using content analysis.

Results: A total of 34 students provided qualitative responses for analysis, these students were 85.3% female, mean age 20.2 (SD=1.4) and 20.6% racial/ethnic minority. Students who reported posting about depression often expressed an emotion or feeling but did not use the word "depression" in the post. Approximately 20% posted language about a bad day, and 15% posted a song or music video. Only one person reported posting a statement that directly asked for help. When friends responded to the posts, students generally perceived the responses as supportive or motivating gestures. Nearly 15% of friends contacted the individual outside of Facebook. One individual received a negative response and no responses suggested that the individual seek help.

Conclusions: This study found that college students who post about depression often do so without directly referencing depression and that friends were generally supportive. However, no participants reported their social network suggested they seek help, which may suggest increasing mental health literacy, for both support seekers and responders, would be an opportunity to improve online mental health-related support.

(JMIR Form Res 2020;4(7):e13650) doi:10.2196/13650

KEYWORDS

social media; depression; college students; qualitative

Introduction

Depression is a common health issue among college students [1-4]. Untreated depression may cause a number of adverse outcomes, including substance use, suicidal behaviors, and other psychiatric conditions [5-10]. Obtaining help with depression can be complicated by students' inability to understand their symptoms, denial of symptoms, unwillingness to reach out for formal help, a lack of available mental health resources, and other complications associated with depression severity [2,3,11]. Innovative approaches are needed to increase the knowledge about depression and help-seeking resources while also reducing the stigma associated with mental health issues.

Internet-based or online social networks may be an important avenue for young people to seek and receive support related to mental health and to access mental health resources. In 2018, nearly all young people in the United States used the internet through home broadband or smartphones [12]. Social network site use is nearly as ubiquitous. In 2018, YouTube was used by the greatest percentage of 18- to 24-year-old Americans (94%), followed by Facebook (80%), Snapchat (78%), and Instagram (71%) [13]. Although social media use has been shown to be associated with increased mental health issues under some circumstances [14-18], a growing body of literature has shown that college students use social network sites, such as Facebook and Twitter, to discuss mental health concerns and seek support [19-22]. Moreno and colleagues [20] found that 25% of respondents had some evidence of depression symptoms in their Facebook posts, and 2.5% posted about symptoms that met the criteria for a major depressive episode. The authors also found that individuals who posted about depression were online more often than those who did not [20]. Support-seeking behaviors may be driving this increased use. One study demonstrated that young adults with a mental health diagnosis were more likely to seek support online and to engage in various other social connection-seeking activities than peers without a mental health diagnosis [21]. However, social media engagement may not be entirely positive. In another study of social media use among adolescents with depression diagnosis, participants used social media in ways that were likely to positively (eg, entertainment, humor) and negatively (eg, sharing risky behaviors, cyberbullying) impact their symptoms [22].

Few studies have examined the type of response that students receive from peers when they post about depression symptoms on Facebook. In one study, a positive association was found between depression symptom posts and peer responses, and peer responses were generally positive and supportive [20]. However, in another study, participants perceived depression symptom posts as “drama” and were reluctant to offer help or support [23]. This discrepancy may be due to differences in how individuals with depression perceive social interactions compared with nondepressed peers. One small study found that individuals who posted about depression *thought* they received less support, but a review of actual social support transactions found that posts about depression symptoms were associated with increased social support from Facebook connections [24].

Despite growing evidence about the objective use of Facebook and other social network sites to share emotions/experiences related to depression, relatively less is known about how young people perceive those experiences [25]. The purpose of this study was to determine how college students *perceived* their help-seeking posts about depression and the types of responses they received.

Methods

Design

Data for this analysis were extracted from a comprehensive mixed method, multisite, cross-sectional study conducted between September and November 2012. The purpose of the overall study was to assess multiple health outcomes among a college student population, including mental health and technology-related behaviors. Data were collected via an online survey with college students from four universities in the United States. The four universities were selected to provide geographic variation and a mix of public and private schools. Of these, three were in the Midwest and one was in the West. All participants in the overall study were invited to respond to all relevant questions, including the qualitative questions reported in this paper. As the Patient Health Questionnaire-9 (PHQ-9) includes questions regarding suicidal thoughts over the past 2 weeks, students were provided a list of Web-based mental health resources as a condition of the institutional review board (IRB) approval. This study was approved by each university’s affiliated IRB.

Procedures

Paper fliers with a link to the online Catalyst WebQ survey (https://itconnect.uw.edu/learn/tools/catalyst-web-tools/webq/) were distributed to 662 college students taking biology, communications, psychology, and nursing classes during the Autumn semester in 2012. These specific classes were selected to be able to reach a diverse set of students, as many were introductory classes required across different majors in order to be able to reach a diverse set of students. Each college class that was included in the study had a general Listserv that the professors used to communicate with students. In all, two to three follow-up emails were sent via the course Listserv to all 662 students who received the flier. Students who accessed the survey were required to complete an informed consent. A US $5 Starbucks gift card was given to those who completed the survey. In one class, students who completed the survey received extra credit in addition to the gift card. To complete the survey, students were required to be between the age of 18 and 23 years, and this inclusion criterion was included in recruitment materials. Any respondents outside this age range were excluded from this analysis. The response rate for the overall survey was 43.3% (n=287). The final sample size for this study was 33 students, which represented 11% of the total sample.

Sample

The sample for this study included all participants (n=33) who provided qualitative responses to the survey question on the nature of their and their friends’ responses to their Facebook posts and who had completed the questions on the PHQ-9.

A *t* test was used to determine if there were demographic and mental health differences between this study sample (n=33) and the sample of students who did not complete the open-ended questions (n=253). There was a statistically significant difference (P=.001) in the mean score on PHQ-9 between those who completed the qualitative Facebook questions (mean 9.06, SD 6.43) and those who did not complete those questions (mean 4.92, SD 4.25). No other statistically significant differences were identified between the full sample and the subset of participants in this study.

Measures

Demographic data were collected from each participant, including age, race/ethnicity, relationship status, sexual orientation, and year in school.

Depression

Depression was measured using the PHQ-9. The PHQ-9 score was based on the instrument guidelines with the following categorization: <4=minimal depression, 5-9=mild depression, 10-14=moderate depression, 15-19=moderately severe depression, and 20-27=severe depression. A total score of 10 or more points was used in this study to indicate at least moderate depression systems, which was consistent with the prior use of the questionnaire [26]. The PHQ-9 was chosen to measure depression in this study because it has been validated in both adolescent and adult populations [27,28].

Qualitative Facebook questions

Participants were asked to identify if they had ever reached out on Facebook when feeling depressed and if yes, to describe the post(s). Participants were then asked to describe the nature of comments their friends made in response to a post, where they relayed some variation of feelings of depression. A copy of the questions are presented in Multimedia Appendix 1. The qualitative Facebook questions were developed by a study investigator, who based these questions on an internet survey about support-seeking behaviors used by adolescents when they were depressed [29]. Other researchers in the team then conducted an initial review of the Facebook questions for face validity and assessed them for question clarity. Prior to implementing the survey, the questions were also vetted by several college students.

Analyses

The first purpose of the study was to identify themes in participant perceptions regarding depression-related Facebook posts and their friend's responses. Classic content analysis [30-32] was used to examine responses to the two open-ended questions. Three investigators defined categories a priori (see Table 1). Investigators then participated in three rounds of practice coding. Discrepancies were discussed after each round of coding. An interrater agreement of 0.79 was obtained on a subset of 10 quotes before coding of the full subsample. The remaining data were divided among three coders and an open coding approach was used to categorize the total sample of open-ended responses. As ambiguous or unclear cases emerged, they were discussed and assigned the category agreed up on by all three coders. The final interrater agreement on all 33 statements was 0.90.

Table 1. Qualitative themes and definitions for participant posts and their friends' responses.

Themes	Definition
Participant post	
Emotion	Emotion or feeling that does not use the word depressed.
Bad day	Reference to having a bad day or a bad one-time experience doing something.
Song	Lyrics of a song or mention of a music video.
Private message	A type of message that is only available to the recipient via Facebook, email, phone, text.
Depressed mood	Reference to symptoms of a major depressive episode, which included (1) depressed mood, (2) decreased interest or pleasure in day-to-day activities, (3) increase/decrease in appetite, (4) increased/decreased sleep, (5) psychomotor agitation or retardation, (6) loss of energy, (7) feelings of guilt, worthlessness, negative self-appraisal, (8) indecisiveness, (9) difficulty in concentrating, (10) recurrent thoughts of death or suicidal ideation.
Emoji/emoticon	The emoticon of a sad face.
Quote	The use of someone other than the profile owner's words to express emotions.
Joke	Making fun of depression or being sarcastic.
Asked for help	Follow-up the content of the post with saying need help.
Friend response	
Support or motivating gestures	An individual saying supporting or motivating words to the participant.
Asked a question	Questioned individual on how they were doing.
Contact in alternative communication	An individual contacted the participant using some other form of communication, such as email, text, phone call, etc.
Liked	An individual clicked like to their post or commented on Facebook.
Private message	Contact the individual through nonpublic methods.
Empathize	Understand and feel the feelings of another, "walk a mile in their shoes."
Sympathize	Compassion or commiserating with another, "I'm sorry that you are feeling this way."
Negative	An individual responding in an unsupportive way.
No response	Received no response from friends after they posted.

We conducted bivariate analyses to assess if the theme of friend's response differed by the theme(s) of the participant's post. We also conducted bivariate analyses to assess differences between themes of the participant's posts, by depression

severity. To test for statistical significance with these analyses, while accounting for the small sample size and cell counts, we used Fisher exact test.

Results

The mean age of the 33 participants was 20.2 (SD 1.4) years. The majority of participants were female (n=28, 85%) and Caucasian (n=26, 79%). Participants were at different stages in their studies, with 24% freshman (n=8), 18% sophomore (n=6), 36% junior (n=12), and 24% senior (n=8). Almost half of the participants were single (n=16, 48.5%), and only one person identified himself as homosexual.

The mean score on the PHQ-9 was 9.06 (SD 6.43; n=33). Based on the normed scoring criteria, those who scored 10 or more points were considered to be clinically depressed. Of note, 12% had minimal depression (n=5), 33% had mild depression (n=11), 27% had moderate depression (n=9), 6% had moderate to severe depression (n=2), and 12% had severe depression (n=4); three respondents had a score of 0 on the PHQ-9. Question 9 on the PHQ-9 assesses suicidal thoughts in the last 2 weeks. Of these, 33% indicated suicidal thoughts on several days (n=11) and 6%, nearly every day (n=2).

We used a multiple-response analysis, where participants could identify multiple ways in which they had posted about their depression on Facebook. Of these, 24% (n=8) most commonly reported that they posted their posting about their emotions, and 18% (n=6) referenced a bad day (Table 2). A total of four participants (12%) referenced a specific form of the word depression or an associated symptom. Of the participants, 15% (n=5) referenced to lyrics of a song or a music video and 6% (n=2) included quote from a song. Of these, 12% (n=4) reported that their friends responded via alternative communication approaches or through a private message; 6% (n=2) included an emoji/emoticon. Only 1% participant reported directly seeking help.

Table 2. Number and percentage of themes present in participant's posts

Theme	n (%)[a]	Example quote
Emotion	8 (24)	"I'm tired of being treated like crap, etc." "I just said I felt so alone."
Bad day	6 (18)	"I made a couple of posts over the course of the past year. In one I posted something similar to 'Oh man, something's gotta give.' Another one I remember saying 'I'm calling all you angels.'" "Terrible day. Things couldn't get any worse."
Song	5 (15)	"Something with sad music, or a comment on how frustrated or hard life can be." "I posted a music video that correlated with my mood that day."
Private message	4 (12)	"I didn't post something about it, I private messaged a friend for some advice." "I did not post publicly. I talked to my best friend via Facebook messaging. I talked to her about my problems and what I should do."
Depressed mood	4 (12)	"Depressed sounding post, which led to friends helping." "Just a sad status."
Emoji/emoticon	2 (6)	"Generic sad faces or 'screw this' posts." "Sad face."
Quote	2 (6)	"Just posted a quote about being stressed."
Joke	1 (3)	"I have posted jokingly about particular days being emotionally difficult."
Asked for help	1 (3)	"If anyone could help me get through a problem."
Other	2 (6)	"It's probably one of the hardest things when you wanna help someone but you know you can't. And you know you should just worry about yourself but you just can't help it."

[a]Qualitative responses may have included multiple themes, so the column does not add up to 100%.

The most common perceived response to the participant's Facebook depression post was gestures of support or motivation (n=13, 39%; Table 3). Almost 15% of the participants (n=4) mentioned that their friends responded via alternative communication approaches or through a private message. A total of four participants (12%) discussed receiving "likes." One described a negative response to the depression-related display. No participants reported that their friends encouraged them to seek help.

Table 3. Number and percentage of themes present in friends' responses.

Theme	n (%)[a]	Example quote
Support or motivating gestures	13 (39)	"All positive, sweet, telling me I will get through anything." "All my close friends were there to encourage me and letting me know that everything will be okay."
Asked a question	7 (21)	"If they are generic posts friends generally have to ask what is wrong. It is hard to tell who cares or who's curious this way though." "They ask what's wrong or text, or message me to talk."
Contact in alternative communication	4 (12)	"My best friend called me when she got the message and talked to me for a long time." "She called me immediately."
Liked	4 (12)	"A few people 'liked' it." "People 'liked' the music video."
Private message	4 (12)	"They make plans to hang with me and take my mind off of it."
Empathize	2 (6)	"They would respond most often by agreeing with me and saying that they were having a hard time as well."
Sympathize	1 (3)	"Tried to make me feel better, 'hang in there,' etc."
Negative	1 (3)	"Negatively"
No response	1 (3)	"Usually they don't respond..."

[a]Qualitative responses may have included multiple themes, so percentages do not add up to 100%.

We also explored the relationship between individuals' Facebook post and their friends' perceived response. When a person posted about having a "bad day," 83% of his/her friends responded with support or motivation and 17% responded with sympathy. When he/she posted a song, song lyrics, music video, 40% received a response coded as support or motivation, and 60% "liked" the post. When the person made an emotional post, 62% received a support/motivation response, 13% were asked to communicate outside of Facebook, and 25% asked a question. When the person wrote that he/she had some depression or depression symptoms, half of his/her friends responded with a negative message and the other half used a private message to communicate. Finally, if the person made a joke about being depressed, he/she received an empathetic response. There was a significant relationship ($P<.001$) between how a person reached out on Facebook and his/her friend's perceived response. We also explored the relationship between the PHQ-9 category and the type of post the person made on Facebook (Table 4). The themes of the participant posts varied by depression severity.

Table 4. Patient Health Questionnaire-9 categories by theme of participant post.

Depression severity	Song, n (%)	Private message, n (%)	Depressed mood, n (%)	Help, n (%)	Emoji, n (%)	Joke, n (%)	Bad day, n (%)	Emotion, n (%)	Quote, n (%)	Other, n (%)
Minimal (n=5)	0 (0)	1 (20)	1 (20)	1 (20)	0 (0)	0 (0)	1 (20)	1 (20)	0 (0)	0 (0)
Mild (n=11)	1 (9)	2 (18)	1 (9)	0 (0)	1 (9)	1 (9)	1 (9)	2 (18)	1 (9)	1 (9)
Moderate (n=9)	1 (11)	0 (0)	1 (11)	0 (0)	1 (11)	0 (0)	1 (11)	44 (44)	0 (0)	1 (11)
Moderate to severe (n=2)	1 (50)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (50)	0 (0)	0 (0)	0 (0)
Severe (n=4)	0 (0)	1 (25)	0 (0)	0 (0)	0 (0)	0 (0)	2 (50)	0 (0)	1 (25)	0 (0)

Discussion

Overview

The purpose of this study was to determine how college students *perceived* their help-seeking posts about depression and the types of responses received. In our sample, no participants reported using specific indicators of depression symptoms, as defined in the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV; eg, guilt/worthlessness, hopeless, depressed mood) [33], which is contrary to the findings of two studies on Facebook posts by depressed individuals that found that some individuals used a language consistent with diagnosable depression symptoms [20,34]. Since our study focused on how individuals perceived their posts, this difference between our studies and the other studies suggests that participants in our study had altered perceptions of the content of and responses to their posts, which is consistent with some prior research on online support seeking [24]. In one prior study, more severe depression symptoms were associated with significantly reduced perceptions of Facebook-related support, despite increased levels of support found in the review of the text of the interactions [24]. A larger body of research has demonstrated a negative bias in how individuals with depression perceive social support and other social perceptions in offline interactions [35-37]. Alternatively, social desirability bias may

account for the difference between participants' perceptions and review of social media posts in prior studies [20,24,38]. Students may not report using specific depression symptoms due to concerns about social stigma. However, it is unclear why participants in this study would report posting about reaching out on Facebook regarding depression and then withhold that they used depression-specific language.

In addition to triangulating prior studies of Facebook-related mental health support, our study has several implications. In 2017, Facebook implemented a new mechanism for concerned friends and family to provide support for individuals who may be experiencing suicidal thoughts [39]. After a report is made, Facebook will send a message indicating someone expressed concern and will provide a list of resources (eg, suicide and crisis hotlines). Facebook also changed the "Like" function, so that individuals can indicate their reactions (eg, "Like," "Happy," "Wow," "Sad," and "Angry"), rather than simply "liking" a post. This change may alter how participants perceive support, as approximately 12% of participants reported friends provided support by liking posts. With this increased specificity, friends may choose a reaction, rather than posting a text response. Additional research is necessary to determine how these changes impact the objective support given and perceptions of support received.

Increasing mental health literacy may be another way to improve how friends provide support and individuals request/receive mental health-related support. Increasing mental health literacy has been shown to have a positive effect on reducing mental health stigma and may encourage friends to respond to depressive posts in more helpful ways [40-42]. In addition, Armstrong and Young [40] found that college students in their study preferred that mental health information be delivered via the internet, public service announcements, and the media. Providing health education and positive social support through social media or other online platforms may be one approach to support individuals experiencing depression.

Limitations

This study was limited by several factors. First, the study was cross-sectional and used a convenience sample. The sample size was small (n=33), which also limits the generalizability of the findings. Thus, the understanding of an individual's perceptions of support seeking and receiving may not generalize to all Facebook users, but are more likely to be generalizable to individuals with depression who have posted about or sought support related to depression, particularly when triangulated with findings from other research studies. In addition, the individuals who responded to the open-ended question about depression-related posts and were thus included in this analysis had higher depression scores than other individuals in the overall multisite study. Thus, perceptions of online support seeking and receipt are likely most generalizable to an individual with more significant depression symptoms. Finally, interactions on Facebook may not be generalizable to other social media platforms. Some young people are likely to disclose sensitive information on social media platforms other than Facebook, due to privacy concerns and concerns about parental presence on Facebook [26,43].

Conclusions

This study explored how individuals perceived support seeking and their perceptions of the responses. The findings, when combined with those of prior studies, indicate that individuals may perceive their depression-related posts and the related responses differently from how others perceive them. As such, there may be a need to increase mental health literacy to help individuals, including both support seekers and support providers, communicate about mental health. Since stigma is often related to a lack of mental health literacy, enhancing mental health literacy may be an approach to reduce mental health stigma and increase support to individuals with depression symptoms.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study questions about depression posts on Facebook.
[DOCX File , 22 KB - formative_v4i7e13650_app1.docx]

References

1. Gallagher RP. National Survey of Counseling Center Directors, 2006 (Monograph Series No. 8P). Alexandria: International Association of Counseling Services, Inc.; 2007.
2. American College Health Association. American College Health Association: National College Health Assessment II: Reference Group Data Report Fall 2008. Baltimore, MD: American College Health Association; 2009.
3. Zivin K, Eisenberg D, Gollust SE, Golberstein E. Persistence of mental health problems and needs in a college student population. J Affect Disord 2009;117(3):180-185 [FREE Full text] [doi: 10.1016/j.jad.2009.01.001] [Medline: 19178949]
4. Hunt J, Eisenberg D. Mental health problems and help-seeking behavior among college students. J Adolesc Health 2010;46(1):3-10 [FREE Full text] [doi: 10.1016/j.jadohealth.2009.08.008] [Medline: 20123251]
5. Kessler RC, Foster CL, Saunders WB, Stang PE. Social consequences of psychiatric disorders, I: educational attainment. Am J Psychiatry 1995;152(7):1026-1032 [FREE Full text] [doi: 10.1176/ajp.152.7.1026] [Medline: 7793438]

6. Bramesfeld A, Platt L, Schwartz FW. Possibilities for intervention in adolescents' and young adults' depression from a public health perspective. Health Policy 2006;79(2-3):121-131 [FREE Full text] [doi: 10.1016/j.healthpol.2005.11.012] [Medline: 16414145]
7. Deas D, Brown ES. Adolescent substance abuse and psychiatric comorbidities. J Clin Psychiatry 2006;67(7):18-23 [FREE Full text] [doi: 10.4088/jcp.0706e02] [Medline: 17107227]
8. Rao U. Links between depression and substance abuse in adolescents: neurobiological mechanisms. Am J Prev Med 2006;31(6):S161-S174 [FREE Full text] [doi: 10.1016/j.amepre.2006.07.002] [Medline: 17175411]
9. Garlow SJ, Rosenberg J, Moore JD, Haas AP, Koestner B, Hendin H, et al. Depression, desperation, and suicidal ideation in college students: results from the American Foundation for Suicide Prevention College Screening Project at Emory University. Depress Anxiety 2008;25(6):482-488 [FREE Full text] [doi: 10.1002/da.20321] [Medline: 17559087]
10. Rao U, Chen LA. Characteristics, correlates, and outcomes of childhood and adolescent depressive disorders. Dialogues Clin Neurosci 2009;11(1):45-62. [Medline: 19432387]
11. Eisenberg D, Golberstein E, Gollust E. Help-seeking and access to mental health care in a university student population. Med Care 2007;45(7):594-601 [FREE Full text] [doi: 10.1097/MLR.0b013e31803bb4c1] [Medline: 17571007]
12. Pew Research Center. 2018. Internet/Broadband Fact Sheet URL: http://www.pewinternet.org/fact-sheet/internet-broadband/ [WebCite Cache ID 6oNny1FMh]
13. Pew Research Center. 2018. Social Media Use in 2018 URL: http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/ [WebCite Cache ID 6ykw0G0Gc]
14. Zhang R. The stress-buffering effect of self-disclosure on Facebook: an examination of stressful life events, social support, and mental health among college students. Comput Hum Behav 2017;75:527-537 [FREE Full text] [doi: 10.1016/j.chb.2017.05.043]
15. Shensa A, Sidani JE, Escobar-Viera CG, Chu KH, Bowman ND, Knight JM, et al. Real-life closeness of social media contacts and depressive symptoms among university students. J Am Coll Health 2018;66(8):747-753 [FREE Full text] [doi: 10.1080/07448481.2018.1440575] [Medline: 29452042]
16. Saunders JF, Eaton AA. Snaps, selfies, and shares: how three popular social media platforms contribute to the sociocultural model of disordered eating among young women. Cyberpsycho Behav Soc Netw 2018;21(6):343-354 [FREE Full text] [doi: 10.1089/cyber.2017.0713] [Medline: 29883209]
17. Vogel EA, Rose JP, Okdie BM, Eckles K, Franz B. Who compares and despairs? The effect of social comparison orientation on social media use and its outcomes. Pers Individ Differ 2015;86:249-256 [FREE Full text] [doi: 10.1016/j.paid.2015.06.026]
18. Frison E, Eggermont S. Exploring the relationships between different types of Facebook use, perceived online social support, and adolescents' depressed mood. Soc Sci Comput Rev 2015;34(2):153-171 [FREE Full text] [doi: 10.1177/0894439314567449]
19. Murrieta J, Frye CC, Sun L, Ly LG, Cochancela CS, Eikey EV. Depression: findings from a literature review of 10 years of social media and depression research. In: Chowdhury G, McLeod J, Gillet V, Willett P, editors. Transforming Digital Worlds. iConference 2018. Lecture Notes in Computer Science. Cham: Springer; 2018:47-56.
20. Moreno MA, Jelenchick LA, Egan KG, Cox E, Young H, Gannon KE, et al. Feeling bad on Facebook: depression disclosures by college students on a social networking site. Depress Anxiety 2011;28(6):447-455 [FREE Full text] [doi: 10.1002/da.20805] [Medline: 21400639]
21. Gowen K, Deschaine M, Gruttadara D, Markey D. Young adults with mental health conditions and social networking websites: seeking tools to build community. Psychiatr Rehabil J 2012;35(3):245-250 [FREE Full text] [doi: 10.2975/35.3.2012.245.250] [Medline: 22246123]
22. Radovic A, Gmelin T, Stein BD, Miller E. Depressed adolescents' positive and negative use of social media. J Adolesc 2017;55:5-15 [FREE Full text] [doi: 10.1016/j.adolescence.2016.12.002] [Medline: 27997851]
23. Egan KG, Koff RN, Moreno MA. College students' responses to mental health status updates on Facebook. Issues Ment Health Nurs 2013;34(1):46-51 [FREE Full text] [doi: 10.3109/01612840.2012.718044] [Medline: 23301569]
24. Park J, Lee DS, Shablack H, Verduyn P, Deldin P, Ybarra O, et al. When perceptions defy reality: the relationships between depression and actual and perceived Facebook social support. J Affect Disord 2016;200:37-44 [FREE Full text] [doi: 10.1016/j.jad.2016.01.048] [Medline: 27126138]
25. Cavazos-Rehg PA, Krauss MJ, Sowles S, Connolly S, Rosas C, Bharadwaj M, et al. A content analysis of depression-related tweets. Comput Human Behav 2016;54:351-357 [FREE Full text] [doi: 10.1016/j.chb.2015.08.023] [Medline: 26392678]
26. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
27. Richardson LP, McCauley E, Grossman DC, McCarty CA, Richards J, Russo JE, et al. Evaluation of the patient health questionnaire-9 item for detecting major depression among adolescents. Pediatrics 2010;126(6):1117-1123 [FREE Full text] [doi: 10.1542/peds.2010-0852] [Medline: 21041282]
28. Allgaier AK, Pietsch K, Frühe B, Sigl-Glöckner J, Schulte-Körne G. Screening for depression in adolescents: validity of the patient health questionnaire in pediatric care. Depress Anxiety 2012;29(10):906-913 [FREE Full text] [doi: 10.1002/da.21971] [Medline: 22753313]

29. Cash SJ, Bridge J. Meeting them where they are: an exploration of technology use and help-seeking behaviors among adolescents and young adults. 2012 Presented at: Poster presented at: 59th Annual Meeting of the American Academy of Child & Adolescent Psychiatry; Oct 25, 2012; San Francisco, CA.
30. Krippendorf K. Content Analysis: An Introduction to its Methodology, 2nd edition.th edition. Thousand Oaks, CA: Sage; 2004. ISBN: 978-0761915454.
31. Miles M, Huberman M. Qualitative Data Analysis: An Expanded Sourcebook. Thousand Oaks, CA: Sage; 1994. ISBN: 978-0803955400.
32. Weber RP. Basic Content Analysis, 2nd editionth edition. Thousand Oaks, CA: Sage; 1990.
33. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders., 4th edition. Text Revisionth edition. Washington, DC: American Psychiatric Association; 2000.
34. Moreno MA, Jelenchick LA, Kota R. Exploring depression symptom references on Facebook among college freshman: a mixed methods approach. Open J Depress 2013;2(3):35-41 [FREE Full text] [doi: 10.4236/ojd.2013.23008]
35. Holahan CJ, Moos RH, Holahan CK, Cronkite RC, Randall PK. Unipolar depression, life context vulnerabilities, and drinking to cope. J Consult Clin Psychol 2004;72(2):269-275 [FREE Full text] [doi: 10.1037/0022-006X.72.2.269] [Medline: 15065960]
36. Rookm KS. The negative side of social interaction: impact on psychological well-being. J Pers Soc Psychol 1984;46(5):1097-1108 [FREE Full text] [doi: 10.1037/0022-3514.46.5.1097]
37. Davila J, Hershenberg R, Feinstein JA, Gorman K, Bhatia V, Starr LR. Frequency and quality of social networking among young adults: associations with depressive symptoms, rumination, and corumination. Psychol Pop Media Cult 2012;1(2):72-86 [FREE Full text] [doi: 10.1037/a0027512] [Medline: 24490122]
38. Krumpal I. Determinants of social desirability bias in sensitive surveys: a literature review. Qual Quant 2013;47(4):2025-2047 [FREE Full text] [doi: 10.1007/s11135-011-9640-9]
39. Facebook. 2019. Report Suicidal Content URL: https://www.facebook.com/help/contact/305410456169423
40. Armstrong LL, Young K. Mind the gap: person-centred delivery of mental health information to post-secondary students. Psychosoc Interven 2015;24(2):83-87 [FREE Full text] [doi: 10.1016/j.psi.2015.05.002]
41. Canadian Alliance on Mental Illness and Mental Health. 2007 . Mental Health Literacy in Canada: Phase One Draft Report Mental Health Literacy Project URL: http://camimh.ca/wp-content/uploads/2012/04/Mental-Health-Literacy_-_Full-Final-Report_EN.pdf
42. Keys S, Cash S. Reachout.com report to California Mental Health Association. San Francisco, CA; 2012.
43. Mullen C, Hamilton NF. Adolescents' response to parental Facebook friend requests: he comparative influence of privacy management, parent-child relational quality, attitude and peer influence. Comput Human Behav 2016;60:165-172 [FREE Full text] [doi: 10.1016/j.chb.2016.02.026]

Abbreviations

IRB: institutional review board
PHQ-9: Patient Health Questionnaire-9

Edited by A Tsanas; submitted 07.02.19; peer-reviewed by J Sidani, K Kaipainen, A Benetoli, S Pollak; comments to author 26.03.19; revised version received 24.06.19; accepted 18.11.19; published 17.07.20.

Please cite as:
Cash S, Schwab-Reese LM, Zipfel E, Wilt M, Moreno M
What College Students Post About Depression on Facebook and the Support They Perceive: Content Analysis
JMIR Form Res 2020;4(7):e13650
URL: https://formative.jmir.org/2020/7/e13650
doi:10.2196/13650
PMID:32706687

Original Paper

Effectiveness of a Voice-Based Mental Health Evaluation System for Mobile Devices: Prospective Study

Masakazu Higuchi[1*], PhD; Mitsuteru Nakamura[2*], PhD; Shuji Shinohara[1*], PhD; Yasuhiro Omiya[3*], PhD; Takeshi Takano[3*], ME; Shunji Mitsuyoshi[1*], PhD; Shinichi Tokuno[1*], MD, DMCC, PhD

[1]Department of Bioengineering, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan

[2]Department of Pharmacology, School of Medicine, The University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

[3]Research and Product Development Department, PST Inc, Kanagawa, Japan

[*]all authors contributed equally

Corresponding Author:
Masakazu Higuchi, PhD
Department of Bioengineering
Graduate School of Engineering
The University of Tokyo
7-3-1 Hongo Bunkyo-ku
Tokyo,
Japan
Phone: 81 3 5841 3432
Email: higuchi@bioeng.t.u-tokyo.ac.jp

Abstract

Background: We developed a system for monitoring mental health using voice data from daily phone calls, termed Mind Monitoring System (MIMOSYS), by implementing a method for estimating mental health status from voice data.

Objective: The objective of this study was to evaluate the potential of this system for detecting depressive states and monitoring stress-induced mental changes.

Methods: We opened our system to the public in the form of a prospective study in which data were collected over 2 years from a large, unspecified sample of users. We used these data to analyze the relationships between the rate of continued use, the men-to-women ratio, and existing psychological tests for this system over the study duration. Moreover, we analyzed changes in mental data over time under stress from particular life events.

Results: The system had a high rate of continued use. Voice indicators showed that women have more depressive tendencies than men, matching the rate of depression in Japan. The system's voice indicators and the scores on classical psychological tests were correlated. We confirmed deteriorating mental health for users in areas affected by major earthquakes in Japan around the time of the earthquakes.

Conclusions: The results suggest that although this system is insufficient for detecting depression, it may be effective for monitoring changes in mental health due to stress. The greatest feature of our system is mental health monitoring, which is most effectively accomplished by performing long-term time-series analysis of the acquired data considering the user's life events. Such a system can improve the implementation of patient interventions by evaluating objective data along with life events.

(JMIR Form Res 2020;4(7):e16455) doi:10.2196/16455

KEYWORDS

mental health; monitoring system; stress evaluation; voice analysis

Introduction

The importance of mental health care is increasing globally in today's "society of stress." Stress negatively affects mood and health in everyday life, and accumulated stress leads to psychiatric and behavioral disorders [1]. These disorders are associated with major societal economic losses owing to lowered lifetime earnings and a decrease in labor productivity [2,3]. Early interventions for depression lead to a higher remission rate [4]; therefore, techniques to easily screen for depression and stress for the early detection of poor mental health are in demand.

Research has been conducted on screening methods for patients with mental health issues using biomarkers such as saliva [5] and blood [6]; however, these are costly as they require special measurement equipment or reagents and are invasive. Thus, noninvasive self-report psychological tests such as the General Health Questionnaire [7,8] and the Beck Depression Inventory (BDI) [9] are generally used. Although these tests are comparatively simple, the effect of reporting bias [10], in which respondents selectively underestimate or overestimate specific information either consciously or unconsciously, cannot be eliminated.

There is empirical evidence showing that mood changes are apparent in one's facial expression and voice, and there is ongoing research estimating the depressive state and stress state using these indicators [11-14]. In addition to being noninvasive, an analysis using voice data has the particular advantage of simplicity, as it does not require specialized equipment and can be used remotely. This method also potentially solves the problems of detecting various psychiatric illnesses and eliminating the reporting bias present in self-report psychological tests. For these reasons, voice-based techniques have garnered attention in recent years.

It is desirable to continuously monitor individuals' status to detect poor mental health as early as possible. Recently, the capabilities of mobile device platforms have rapidly improved, and smartphones can perform moderate-level arithmetic processing. Furthermore, smartphones have already become an indispensable part of our daily lives, making them the optimal tool for continuously and noninvasively monitoring elements such as users' biological information. As such, research utilizing smartphones to detect stress and develop mental health apps for mobile terminals is actively being pursued [15-17].

In this study, we developed a system for monitoring mental health using voice data from daily phone calls, termed Mind Monitoring System (MIMOSYS) [18], by implementing a method for estimating mental health from voice data [12] through Android smartphones. It is hoped that using MIMOSYS will allow for the monitoring of daily mental health levels and the ability to circumvent conditions of poor mental health such as depression before they arise. This system was made available to the public as a prospective study in which MIMOSYS data were collected over a period of 2 years from an unspecified large sample of users who provided consent for research participation. The aim of this study was to use the data obtained to verify the effectiveness of MIMOSYS for detecting depressive states and monitoring stress-induced mental changes.

Methods

Ethical Considerations

This study was conducted with approval of the Research Ethics Committee of the Faculty of Medicine of the University of Tokyo (no. 10860).

MIMOSYS

The method of voice analysis used by MIMOSYS is based on sensibility technology [19], which analyzes patterns of change in the fundamental frequency of voice data to indicate the extent of the emotions "calmness," "anger," "joy," "sorrow," and "excitement" included therein. MIMOSYS uses the emotions analyzed by sensibility technology to calculate "vitality," the quantified mental health at the time of the call, and "mental activity," which includes the moving average and variation in vitality measured over a 2-week period prior to the call. Mental activity is an index that is useful in monitoring changes in mental health over the mid to long term. Moving averages were taken from the previous 2 weeks because the Diagnostic and Statistical Manual of Mental Disorders-IV [20] diagnostic criteria for depression state that primary symptoms must persist for at least 2 weeks.

Vitality and mental activity were calculated using algorithms as values between 0 and 1 but were multiplied by 100 and displayed over the range of 0-100 in the app to make it easier for users to understand. Extremely low or high values were judged to represent an abnormality in mental health.

The smallest unit of speech analyzed by sensibility technology is an "utterance"; that is, continuous voice data divided by breaths and other metrics. The start of an utterance is detected as the point at which a state of silence changes to a state of speech and continues for a set period of time. The end of an utterance is detected as the point at which a state of speech changes to a state of silence for a set period of time. Distinguishing between states of speech and silence is performed by thresholding the amplitude of temporal waveforms in voice data. MIMOSYS calculates vitality for each individual utterance in voice data and outputs the average of these values as the vitality for that voice data.

Data Collection

User consent for research participation was requested the first time MIMOSYS was used. Users who gave consent were assigned an anonymized personal ID that was saved to a dedicated server. Attribute data from a questionnaire were also saved when users registered. Attribute data included sex, age, medical history, history of present illness, and residence. All of these were self-declared. Each time a user made a call with their smartphone, MIMOSYS was automatically executed, and the analysis results were recorded consecutively on the same server. Voice data from the call were temporarily recorded on the smartphone. The analysis was performed simultaneously with the end of the call, and the voice data were immediately erased after the analysis results had been stored on the server. MIMOSYS also implemented a BDI test on the device screen every 3 months after initial use. This score was also recorded on the same server. Data were collected from July 20, 2015 to July 20, 2017.

The calls were private calls made by the users irregularly on their own smartphones. We did not impose any particular restrictions concerning the number of calls or intervals between calls.

Participants

There were approximately 3800 total downloads of MIMOSYS during the data collection period, and consent for research participation was obtained from 2462 users. Of these, 1814 users were aged 18 and older (maximum age 81 years), for

whom at least one call was analyzed by MIMOSYS. Our previous research showed that MIMOSYS does not depend on age [21], which is why this study involved users of a wide age range. MIMOSYS was downloaded by the users to their smartphones. The users had come to know about MIMOSYS through a variety of media; thus, nearly all had never met any of the researchers previously. As such, almost none of the users had attended lectures about how to use MIMOSYS in advance, and it is likely that many users had never been exposed to the BDI test conducted in the app previously. However, we assumed that they could navigate it easily as smartphone users. Anyone who agreed was qualified to participate. Users were not provided any particular reward except for being allowed to use MIMOSYS for free throughout the experimental period.

Data Analysis

The value for vitality becomes unstable if there are few utterances in the voice data. To ensure sufficient accuracy, empirically, it is desirable that voice data include at least 6 utterances. Furthermore, we have found that to improve the accuracy of mental activities, a mental activity averaged from at least 5 vitality scores is considered empirically effective. The data meeting the above conditions were extracted from MIMOSYS user data. The data from extremely short calls (less than 10 seconds) or with unsuitable call information were excluded. Therefore, there were 183,490 (123,860, 67.50% men) entries for analysis of effective vitality and 167,610 (113,600, 67.78% men) entries for analysis of effective mental activity. However, to match users for vitality and mental activity, vitality data were excluded for users for whom effective mental activities could not be calculated. The remaining data were used as the call dataset.

Furthermore, from the above data, the first vitality score measured after the initial BDI test and the most recent mental activity score within 2 weeks of the initial BDI test were selected for each user. These data were used as the user data set and included 1015 users (651, 64.14% men).

Analyzed Items

The following items were analyzed for the abovementioned datasets using the free statistical analysis software R version 3.4.2 [22] and G*Power version 3.1.9.2 [23].

MIMOSYS Rate of Continued Use

The period from the date and time each user consented to the study to the date and time of the most recent call analysis was calculated as the number of days used. Note that the number of days used was set as 0 for users who consented but did not have even one call analyzed, and for users who had a call analyzed but stopped use on the same day.

Normality of Vitality and Mental Activity Distributions

We examined the statistical features of the distributions of vitality and mental activity for the call dataset and for the user dataset by comparing them to the normal distribution.

Sex Differences in Vitality, Psychological Tests, and Vitality vs Mental Activity Comparison

The users in the user dataset were divided into those deemed healthy by the questionnaire and those deemed to have poor health, and sex differences in the vitality distribution were explored. Users were classified using the score from the initial BDI test executed by MIMOSYS and judged according to BDI evaluation criteria [24]. Based on these evaluation criteria, the foremost data after the BDI test of users with a BDI score of less than 11 were deemed to represent data from healthy individuals, whereas the foremost data after the BDI test of users with a BDI score of 11 or above were deemed to represent data from individuals with a mental health issue.

Moreover, we examined the correlation between user BDI and vitality/mental activity for the user dataset.

Temporal Changes in Vitality and Mental Activity Before and After a Stressful Life Event

On the nights of April 14 and 16, 2016, during the data collection period, there were two earthquakes of intensity 6 upper and 7 lower, with the epicenter in Kumamoto, Japan. According to the release by the Japanese Meteorological Agency, the earthquake on April 14 was a foreshock and that on April 16 was the main shock. Several of those affected by these earthquakes were MIMOSYS users, which allowed us to measure changes in vitality and mental activity before and after the disaster.

We then examined changes in vitality and mental activity around the time of the earthquake by area in Japan. We calculated an average of the vitality values in the 2-week periods before and after the earthquake. We also evaluated the most recent mental activity values from before the earthquake and from the 2-week period following the earthquake. The date of the earthquake was set as that of the foreshock on April 14, 2016.

Results

Participants

Table 1 shows the detailed participant information.

Table 1. Detailed information of MIMOSYS[a] participants (N=1814).

Characteristic	Value
Age (years), mean (SD)	39.97 (12.14)
BDI[b], mean (SD)	11.91 (10.37)
Number of calls, mean (SD)	121.10 (256.58)
Interval between calls (hours), mean (SD)	6.37 (7.73)
Sex (men), n (%)	1135 (62.57)
Medical history[c], n (%)	
Yes	855 (47.13)
Depression	226 (12.46)
History of present illness[c], n (%)	
Yes	573 (31.59)
Depression	138 (7.61)
Residence, n (%)	
Hokkaido	50 (2.76)
Tohoku	89 (4.91)
Kanto	979 (53.97)
Chubu	213 (11.74)
Kansai	204 (11.25)
Chugoku	71 (3.91)
Shikoku	27 (1.49)
Kyushu	151 (8.32)
Other	30 (1.65)

[a]MIMOSYS: Mind Monitoring System.

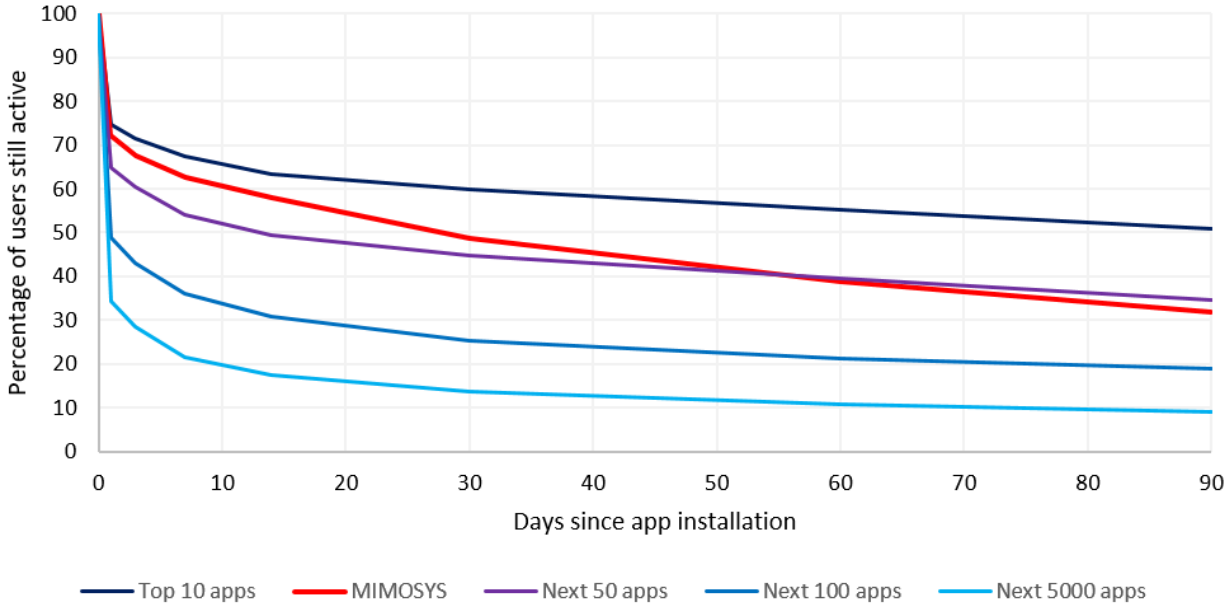
[b]BDI: Beck Depression Inventory.

[c]"Medical history" and "history of present illness" indicate the illness(es) the user has contracted in the past and presently, respectively. The item "yes" for these categories denotes the number of users who reported a specific illness(es), and the item "depression" denotes the number of users whose specific illnesses included depression.

MIMOSYS Rate of Continued Use

Figure 1 shows the rate of continued use curves for MIMOSYS and other smartphone apps. The rates of continued use for other smartphone apps referred to prior data [25]. The curves in Figure 1 for apps other than MIMOSYS show the average rate of continued use for the apps based on download rankings in Google Play. Since MIMOSYS has only been released in Japan, most MIMOSYS users live in Japan; however, a minority of people living abroad also downloaded the app in Japan.

Figure 1. Rate of continued use curves for Mind Monitoring System (MIMOSYS) and other smartphone apps.



Normality of Vitality and Mental Activity Distributions

Figure 2 shows the distribution of vitality of the call dataset. The mean of the full dataset was 0.56 (SD 0.094), whereas the mean for men and women was 0.57 (SD 0.095) and 0.53 (SD 0.084), respectively. The vitality distribution was lower for women than for men. Goodness-of-fit testing for the ideal normal distribution of the full set was significant with an effect size of 0.068 (P<.001) for the difference between the ideal and actual distributions [26]; therefore, the null hypothesis was rejected. Although it can be concluded that the vitality distribution does not follow a normal distribution, the effect size is extremely small. According to the effect size of evaluation criteria for the goodness-of-fit [26], there is an insignificant difference, and vitality can be regarded as having a normal distribution. Concerning the distribution for men and women separately, the goodness-of-fit effect size was 0.092 (P<.001) and 0.094 (P<.001), respectively. Thus, similar to the distribution of the full set, vitality was considered to have a normal distribution for both men and women.

Figure 2. Call dataset vitality distribution for the full set and by sex.

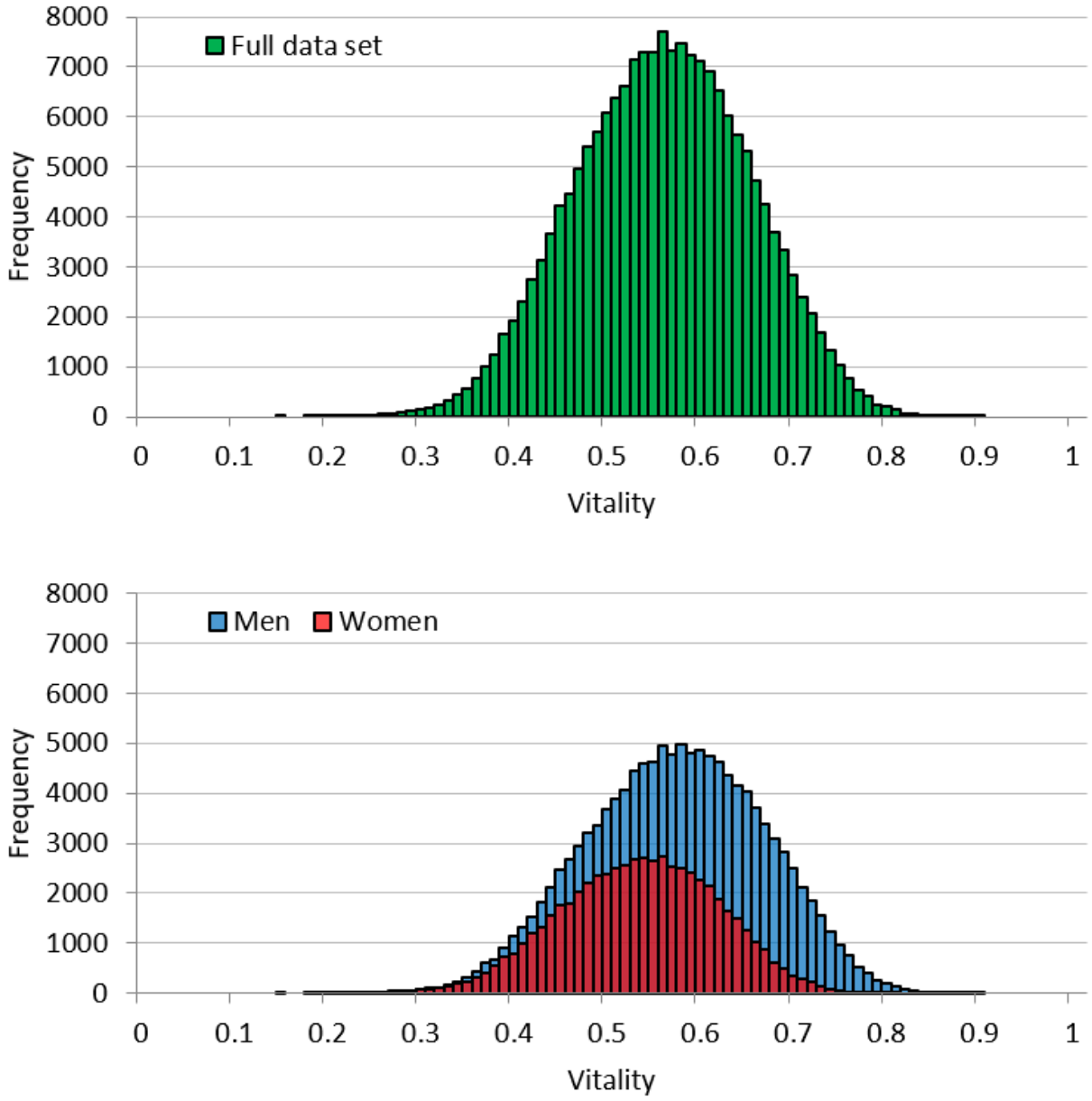


Figure 3 shows the vitality distribution for the user dataset. The mean vitality of the full dataset was 0.54 (SD 0.094), whereas the means for men and women were 0.56 (SD 0.096) and 0.52 (SD 0.085), respectively. As for the call dataset, the vitality distribution was lower for women than for men. Goodness-of-fit testing for the ideal normal distribution of the full set showed an effect size of 0.11 (P=.77), demonstrating that vitality follows a normal distribution. Regarding the distribution for men and women separately, the goodness-of-fit effect size was 0.15 (P=.68) and 0.23 (P=.39), respectively. Thus, similar to the distribution of the full set, it was concluded that vitality follows a normal distribution for both men and women.

Figure 3. User dataset vitality distribution for the full set and by sex.

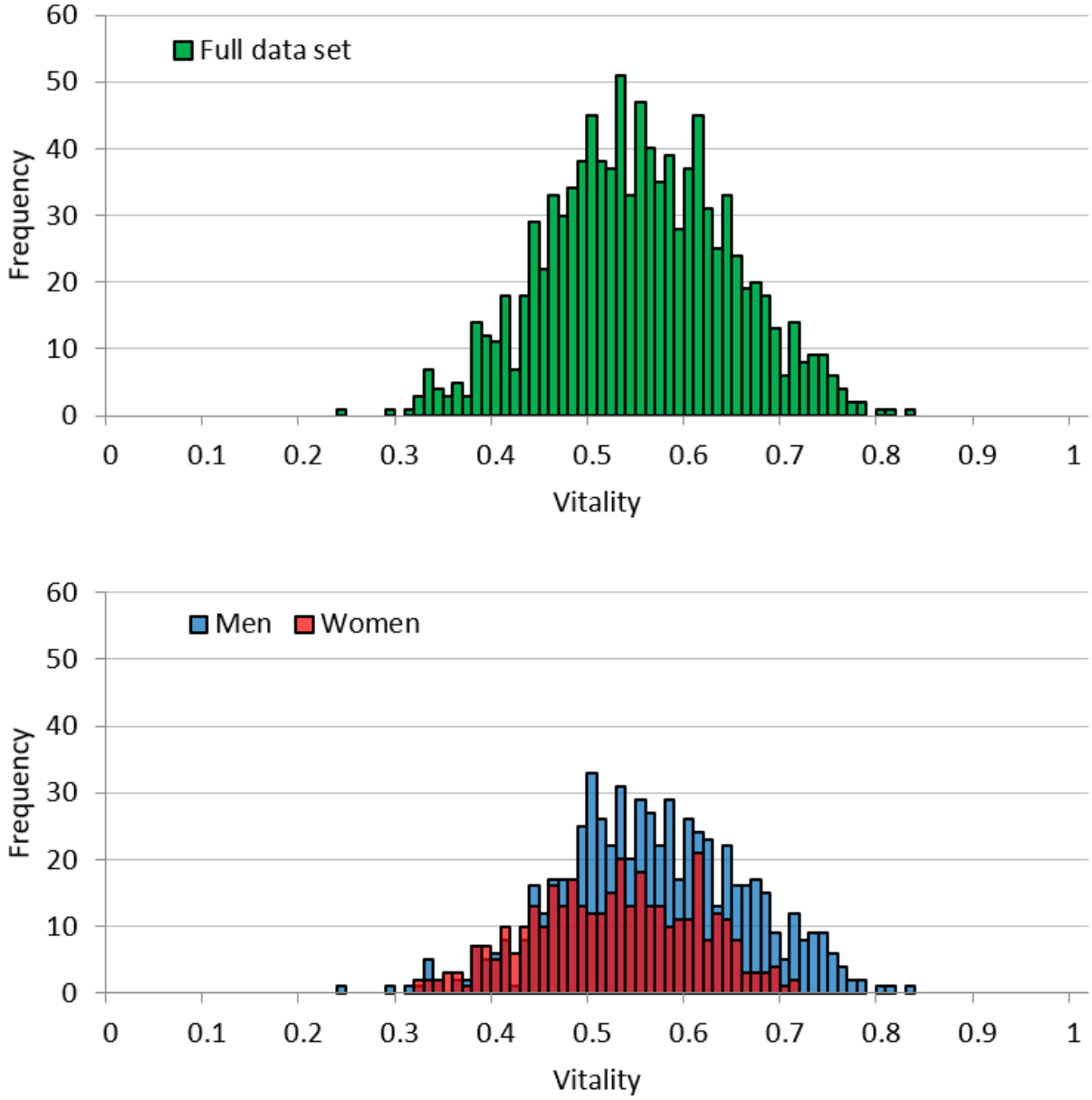


Figure 4 shows the mental activity distribution of the call dataset. The mean for the full dataset was 0.46 (SD 0.049), while the means for men and women were 0.47 (SD 0.050) and 0.44 (SD 0.040), respectively. As observed for vitality, the distribution for mental activity was lower for women than for men. Further, and also similar to vitality, goodness-of-fit testing for the normal distribution of the full set showed an effect size of 0.067 (P<.001). Therefore, mental activity was considered to follow a normal distribution. Regarding the distribution for men and women separately, the goodness-of-fit effect size was 0.085 (P<.001) and 0.058 (P<.001), respectively. Thus, mental activity was also regarded as following a normal distribution for both men and women.

Figure 4. Call dataset mental activity distribution for the full set and by sex.

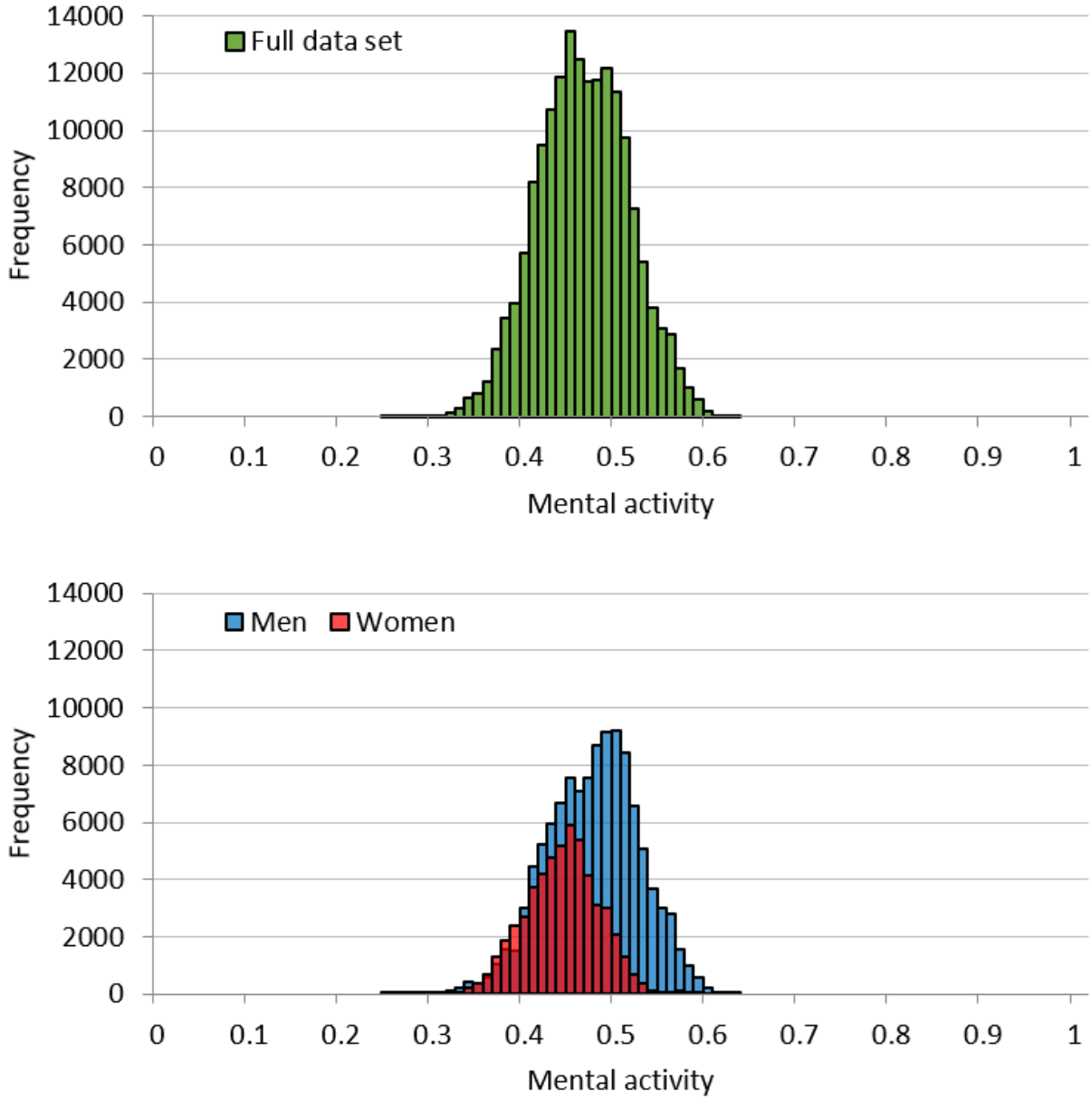
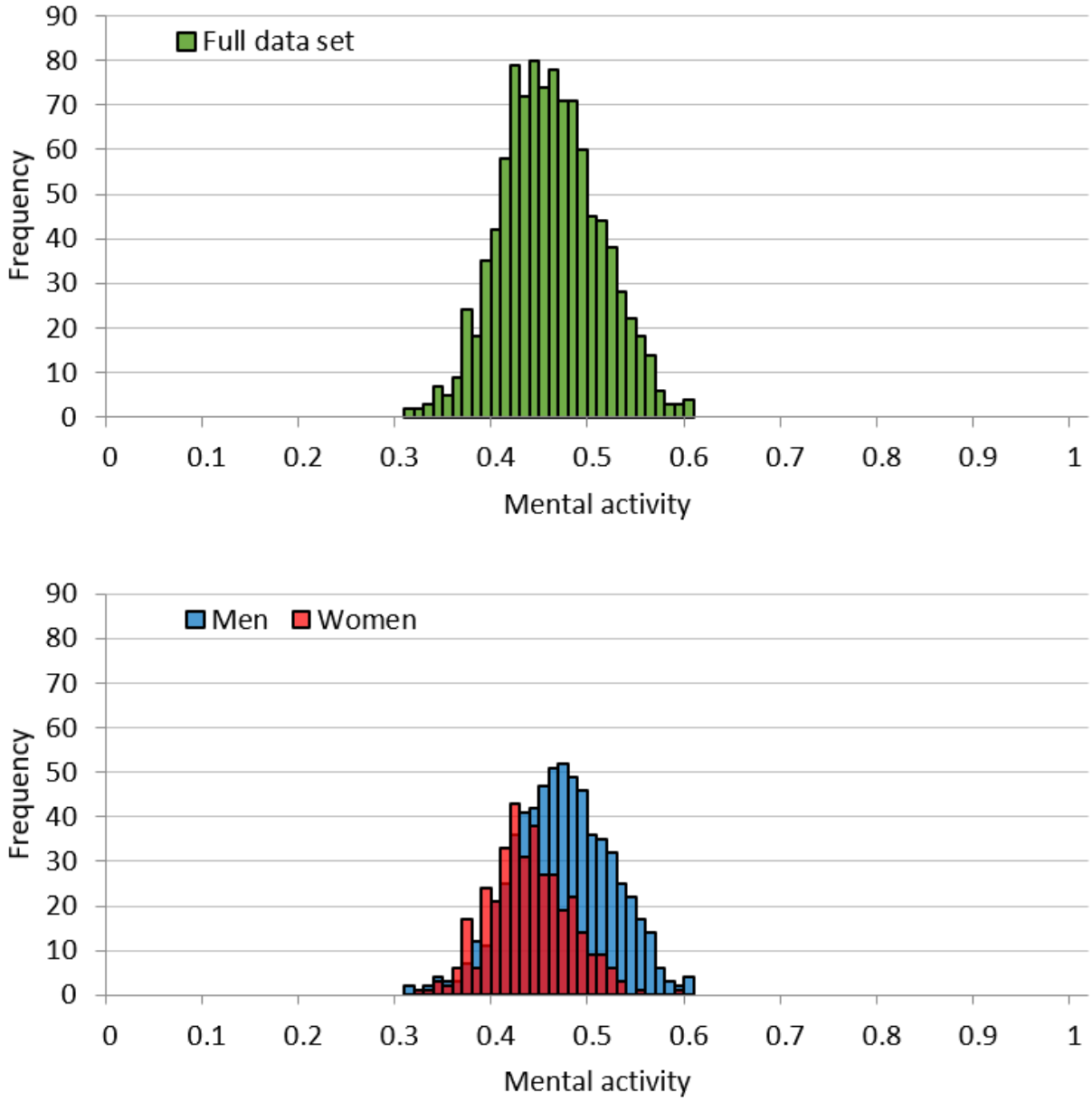


Figure 5 shows the mental activity distribution of the user dataset. The mean for the full dataset was 0.45 (SD 0.051), while the means for men and women were 0.46 (SD 0.051) and 0.43 (SD 0.041), respectively. Similar to the call dataset, the distribution was lower for women than for men. In addition, like vitality, the effect size of the goodness-of-fit test for a normal distribution for the full set was 0.09 (P=.98). Therefore, we concluded that mental activity follows a normal distribution. Regarding the distribution for men and women separately, the goodness-of-fit effect size was 0.096 (P>.99) and 0.084 (P>.99), respectively. Thus, like the distribution of the full set, it was concluded that mental activity follows a normal distribution for both men and women as well.

Figure 5. User dataset mental activity distribution for the full set and by sex.



Sex Differences in Vitality, Psychological Tests, and Vitality vs Mental Activity Comparison

Figure 6 shows the frequency distribution of BDI scores by sex. There were more men than women with a BDI score of less than 11. In addition, the men-to-women ratio was roughly equivalent. Figure 7 shows the distribution of BDI test completion time (seconds). Approximately 10% of the participants completed the BDI test in 1.5 minutes or less. Another 10% or so of the participants took 5 minutes or more to complete the BDI test. The shortest time was 9 seconds and the longest time was about 22 hours.

Figure 6. Beck Depression Inventory (BDI) score distribution by sex.

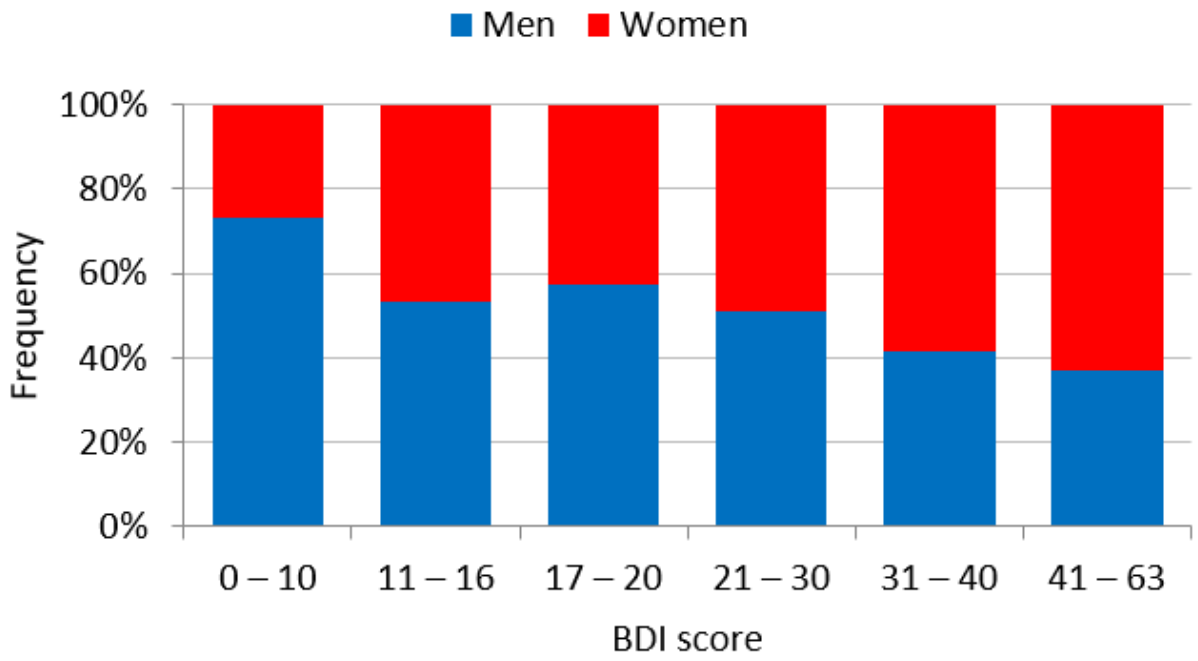


Figure 7. Beck Depression Inventory (BDI) test completion time distribution.

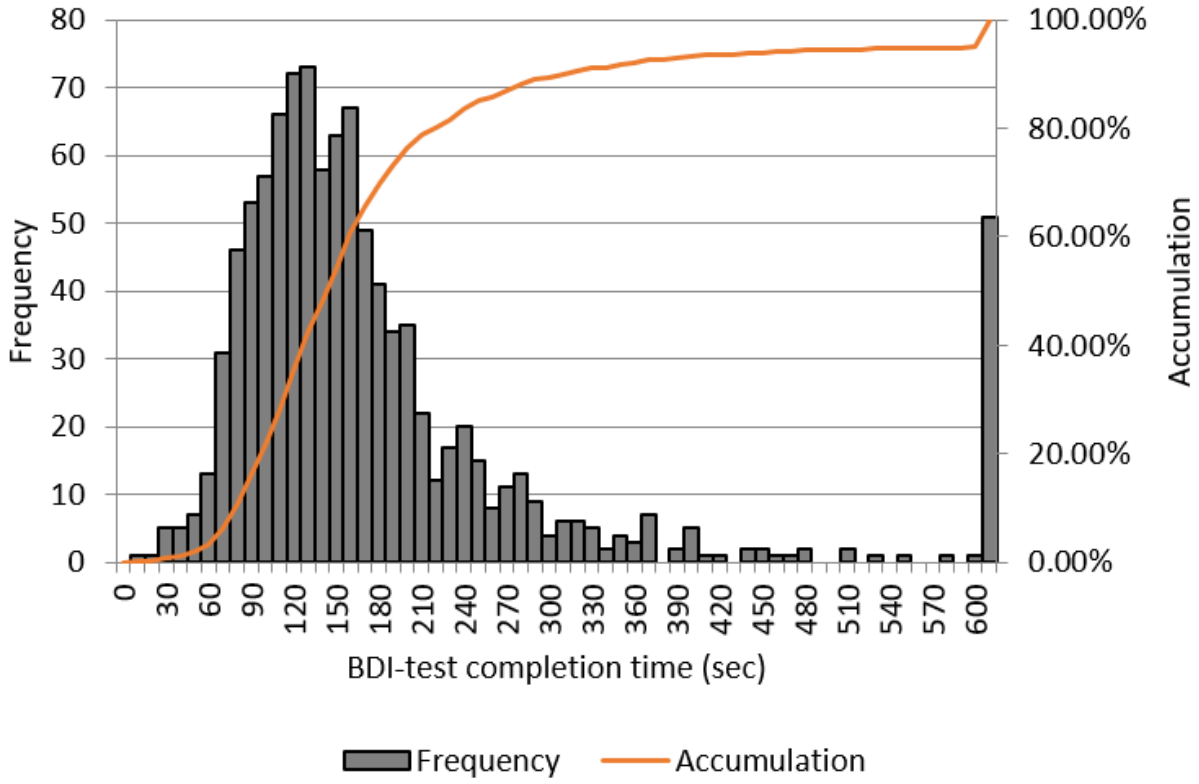


Figure 8 shows the vitality distribution by sex according to the evaluation criteria of Beck et al [24]. The *P* values shown in Table 2 represent the comparison of the mean vitality for two groups of men and women with a *t* test, and the effect size represents the size of the difference between the two groups for the *t* test [26]. Significant differences were observed in tests of the vitality for men and women in both the normal BDI score category and in the abnormal BDI score category. However, because the effect size of the difference between men and women in the normal category was considered to be small according to prior standards [26], there was no substantial difference between men and women in this category.

Conversely, the effect size of the sex difference in the abnormal category was medium; thus, it was considered to indicate a substantial difference.

Figure 8. Vitality distribution by sex with Beck Depression Inventory (BDI) score evaluation criteria.

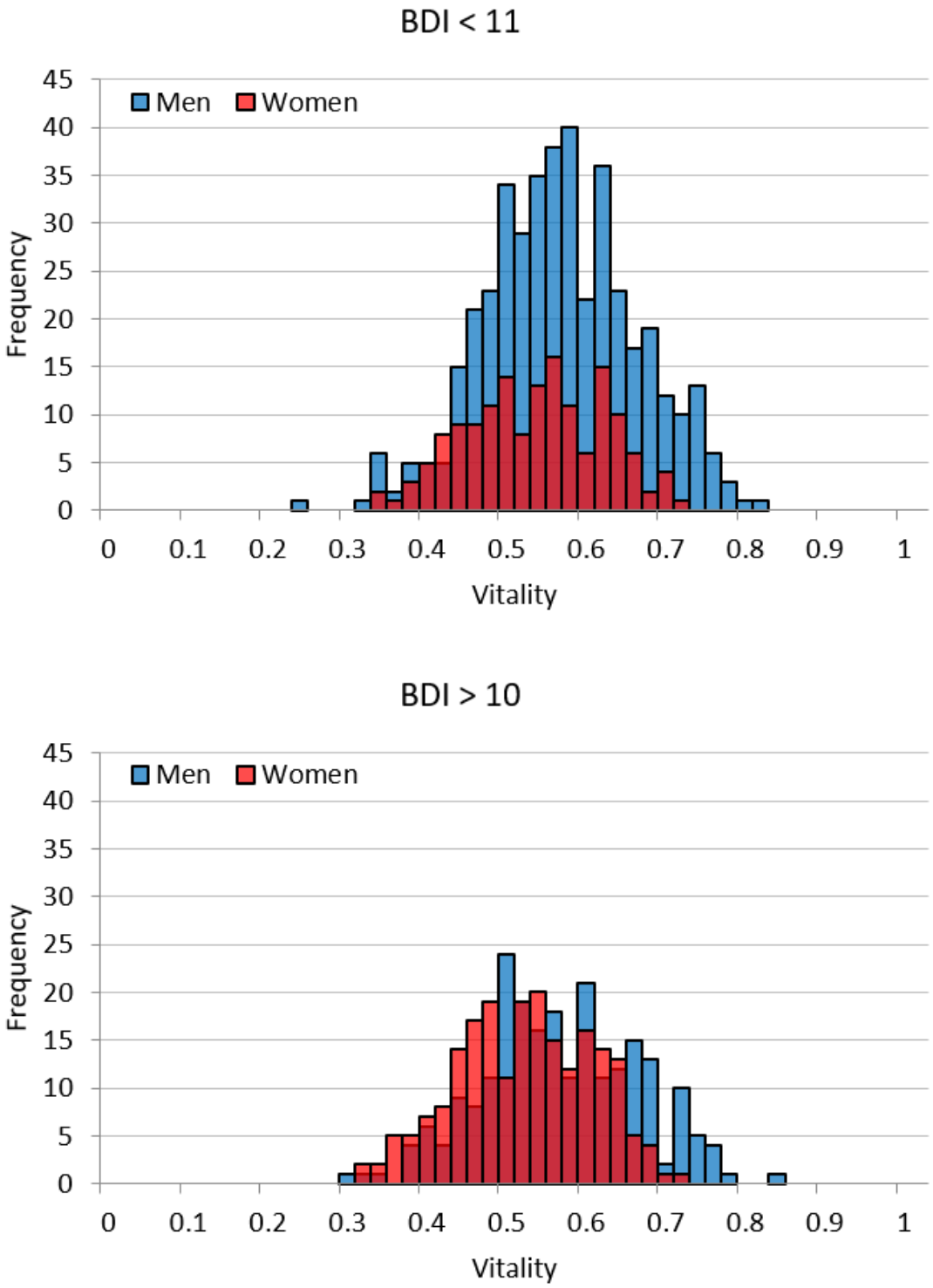


Table 2. Comparison of mean vitality between two groups of men and women according to BDI[a] score evaluation criteria.

Category	Sex (men), n (%)	Men, mean (SD)	Women, mean (SD)	*P* value[b]	Effect size
BDI<11 (N=577)	423 (73.31)	0.56 (0.095)	0.53 (0.085)	<.001	0.34
BDI>10 (N=438)	228 (52.05)	0.56 (0.097)	0.51 (0.084)	<.001	0.48

[a]BDI: Beck Depression Inventory.

[b]Based on *t* tests.

Figure 9 and Figure 10 show the correlations of BDI score with vitality and mental activity, respectively. Graphs A1 in both figures show the data for all users in the user dataset, graphs A2 show data for users in A1 who declared depression as a history of present illness, graphs A3 show data for users with a BDI test completion time of 80-300 seconds (all users except the top and bottom 10%), and graphs A4 show data for users in A3 who declared depression as part of the history of present illness. Moreover, graphs B show data for users from the user dataset with a completion time of 80-300 seconds in the second BDI test that was performed 3 months later. Graphs C show data for users with a completion time of 80-300 seconds in the third BDI test that was performed another 3 months later. The vitality measurements in graphs B and C were the first to be performed after the BDI test in question, whereas the mental activity measurements were the most recent values within 2 weeks of the BDI test in question. Regarding vitality, we found a significant yet weak negative correlation between the initial BDI score and vitality for users with a completion time of 80-300 seconds in the first BDI test and who declared a history depression. We did not find any sufficiently significant correlation between mental activity and the BDI score.

Figure 9. Correlation between Beck Depression Inventory (BDI) score and vitality for users.

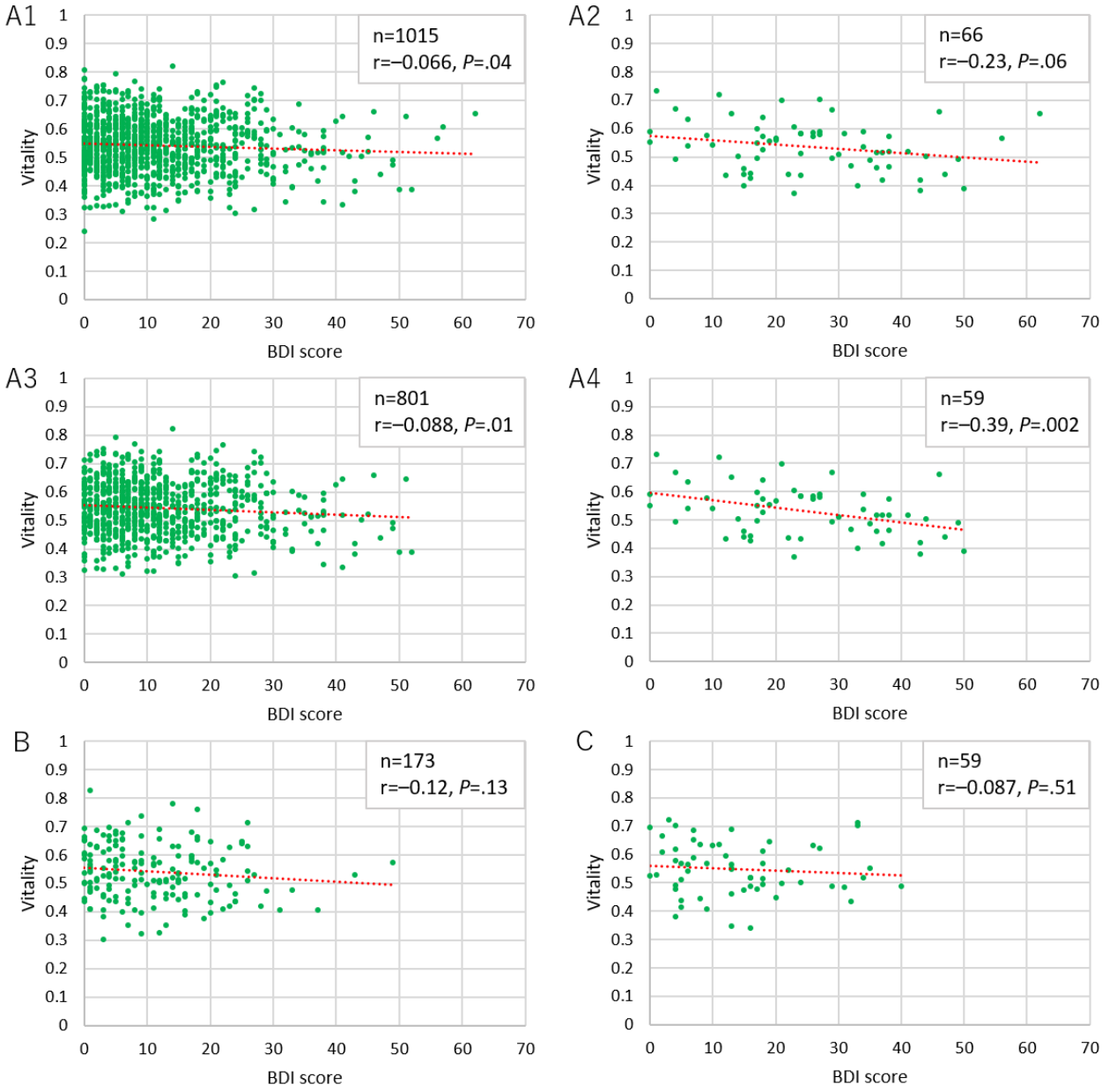
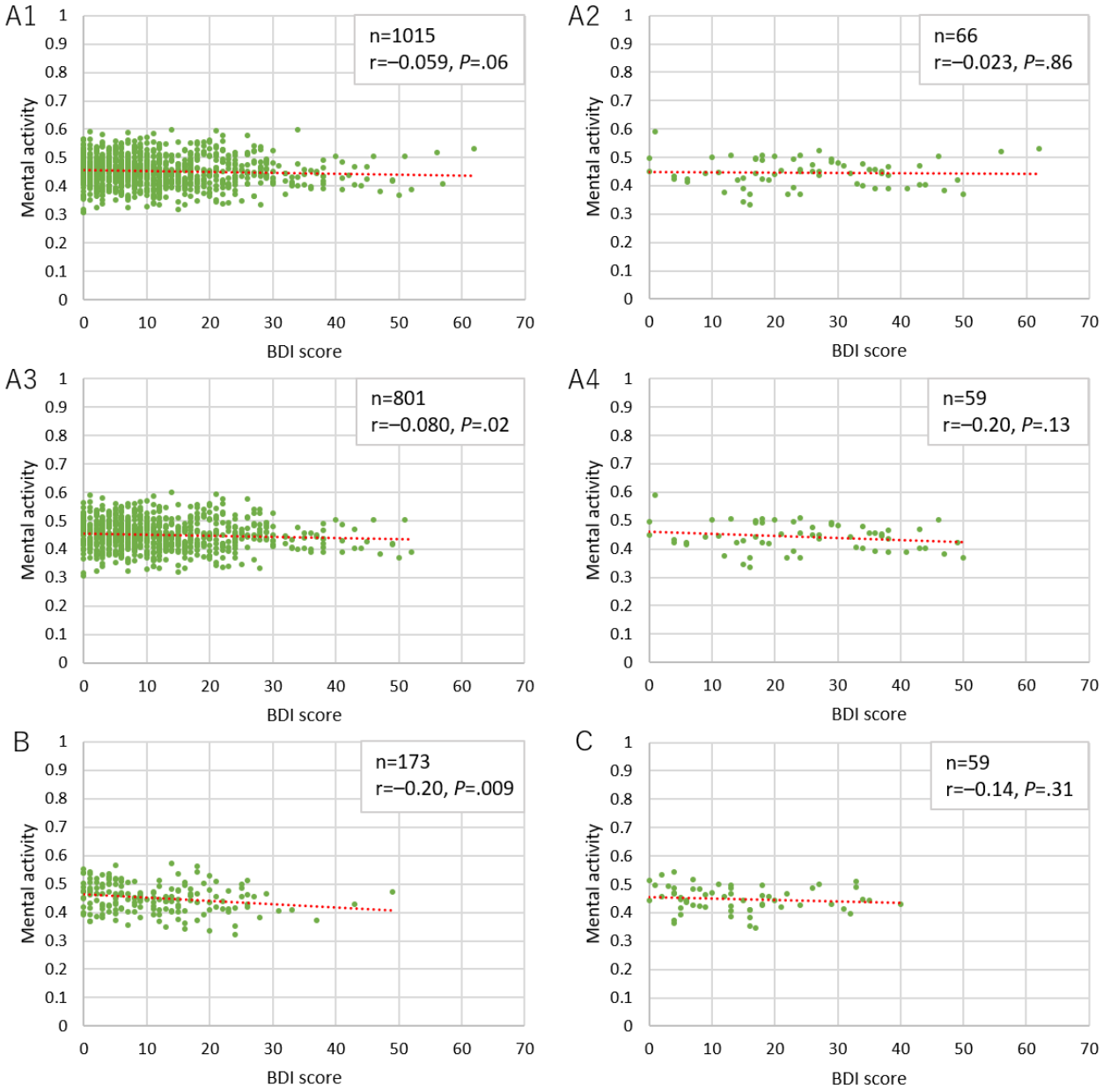


Figure 10. Correlation between Beck Depression Inventory (BDI) score and mental activity for users.



Temporal Changes in Vitality and Mental Activity Before and After a Stressful Life Event

Figure 11 shows the changes in vitality over a 2-week period before and after the earthquakes by region. Figure 12 shows the changes in the most recent mental activity before the earthquakes and within the 2 weeks after the earthquakes. After the earthquakes, vitality decreased drastically in Shikoku, Kyushu, and Kumamoto, but a slight downward trend in mental activity was observed in Kumamoto only.

Figure 11. Changes in vitality before and after the Kumamoto earthquakes.

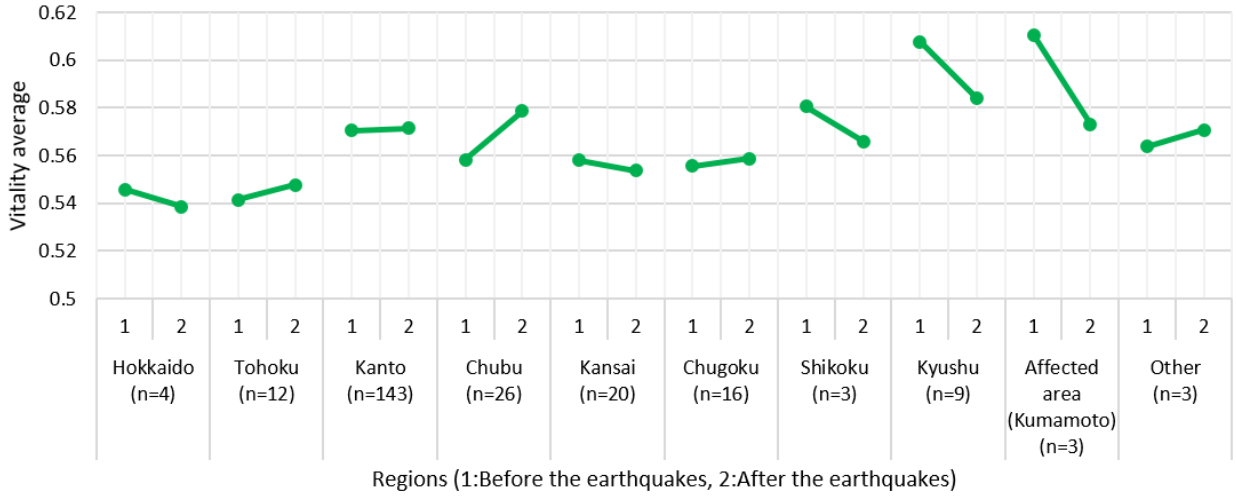
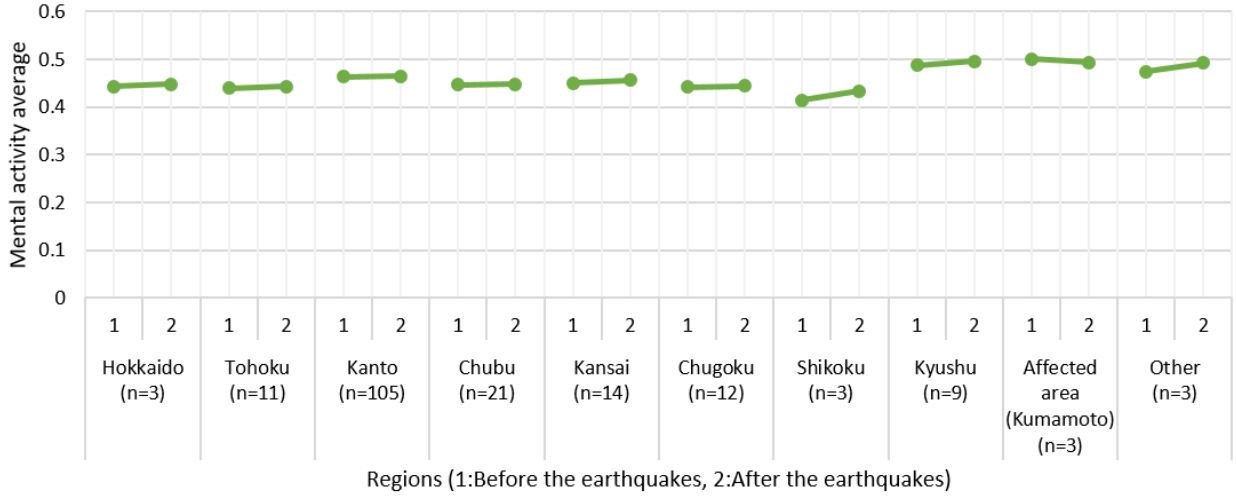


Figure 12. Changes in mental activity before and after the Kumamoto earthquakes.



Additionally, Figure 13 shows the change over time in mental activity for one disaster victim living in Kumamoto. The vertical line in the graph represents the day and time of the foreshock. An obvious drop in mental activity can be seen before and after the earthquake. The mental activity did not return to the level before the earthquake by the end of the data collection period.

Figure 13. Change over time in mental activity for one victim of the Kumamoto earthquakes.

Discussion

Principal Findings

The rate of continued use for MIMOSYS rivaled that of the top 60 apps in Google Play's rankings. It is believed that the primary reason contributing to long-term maintenance of a high rate of continued use is because MIMOSYS can be used effortlessly, as the app begins processing automatically when a call is made. MIMOSYS is effective at maintaining long-term use, which is a prerequisite for a monitoring system.

The distribution of vitality is considered to approximate a normal distribution because it is the average vitality for each individual utterance in the voice data. When it comes to effect size in the call dataset, it is acceptable to say that vitality follows a normal distribution. Regarding the user dataset, from the goodness-of-fit testing, we concluded that vitality follows a normal distribution. Most parametric statistical analyses assume that data for an analysis will follow a normal distribution; thus, whether the obtained data actually follow a normal distribution is vital for the statistical tests performed thereafter. From this perspective, vitality is advantageous for statistical analyses. The same can be said for mental activity. However, some distortion was detected in the distribution of mental activity. This is likely influenced by the fact that mental activity is an indicator that is largely controlled by life events unique to the user.

We have continuously observed low vitality values for people with depression in our research, whereas the vitality of healthy individuals is more widely scattered [27]. According to Kessler et al [28], approximately 20% of Japanese people in 2007 would have depression or another mental illness at some point in their lives. Since the normality of the vitality distribution showed a mean range of 0.46-0.65 for about 70% of the total, we determined that vitality values scattered within this range indicate good mental health, which is in line with the report of Kessler et al [28].

Women tended to have lower vitality and mental activity than men. There are two probable reasons for this. The first is a problem with the algorithm. Women are generally more expressive of their emotions than men [29]. The differing trends in emotional expression between men and women were accounted for in the development of MIMOSYS through the proposal of an algorithm designed to eliminate sex differences [27]. However, the difference may have surfaced with the increase in the size of the dataset analyzed because the number of data entries from teachers used in the development of the algorithm was insufficient.

The second reason is the underlying sex difference in the rate of depression, which is higher in women than men [30]. In the distribution of BDI scores, there were overwhelmingly fewer women than men in the normal category, and the more severe category also had slightly more women than men. It is possible that MIMOSYS is detecting this difference. To investigate this, data were classified based on the evaluation criteria for BDI score and sex differences in vitality explored for each category. In the normal category, from the viewpoint of effect size, there was no substantial sex difference considered to exist. Conversely, in the abnormal category, from the viewpoint of effect size, there was a substantial sex difference determined. In other words, the MIMOSYS algorithm was functioning correctly in the normal group, whereas the abnormal group may be reflecting differences in the rate of depression. Nonetheless, the difference in effect size between the normal and abnormal groups was not overly large, and this point requires additional investigation.

We could not find any sufficient correlation between BDI score and vitality/mental activity in this study. A possible reason for this is that it typically takes 2-3 minutes to complete the BDI test if conducted normally; however, we observed a wide spread of completion times that might not have yielded reliable scores. As such, we conducted the analysis after excluding the top and bottom 10% of completion times as they were considered to have low reliability; however, this did not lead to any considerable changes. The reason for this is unknown. We observed the same tendencies for the BDI scores in the second and third tests performed at 3-month intervals. We believe that

users with depression responded to the test relatively seriously and found a significant but weak correlation between vitality and initial BDI score; yet, these data cannot be considered sufficient. As discussed below, the data from the Kumamoto earthquakes clearly traced changes in mental state. This suggests the possibility that BDI testing through smartphones, the gap between reality and self-reported conditions because of reporting bias, and the scattering of voice evaluation mutually influenced the results. As such, it is possible that this system is more suitable for monitoring rather than for screening if we take the BDI test as the standard. Furthermore, the reason we did not refer to only users with depression in the second and third tests is that we could not find a sufficient number of users, which undermined the reliability of the statistical analysis.

The reason we could not find any correlation between BDI score and mental activity for users with depression is likely the same as that for vitality. Moreover, since we did not have any vitality data collected prior to BDI score measurement, it is possible that the actual tendency for the correlation with the initial BDI score was not reflected in the mental activity. In fact, we did find an extremely weak yet significant correlation with the second BDI score. Nevertheless, any further investigation would be limited since we suspect a major influence from the reliability of the user data.

We previously reported a correlation between BDI score and vitality/mental activity [31,32]. We also reported a correlation between vitality and the Hamilton Depression Rating Scale [33] for individuals with depression [34]. These results were obtained from participants who were recruited appropriately, which supports that this method exhibits adequate performance under specified conditions. Although it may also be possible to obtain a favorable result from these data by extracting information from users whose day-to-day background is clear, it is difficult to know a user's detailed state from their attributes; thus, reexamining the study design remains a future research task. However, the design used in this study has shown us that it is difficult to obtain our intended result.

As the BDI test is intended for clinical applications, it would have been more appropriate to use Patient Health Questionnaire 9 [35] and Center for Epidemiologic Studies Depression Scale [36] for this study. However, it was not possible to additionally administer these tests to the participants; therefore, they should be utilized in future studies.

As the primary purpose of MIMOSYS is mental health monitoring, analyzing changes in vitality and mental activity in relation to the user's life events is desired. We observed a deterioration of mental health among users in the areas affected by the Kumamoto earthquakes around the time of the disaster. This was thought to reflect the heightened stress caused by the earthquake. It is easiest to understand this finding by considering the results of one disaster victim presented as an example. This user's mental activity was good for some time after starting to use MIMOSYS; however, there was a steep drop immediately after a certain point, following which mental activity did not recover by the time the user stopped using MIMOSYS. The Kumamoto earthquakes occurred around the time of the marked change in mental activity, and it is hypothesized that the cause of the drop in mental activity was stress from the earthquakes. By observing changes in vitality and mental activity along with life events over the mid to long term in this manner, we can not only detect poor mental health in the early stage but can also hypothesize the cause of that state, leading to a solution at the root of the problem.

From this perspective, although MIMOSYS is not well-suited for taking measurements upon every screening, the results suggested that it could be an effective mental health monitoring system. A separate study in which we compared changes in vitality and mental activity in the affected area and other regions at the time of the Kumamoto earthquakes [37] showed that vitality and mental activity increased in a region in which a similar earthquake had occurred in the past (Tohoku) because of excitement from flashbacks to memories of that time (trends similar to those can be observed in Figures 11 and 12 as well). This result may reflect detecting symptoms of posttraumatic stress disorder such as hypervigilance or irritability, and indicates that MIMOSYS may be useful in detecting posttraumatic stress disorder. Nevertheless, we did not conduct any examination using the Impact of Event Scale-Revised [38] or the Post-traumatic Diagnostic Scale [39]; thus, such comparisons are a task for future studies.

This discussion assumes the legitimacy of users' attribute data; however, because these data were reported by users for the current study, it is difficult to verify the reliability. Therefore, it is possible that accuracy was lost when categorizing the data. Moreover, since the voice analysis was conducted using smartphones privately owned by the users, we could not exclude dependence on the model. This is a limitation of the current study and a challenge for future research. Data collection via smartphone is simple and allows for the accumulation of large-scale data; however, we encountered the problem of necessarily rejecting the null hypothesis in statistical hypothesis testing because of a too-large sample size. In the future, it will be necessary to include not only analyses based on frequentism but also Bayesian statistics.

Conclusions

In this prospective study, a system for monitoring mental health based on voice data from smartphone calls (MIMOSYS) was opened to the public, data were collected over a period of 2 years from a large unspecified sample of users, and the characteristics and potential of MIMOSYS were examined from various perspectives to verify the effectiveness of the system.

MIMOSYS was found to be effective concerning the rate of continued use. The greatest feature of MIMOSYS is mental health monitoring, and this is done most effectively by performing long-term time-series analysis of mental activity considering users' life events. Fluctuation in mental activity because of stress from life events was observed. It is likely that patient interventions will be facilitated by evaluating objective data such as those collected by MIMOSYS together with life events. However, it is difficult to collect information on users' personal life events with this system alone. Hereafter, we wish to investigate a method of data collection that accounts for this and leads to further verification of MIMOSYS.

Acknowledgments

This research was partially supported by the Center of Innovation Program from Japan Science and Technology Agency. We thank Editage (www.editage.com) for English language editing.

Conflicts of Interest

MH, MN, and ST had received financial support from PST Inc until 2019 and currently report no financial support from the company. All other authors declare no conflicts of interest.

References

1. Cohen S, Kessler RC, Gordon LU. Measuring stress: A guide for health and social scientists. Oxford: Oxford University Press; 1997.
2. Perkins A. Saving money by reducing stress. Harvard Business Review 1994;72(6):12.
3. Okumura Y, Higuchi T. Cost of depression among adults in Japan. Prim Care Companion CNS Disord 2011;13(3):e1-e9 [FREE Full text] [doi: 10.4088/PCC.10m01082] [Medline: 21977377]
4. Okuda A, Suzuki T, Kishi T, Yamanouchi Y, Umeda K, Haitoh H, et al. Duration of untreated illness and antidepressant fluvoxamine response in major depressive disorder. Psychiatry Clin Neurosci 2010 Jun;64(3):268-273. [doi: 10.1111/j.1440-1819.2010.02091.x] [Medline: 20602727]
5. Izawa S, Sugaya N, Shirotsuki K, Yamada KC, Ogawa N, Ouchi Y, et al. Salivary dehydroepiandrosterone secretion in response to acute psychosocial stress and its correlations with biological and psychological changes. Biol Psychol 2008 Dec;79(3):294-298. [doi: 10.1016/j.biopsycho.2008.07.003] [Medline: 18706968]
6. Suzuki G, Tokuno S, Nibuya M, Ishida T, Yamamoto T, Mukai Y, et al. Decreased plasma brain-derived neurotrophic factor and vascular endothelial growth factor concentrations during military training. PLoS One 2014 Feb;9(2):e89455 [FREE Full text] [doi: 10.1371/journal.pone.0089455] [Medline: 24586790]
7. Goldberg DP. Manual of the General Health Questionnaire. Windsor: NFER Publishing; 1978.
8. Goldberg DP, Hillier VF. A scaled version of the General Health Questionnaire. Psychol Med 1979 Feb;9(1):139-145. [doi: 10.1017/s0033291700021644] [Medline: 424481]
9. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry 1961 Jun;4:561-571. [Medline: 13688369]
10. Delgado-Rodríguez M, Llorca J. Bias. J Epidemiol Community Health 2004 Aug;58(8):635-641 [FREE Full text] [doi: 10.1136/jech.2003.008466] [Medline: 15252064]
11. Jan A, Meng H, Gaus YFA, Zhang F, Turabzadeh S. Automatic depression scale prediction using facial expression dynamics and regression. 2014 Presented at: The 4th International Workshop on Audio/Visual Emotion Challenge; November 7, 2014; Orlando, FL p. 73-80. [doi: 10.1145/2661806.2661812]
12. Tokuno S, Mitsuyoshi S, Suzuki G, Tsumatori G. Stress evaluation by voice: A novel stress evaluation technology. 2014 Presented at: The 9th International Conference on Early Psychosis; November 17-19, 2014; Tokyo, Japan p. 17-19.
13. Jiang H, Hu B, Liu Z, Yan L, Wang T, Liu F, et al. Investigation of different speech types and emotions for detecting depression using different classifiers. Speech Communication 2017 Jun;90:39-46. [doi: 10.1016/j.specom.2017.04.001]
14. Mundt JC, Vogel AP, Feltner DE, Lenderking WR. Vocal acoustic biomarkers of depression severity and treatment response. Biol Psychiatry 2012 Oct 1;72(7):580-587 [FREE Full text] [doi: 10.1016/j.biopsych.2012.03.015] [Medline: 22541039]
15. Donker T, Petrie K, Proudfoot J, Clarke J, Birch MR, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013 Nov 15;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
16. Lu H, Rabbi M, Chittaranjan GT, Frauendorfer D, Mast MS, Campbell AT, et al. Stresssense: Detecting stress in unconstrained acoustic environments using smartphones. 2012 Presented at: The 2012 ACM Conference on Ubiquitous Computing; September 5-8, 2012; Pittsburgh, PA p. 351-360. [doi: 10.1145/2370216.2370270]
17. Shimizu E, Nakamura T, Kim J, Yoshiuchi K, Yamamoto Y. Application of Empirical Mode Decomposition to Mother and Infant Physical Activity. Methods Inf Med 2018 May;57(3):152-157. [doi: 10.3414/ME18-02-0001] [Medline: 29719923]
18. Omiya Y, Hagiwara N, Shinohara S, Nakamura M, Mitsuyoshi S, Tokuno S. Development of Mind Monitoring System using call voice. 2016 Presented at: Neuroscience; November 12-16, 2016; San Diego, CA.
19. Mitsuyoshi S, Tanaka Y, Ren F, Shibasaki K, Kato M, Murata T, et al. Emotion voice analysis system connected to the human brain. 2007 Presented at: The 12th International Conference on Natural Language Processing and Knowledge Engineering; December 7-10, 2007; Beijing, China p. 479-484. [doi: 10.1109/nlpke.2007.4368074]
20. Diagnostic and statistical manual of mental disorders 4th edition, Text Revision. Washington, DC: American Psychiatric Association; 2000.

21. Omiya Y, Hagiwara N, Nakamura M, Shinohara S, Higuchi M, Takano T, et al. Verification of the effects of age and gender on mental health status measurement using voice [Translated from Japanese]. 2018 Presented at: The 91st Japan Society for Occupational Health; May 16-19, 2018; Kumamoto, Japan.
22. R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL: https://www.R-project.org/ [accessed 2019-02-03]
23. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/BF03193146]
24. Beck AT, Steer RA, Brown GK. Manual for the Beck depression inventory. 2nd ed. San Antonio, TX: Psychological Corporation; 1996.
25. Chen A. New data shows losing 80% of mobile users is normal, and why the best apps do better. URL: https://andrewchen.co/new-data-shows-why-losing-80-of-your-mobile-users-is-normal-and-that-the-best-apps-do-much-better/ [accessed 2019-02-03]
26. Cohen J. Statistical power analysis for the behavioral sciences. London: Routledge; 1988.
27. Tokuno S. Pathophysiological voice analysis for diagnosis and monitoring of depression. In: Kim YK, editor. Understanding depression. Singapore: Springer; 2018:83-95.
28. Kessler RC, Angermeyer M, Anthony JC, De Graaf R, Demyttenaere K, Gasquet I, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007 Oct;6(3):168-176 [FREE Full text] [Medline: 18188442]
29. LaFrance M, Banaji M. Toward a reconsideration of the gender-emotion relationship. In: Clark MS, editor. Emotion And Social Behavior (The Review Of Personality And Social Psychology) (v. 14). Newbury Park, CA: Sage Publications; 1992:178-201.
30. Weissman MM, Klerman GL. Sex differences and the epidemiology of depression. Arch Gen Psychiatry 1977 Jan;34(1):98-111. [doi: 10.1001/archpsyc.1977.01770130100011] [Medline: 319772]
31. Hagiwara N, Omiya Y, Shinohara S, Nakamura M, Yasunaga H, Higuchi M, et al. Validity of the Mind Monitoring System as a mental health indicator. 2016 Presented at: The 16th International Conference on BioInformatics and BioEngineering; October 31-November 2, 2016; Taichung, Taiwan p. 262-265. [doi: 10.1109/bibe.2016.62]
32. Hagiwara N, Omiya Y, Shinohara S, Nakamura M, Higuchi M, Mitsuyoshi S, et al. Validity of Mind Monitoring System as a Mental Health Indicator using Voice. Adv Sci Technol Eng Syst J 2017 May;2(3):338-344. [doi: 10.25046/aj020343]
33. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960 Feb;23:56-62 [FREE Full text] [doi: 10.1136/jnnp.23.1.56] [Medline: 14399272]
34. Higuchi M, Nakamura M, Shinohara S, Omiya Y, Hagiwara N, Takano T, et al. Effectiveness of voice-based mental health evaluation indicator in estimating depression. 2018 Presented at: Human Communication Group Symposium; December 12-14, 2018; Mie, Japan.
35. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
36. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. Appl Psychol Meas 1977 Jun;1(3):385-401. [doi: 10.1177/014662167700100306]
37. Tokuno S, Omiya Y, Shinohara S, Nakamura M, Hagiwara N, Mitsuyoshi S. Psychological impact of Kumamoto earthquake by voice analysis using a smart phone application. 2016 Presented at: Neuroscience; November 12-16, 2016; San Diego, CA.
38. Weiss DS, Marmar CR. The impact of event scale–revised. In: Wilson JP, Keane TM, editors. Assessing psychological trauma and PTSD. New York: Guilford Press; 1997:399-411.
39. Foa EB, Cashman L, Jaycox L, Perry K. The validation of a self-report measure of posttraumatic stress disorder: The Posttraumatic Diagnostic Scale. Psychol Assess 1997;9(4):445-451. [doi: 10.1037/1040-3590.9.4.445]

Abbreviations

BDI: Beck Depression Inventory
MIMOSYS: Mind Monitoring System

Edited by J Torous, G Eysenbach; submitted 04.10.19; peer-reviewed by P Batterham, L Bonet; comments to author 27.10.19; revised version received 26.02.20; accepted 15.06.20; published 20.07.20.

Please cite as:
Higuchi M, Nakamura M, Shinohara S, Omiya Y, Takano T, Mitsuyoshi S, Tokuno S
Effectiveness of a Voice-Based Mental Health Evaluation System for Mobile Devices: Prospective Study
JMIR Form Res 2020;4(7):e16455
URL: http://formative.jmir.org/2020/7/e16455/
doi:10.2196/16455
PMID:32554367

Original Paper

Sedentary Work in Desk-Dominated Environments: A Data-Driven Intervention Using Intervention Mapping

Nathalie M Berninger[1], MSc; Gill A ten Hoor[1], PhD; Guy Plasqui[2], PhD; Gerjo Kok[1], PhD; Gjalt-Jorn Ygram Peters[1,3], PhD; Robert A C Ruiter[1], PhD

[1]Department of Work and Social Psychology, Maastricht University, Maastricht, Netherlands

[2]Department of Nutrition and Movement Sciences, Maastricht University Medical Centre, Maastricht, Netherlands

[3]Faculty of Psychology and Education Science, Open University of The Netherlands, Heerlen, Netherlands

Corresponding Author:
Nathalie M Berninger, MSc
Department of Work and Social Psychology
Maastricht University
PO Box 616, WSP
Universiteitssingel 40
Maastricht
Netherlands
Phone: 31 433881908
Email: nathalie.berninger@maastrichtuniversity.nl

Abstract

Background: Since desk-dominated work environments facilitate sedentary behavior, office workers sit for 66% of their working days and only 8% succeed in interrupting their prolonged periods of sitting within the first 55 minutes. Yet stretches of long and uninterrupted sitting increase the likelihood of several chronic metabolic and cardiovascular diseases.

Objective: We therefore developed a computer-based app designed to interrupt periods of prolonged sitting among office employees.

Methods: When developing the intervention, we applied the intervention mapping protocol. This approach for the systematic design of theory and evidence-based behavior change programs consists of 6 steps: creation of a logic model of the problem, creation of a logic model of change, program design, program production, design of an implementation plan, and development of an evaluation plan.

Results: Working through all 6 steps has resulted in an individually adaptable intervention to reduce sedentary behavior at work. The intervention, UPcomplish, consists of tailored, half-automatized motivational components delivered by a coach. To register sedentary behavior, the VitaBit (VitaBit Software International BV) toolkit, a wearable accelerometry-based monitoring device, is used. Among others, UPcomplish includes personalized goal setting, tailored suggestions to overcome hurdles, and weekly challenges. The VitaBit toolkit supports the participants to monitor their behavior in relation to self-set goals.

Conclusions: Intervention mapping is a useful protocol not only for the systematic development of a comprehensive intervention to reduce sedentary behavior but also for planning program adherence, program implementation, and program maintenance. It facilitates obtaining the participation of relevant stakeholders at different ecological levels in the development process of the intervention and anticipating facilitators to and barriers of program implementation and maintenance.

Trial Registration: Netherlands Trial Register NL7503; https://www.trialregister.nl/trial/7503

(JMIR Form Res 2020;4(7):e14951) doi:10.2196/14951

KEYWORDS

intervention mapping; sedentary behavior; sedentary work; computer-based; occupational health; eHealth; mHealth; data-driven programs

Introduction

Background

Frequent and uninterrupted sedentary behavior is highly prevalent among office workers [1,2] and negatively impacts workers' health and well-being by increasing the risk of noncommunicable diseases such as cardiovascular disease, type 2 diabetes [3-5], obesity [6], and mental health problems [7,8]. This is reflected in the higher mortality rates among office workers as compared with those in more active occupations [9]. Sedentary behavior is defined as sitting, lying, or reclining awake behaviors with low-energy expenditures (≤1.5 metabolic equivalents) [10]. Compensating for the negative effects of sitting time by meeting the recommended levels of physical activity may not be possible [11-15]. Moreover, the accumulation of long uninterrupted sitting bouts and/or a daily sitting time of more than 10 hours has been defined as an unhealthy sitting pattern resulting in increased metabolic risk [15,16]. Research suggests that prolonged sitting should be interrupted by bouts of light to moderate physical activity [16,17] and standing [18,19].

Few studies described the long-term positive effects of interventions to reduce sedentary behavior. Interventions mostly incorporated multiple behavior change methods targeting multiple behavioral determinants [20,21]. Behavior change methods are defined as "general techniques or processes that have been shown to be able to change one or more determinants of behavior" and the behavior, if parameters for use are respected [22,23]. For instance, behavior change methods providing information about health consequences and self-monitoring help build the attitude required to decide to change; instructions about how to perform the behavior and social support help build the self-efficacy required to translate the intention into behavior. Establishing a clear link between the identified determinants of behavior and behavior change methods targeting these determinants is a key component of effective behavior change, according to the intervention mapping (IM) protocol [22]. Worksite physical activity interventions designed using IM have revealed positive long-term effects [24-26]. However, current effective sedentary behavior interventions are quite cost-intensive requiring a personal coach and/or environmental changes [26-28]. This paper describes the systematic development of a low-cost data-driven worksite sedentary behavior intervention designed with the IM protocol.

Intervention Mapping

IM is a framework for planning intervention development, implementation, and evaluation with six iterative steps. In each step, the program designer applies findings from theory, evidence, and their own research: (1) conducting a needs assessment, (2) stating program outcomes and objectives, (3) designing the program, (4) preparing program production, (5) planning program implementation, and (6) developing an evaluation plan (see Figure 1) [22,29].

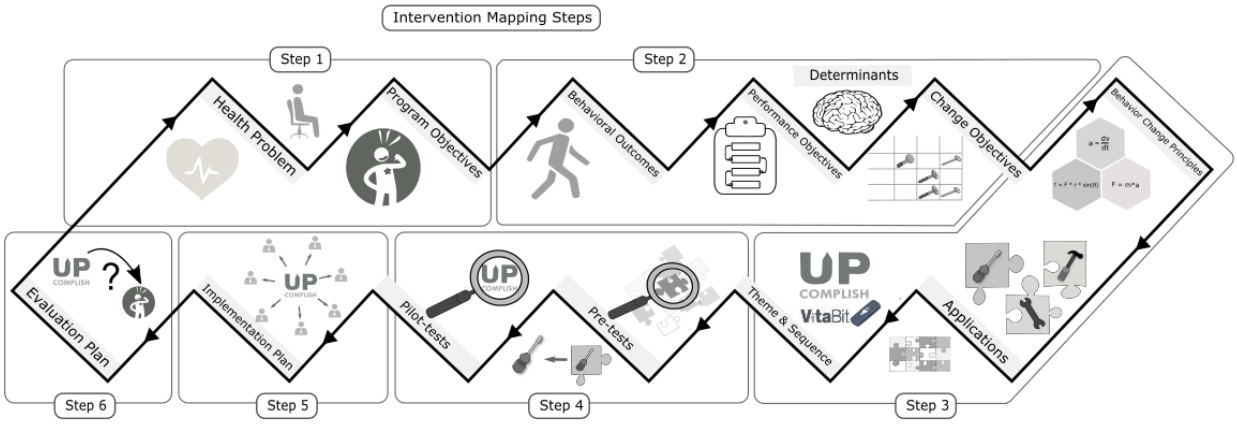
Sedentary behavior can be embedded at both the interpersonal (ie, support by colleagues and managers) and the individual (ie, office workers) level. For example, if an employee would like to interrupt sitting time more often during working hours but is devaluated by their colleagues for not working enough, the new behavior might disappear. Higher levels (ie, organization, community, and society) were not considered in this study for reasons of cost-effectiveness and given that the target high-income Western countries provide sufficient opportunities (such as safe pathways) for individuals to sit less during working hours.

An intervention planning group includes stakeholders who can make relevant contributions to the development, implementation, and evaluation, such as members of the target group and future implementers. This ensures that issues pertinent to the target group are addressed by the intervention or that future implementation issues are anticipated ahead of time [22].

Computer and smartphone technologies can create platforms that support interactions between individuals, making it possible to exchange both print and more complex multimedia files (eg, a coaching procedure at reduced costs that allows for individually adapted suggestions) [30-32]. Since a permanent reduction of sedentary behavior requires the personal assistance of a professional [33], the main component of our intervention is UPcomplish, which is partly automated, with tailored feedback and motivational support remotely provided by a coach. The VitaBit monitoring toolkit is part of the intervention; participants can monitor their own sedentary behavior related to their personal goals, and the UPcomplish coach can use those data to give almost real-time tailored advice.

In this paper, we describe the systematic development of UPcomplish and the design of the VitaBit monitoring toolkit. IM guided important decisions with regard to objectives, behavior change methods, program production, implementation, and evaluation. The decisions were informed by relevant theoretical and empirical literature including our own empirical research. With UPcomplish and VitaBit, we aim to reduce the number and length of sitting bouts among office workers in the short term [3] and increase the vitality and mental health of employees, as well as minimize their risks for noncommunicable diseases in the longer term.

Figure 1. Overview of the steps and products in the intervention mapping protocol.



Methods

All materials and supporting documents are available at the Open Science Framework (OSF) repository [34]. The target population consists of office workers in high-income countries [35]. The trial was registered with the Netherlands Trial Register [NL7503].

Intervention Mapping Steps 1 and 2: Needs Assessment and Program Objectives

The first two IM steps cover problem identification and the logic model of change (problem behaviors and desired behaviors, as well as environmental outcomes). The health problem of sedentary behavior, its impact on quality of life, and the context of the intervention were specified (Figure 1). Individual and environmental factors causing sedentary behavior were identified, and behavioral and psychological outcomes stated for the target group (office workers) and the actors at the interpersonal level (colleagues and managers). Behavioral outcomes often comprise more specific subbehaviors (eg, deciding, planning, monitoring), performance objectives, which are influenced by psychosocial determinants (eg, attitude) consisting of subdeterminants (eg, specific beliefs). Only relevant and changeable determinants were identified. Relevance of a determinant refers to the strength of its association with the outcome behavior; changeability refers to the likelihood that the intervention will influence a change in the determinant [22]. We created a matrix, in which performance objectives constitute the rows, and the relevant and changeable determinants the columns. The cells represent the change objectives and provide detailed and measurable information on who and what will change, providing the basis of our intervention.

Intervention Mapping Steps 3 and 4: Program Design and Production

During IM step 3, we selected behavior change methods based on their suitability to cause change in the determinants that needed to be targeted. These were then translated into practical applications by matching the methods to change objectives considering the parameters of use. We focused on a tailored intervention based on two components (each with several objectives), the VitaBit measurement toolkit and the content of UPcomplish (supplied by the personal coach). We further specified scope and sequence of the program and the program theme. In IM step 4, the practical applications were arranged into a coherent program. Program messages and intervention components were drafted and pilot-tested before being refined and produced.

Intervention Mapping Step 5: Adoption and Implementation Plan

In IM step 5, an adoption, implementation, and sustainability plan was created to maximize the likelihood of maintaining behavioral effects and address program dissemination, structural implementation, and maintenance of the intervention. Relevant stakeholders were identified. Behavioral outcomes were formulated and linked to important determinants. The resulting change objectives were used to map an intervention for adopters, implementers, and maintainers by reapplying IM steps 3 and 4.

Intervention Mapping Step 6: Evaluation Plan

IM step 6 focuses on planning an evaluation to determine behavioral and health effects and underlying mechanisms of intervention effectiveness. We collected and designed indicators and measures and planned the design and procedure of the evaluation study.

Results

Intervention Mapping Steps 1 and 2: Needs Assessment and Program Objectives

Program Objectives

Different sedentary behavior parameters have been recommended [36]. This lack of consensus is rooted in both differences in predicted health outcomes (ie, coronary heart diseases vs type 2 diabetes) and recommended behavioral outcomes (ie, daily sitting time vs daily amount of light activity). As a behavioral outcome regarding sedentary behavior, we considered the recommended values from three cohort studies investigating diseases relevant to the target group (ie, heart diseases, diabetes, and all-cause mortality) [9,14,37]. The program objective includes three subobjectives: reduction in daily sitting time, increase in daily light activity, and attainment of a healthy sitting pattern (including fewer long and uninterrupted sitting bouts). The first two subobjectives were

set at a daily sitting time of less than 8 hours per day per person [9,14,37] and a minimum of 4 hours standing and light activity per day [36].

While not only total sitting time is important but also regular sitting interruptions, there is no direct empirical support for the recommendation of a particular sitting pattern. In order to represent the daily sitting pattern, we propose to square the lengths of the daily sitting bouts and to sum them up (summed squared sitting bouts [SSSB]) as shown in Figure 2.

As this is a new representation, a cutoff recommendation relating this value to health outcomes has not yet been investigated. Therefore, based on our baseline activity data (n=69, see OSF repository), we distinguished between healthy and unhealthy sitting patterns by using the median across all days of SSSB as the cutoff ($18.8 * 10^3$ min²). We used the median because the first two subobjectives (sitting and light activity time) were met on about 50% of the days. However, this still needs to be investigated with health outcomes. In spite of similar daily absolute and relative sitting times (see OSF repository), the average duration of sitting bouts collected in longer sitting bouts is significantly smaller on healthy SSSB days, while the amount of sitting in shorter bouts seems to be similar (Figure 3). An SSSB below $18.8 * 10^3$ min² will constitute a healthy sitting pattern according to this pilot study.

Figure 3 represents different average daily sitting minutes collected in certain bout durations on healthy and unhealthy SSSB days (below and above $18.8 * 10^3$ min²) in the pilot study.

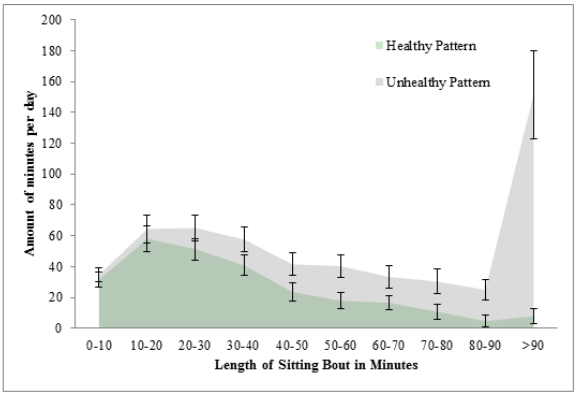
The longer the sitting bout, the less it is represented in a healthy pattern, while time spent in very short sitting bouts is similar between healthy and unhealthy SSSB days. For example, on healthy SSSB days, the individuals spent on average 7.8 minutes of the day in long sitting bouts over 90 minutes (including days without any of these long bouts), while on unhealthy SSSB days, the average time spent in those long bouts was 151.4 minutes. The areas under the curve, therefore, represent the averages of total daily sitting time. Although the average overall sitting time does not differ significantly between healthy and unhealthy SSSB days, this graph clearly shows that on a healthy SSSB day, fewer minutes were collected in longer sitting bouts. We assume that the two sitting patterns differ in terms of health outcome.

The participants in our pilot study met the sitting time objective (maximum 8 hours) with an average of 3.1 days (58.8% of their wearing days), the standing and light activity time objective (minimum 4 hours) with an average of 3.3 days (50.9% of their wearing days), and the SSSB objective (maximum $18.8 * 10^3$ min²) with an average of 2.9 days (54.0% of their wearing days). All three subobjectives were met on an average of 1.4 days (22.8%). Consequently, we specified the following program goal: Participants should achieve all three recommendations on at least 30% of the wearing days in a week (including weekend days). This, at the baseline measurement, was achieved by 26.1% of the participants (control event rate [38]). We would therefore determine effectiveness by the difference of the proportion of participants who meet the program goal after receiving the intervention compared with baseline.

Figure 2. Equation summed squared sitting bouts.

$$SSSB = SitBout_1^2 + SitBout_2^2 + \cdots + SitBout_n^2 = \sum_1^n SitBout_i^2$$

Figure 3. Healthy versus unhealthy summed squared sitting bouts days in the pilot study.



Behavioral Outcomes and Performance Objectives for the Individual Office Worker

At the individual level, the behavioral outcomes were split into a preintentional motivational phase, building an intention to reduce sedentary behavior and preparing for change, and a postintentional volitional phase, translating the intention into behavior [39]. The first behavioral outcome: employees launch a self-regulatory process of controlling their sedentary behavior. This starts with questioning the current behavior and forming an intention to change. It includes monitoring behavior and ends with concrete action planning as indicated by self-set goals. The second behavioral outcome: employees engage in activities in accordance with their previously formulated goals. This focuses on the translation of intentions into behavior by overcoming barriers and actual regular interruptions of sedentary behavior. In addition to this self-regulatory process, other desired behavioral outcomes of the program include establishing good habits and preparing participants for relapses [26,40].

Behavioral Outcomes and Performance Objectives at the Interpersonal Level

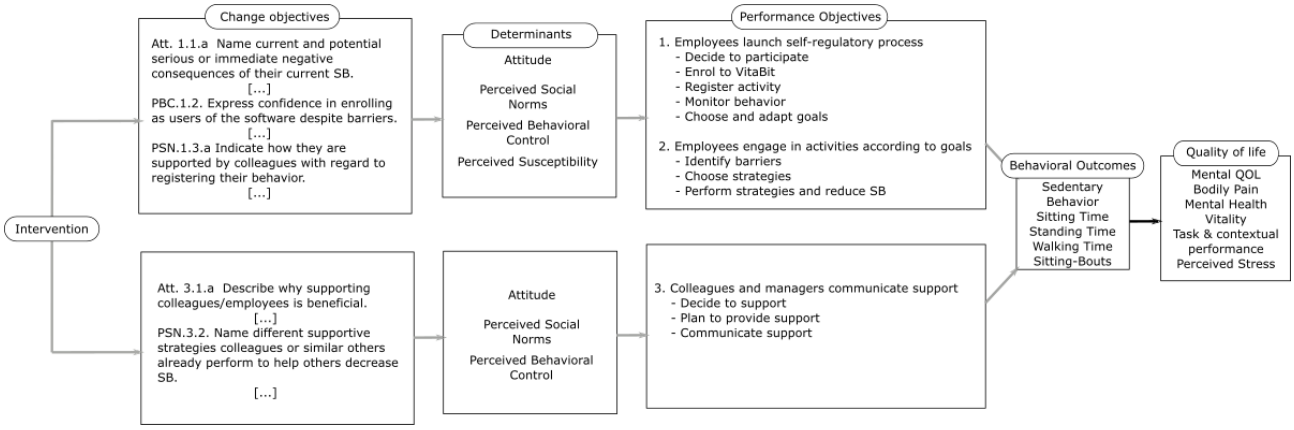
At the interpersonal level, support by colleagues and supervisors is important [30]. Approval from both stakeholders therefore needs to be encouraged and clearly demonstrated. After colleagues and supervisors have decided to show their support, they can apply different supporting strategies. They could decide to participate in a challenge sharing effective strategies for reducing sitting time, as well as joining in and/or initiate standing or walking meetings [41,42]. The support of the supervisors and managers is additionally reflected in the allocation of a room for the kick-off meeting and provision of the funding for the intervention. More information about these two behavioral outcomes can be found in the adoption plan in IM step 5. Supervisors and managers can participate in the program themselves providing similar support to that of the colleagues of the target group [43].

Determinants and Change Objectives at the Individual and Environmental Levels

Empirical evidence from previous sedentary behavior studies was garnered to discover determinants for each performance objective. Since standing is often perceived as being more exhausting than sitting, we included evidence from physical activity research [26]. Identified determinants and their synonyms were covered by the reasoned action approach [44] and the extended parallel process model [41]. The temporal self-regulation theory for physical activity [45] was considered to facilitate the translation of intentions into actual behaviors. Attitudes, perceived social norms, and perceived behavioral control have been shown to explain about 33% of the variance of intention to be less sedentary at work, while 37% of the variance of actual sedentary behavior at work is explained through intention [46]. Since the act of providing support (at an interpersonal level) is a reasoned action, those determinants were also used for agents at the interpersonal level. At the individual level, perceived susceptibility was added as a determinant. A person might only consider making a change if they feel that the threat of negative health outcomes from too much sitting is likely to impact them [47].

Specific underlying beliefs were used to develop change objectives, informed by qualitative literature [26,48] and focus group interviews. For example, in order for an individual to participate, the perceived need to be more active (attitude) and the outlook to receive support (injunctive norm) are critical [49]. The concerning change objective: employees name current and potential serious or immediate negative consequences of their current sedentary behavior. From the temporal self-regulation theory for physical activity, the change objectives related to attitude included the importance of the perceived benefits as being greater and sooner, while the perceived costs were smaller and later. Making those benefits and costs salient at choice time was addressed by the change objectives listed under perceived susceptibility [45]. All change objectives are displayed in the matrices of change objectives (see OSF repository for the matrices and the complete logic model of change). Figure 4 illustrates the logic model of change.

Figure 4. Illustration of the logic model of change.

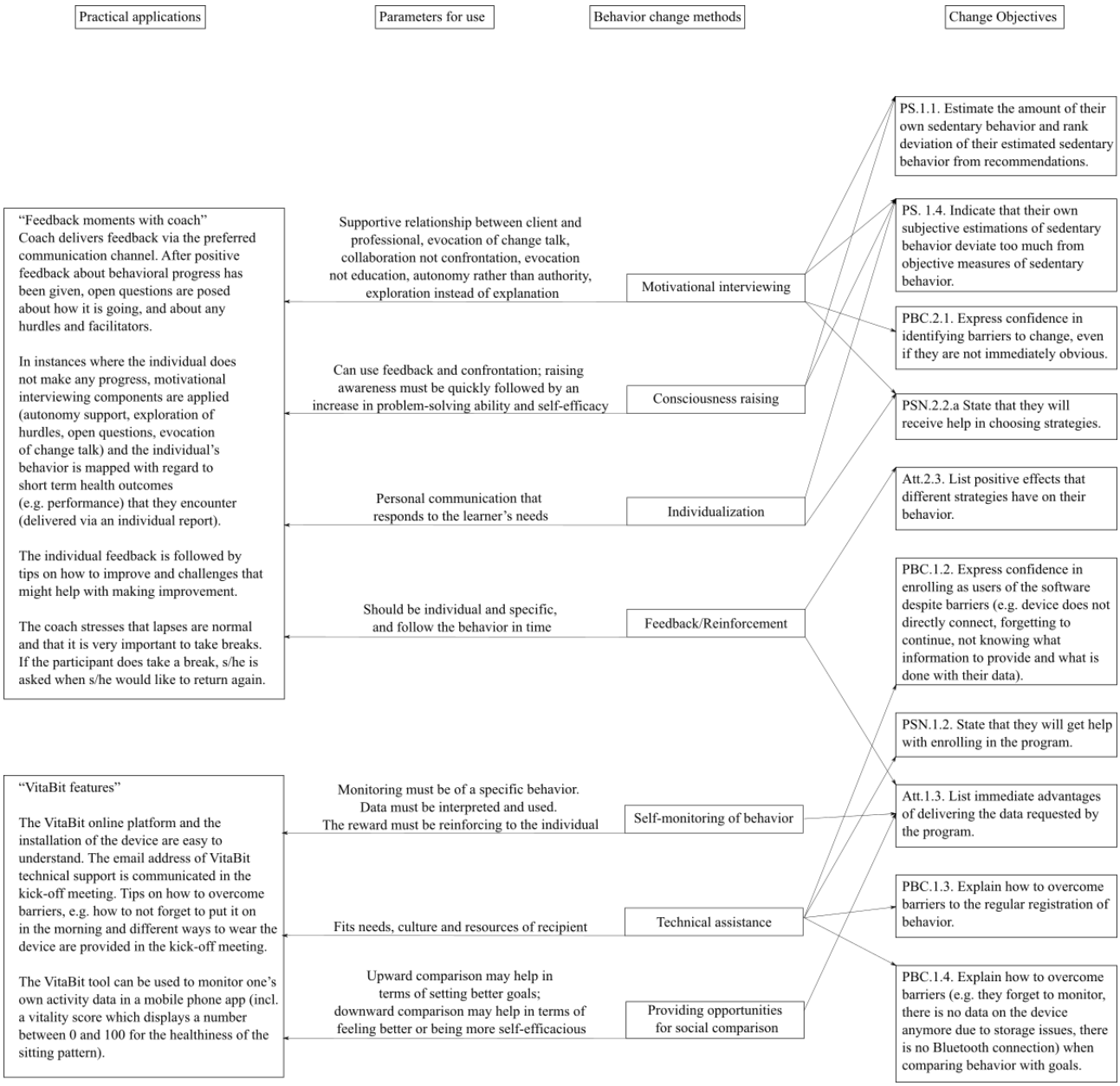


Intervention Mapping Steps 3 and 4: Program Design and Production

Behavior Change Methods and Practical Applications

VitaBit provides the basis for monitoring and delivering individual data, while UPcomplish is provided by a coach to help participants improve their sitting pattern by overcoming individual hurdles. Health professionals and vitality coaches from the field will be the implementers of the intervention, using partly automatized components of UPcomplish (IM step 5). The practical applications can be found in the acyclic behavior change diagrams in the OSF directory, and Figure 5 illustrates examples of important practical applications.

Figure 5. Examples of practical applications.



Program Theme and Sequence

The theme of UPcomplish is based on the assumption that behavioral change in a professional setting should not be too invasive but still motivational. Therefore, the main factors are challenge and low invasiveness. UPcomplish consists of the word up, indicating the goal of the program is supporting desk workers to stand up, and the word accomplish, which reflects the challenging character of the intervention. Getting UP will be accomplished.

The initial phase of preparation and kick-off provides the foundation for the relationship between participant and coach. Participants are introduced to the VitaBit toolkit, familiarize themselves with their own behavior, and get to know the coach. During the kick-off meeting, individualized goals are set, the importance of interrupting sitting is explained, and the preferred communication channel between coach and participant is agreed upon. The baseline phase continues with behavioral and vitality measurements; participants use the VitaBit device for at least 1 week and complete vitality, health, and performance questionnaires including the task and contextual performance subscale of the Individual Work Performance Questionnaire, the Perceived Stress Scale, and the bodily pain, mental health and vitality subscales of the 36-item Short Form Health Survey [50-53] (first of 3 times). During the 3-month trajectory with the coach, participants are provided with activity challenges in biweekly circles. They receive feedback about their behavior 2 times per week and discover facilitators of and hurdles to their behavior through motivational interviewing components. Goals are adjusted after 4 and 8 weeks. In the middle of the intervention, after 6 weeks, participants complete the vitality questionnaire for the second time. In the last 2 weeks, there is a focus on building up habits supported by implementation intentions and the use of buddy systems. At this stage, the vitality questionnaire is completed for the last time [45,54,55].

A group report and individual vitality feedback provide an overview of the participant's achievements (see OSF directory).

Pretests of Program Materials

In order to determine whether the program can be implemented, it needs to be pretested and pilot tested. Pretesting refers to the process whereby specific components of the intervention are tested among the intended population before final production. The goal of pretesting is to safeguard the conditions for effectiveness of the behavior change methods in each component.

Pilot testing is the last evaluation involving all program components, the intended population, and implementers prior to the actual implementation. The goal is to assess the acceptability of the entire program and anticipate any problems in implementation [22].

VitaBit Monitoring Toolkit Pretest

The VitaBit toolkit consists of an accelerometer, mobile phone app, and complementary online platform. These provide the user with tools to monitor their posture patterns with the help of a vitality score (0 = unhealthy, 100 = healthy), set short- and long-term goals, and compare their performance with that of other users. The VitaBit device is a small (3.9 × 1.4 × 0.85 cm, 4.8 g) triaxial wearable accelerometer that monitors sitting, standing, and activity behavior on a half-minute-by-half-minute basis. With regard to sitting, it shows sensitivity and specificity values of 85.7% and 91.2%, respectively [56].

Before the release of the VitaBit toolkit, over 50 pretesters (exact number was not documented) from potential organizations were allocated the VitaBit device, asked to use the device for as long as they liked, and later contacted to provide feedback. This feedback provided information about functionality, design, and features and was translated into improving software components by the VitaBit development team.

UPcomplish Pretest

Initial UPcomplish components were pretested in 11 dispatchers from a German control center. Standing desks were available to these individuals, whose duties mainly involved desk work. A kick-off meeting entailed discussions about the importance of being less sedentary and a short explanation about the intervention and its development. Participants received a weekly progress report. Individual hurdles and facilitators were discussed via their preferred communication channel. Each week, participants received a message in which different performance objectives were addressed, depending on former behaviors and/or reactions to messages (ie, week 1: monitoring behavior; week 2: goal setting; week 3: identifying barriers; etc; see OSF directory). Challenges and other aspects of gamification were not yet included. Two focus group discussions and individual phone calls with participants of this pretest provided feedback about the intervention suggesting that the videos were not watched because they were perceived as being too long, too difficult to download, or too difficult to understand. These video clips were therefore removed from UPcomplish. The kick-off created an atmosphere of trust. However, due to the information about the intervention development being perceived as too lengthy, we decided to shorten the session. The kick-off meeting was also used to help the participants who had not yet tried or succeeded in connecting their device. We decided to split these program components up and call them account creation and pairing the device and that account creation should already be covered before future kick-off meetings to avoid some participants having to wait around. Pairing the device should be handled after the kick-off meeting, in case participants want to directly pair their device with support. The inclusion of challenges and aspects of gamification were not included in this pretest. However, we assumed that these would be attractive and helpful elements. In addition to tailored psychological advice, tailored health advice on individual health outcomes was perceived to be potentially helpful. We decided that motivational interviewing questions should be shortened and performance objectives addressed more frequently, resulting in more frequent delivery of more concise information. Participants showed interest in the vitality score, which provided them with a value between 0 and 100 of how healthy their sitting pattern was.

Pilot Test of UPcomplish

After all adaptations had been made, based on the results of our pretesting, 23 public service desk workers from the Netherlands (5 in the UPcomplish group, who explicitly asked to receive the intervention) took part in our pilot test. After the kick-off meeting, each participant in the UPcomplish group received feedback 2 times per week via their preferred communication channel: individual feedback about goal achievement over the previous week and information regarding sitting patterns on certain weekdays. Furthermore, facilitators of and barriers to sitting less were discussed. Every 2 weeks, participants received gamified challenges. After 4 and 8 weeks, individual goals were revised, if necessary. Summarizing reports completed the intervention. All participants of UPcomplish remained in the program until the end and perceived the coaching to be helpful in terms of reducing sitting time. On average per week, they wore the VitaBit on 74.6% of the days. We observed improvements of sitting, standing, and activity time but cannot interpret them due to the low number of participants and the selectivity of the sample.

Intervention Mapping Step 5: Adoption and Implementation Plan

We expect the managers of our target companies to adopt the intervention, as indicated by the provision of financial funding for the intervention and provision of a room for the kick-off meeting. Additionally, they will supervise and oversee the sustainability of the intervention and its effects. In order to adopt the intervention, the managers first should identify a need to make a decision (eg, determinant: attitude). Second, they should prioritize UPcomplish for individual reasons, such as for an improved reputation of the company (eg, determinant: attitude). Eventually, they should subscribe to the program and continue the subscription for the long run (eg, determinant: perceived behavioral control and attitude) while supervising behavioral maintenance of their employees or institutionalizing the program [43].

Personal talks with the management will address diverse underlying buying preferences. A regular report linking average activity and rates of dropout and commitment, among others, to short-term effects on vitality, performance, mental health, and perceived stress will facilitate positive outcome expectancies. A separate study linking the health outcomes to return on investment is in the planning.

Health professionals and vitality coaches from the field are the implementers of this intervention. It is essential that every component is delivered in the suggested tailored and supportive way in order to maintain program fidelity. Completeness will be accomplished if users receive all of the program components. A workshop for data-driven coaching and meetings with the coaches that implement UPcomplish will maximize fidelity and completeness. A coaching portal in the VitaBit dashboard helps the coach to easily supervise their participants by getting an overview of individual activity patterns. Buttons next to the values of the participants make it possible to deliver the coaches' suggestions directly to the relevant participants or to get an overview of their dashboards. Multimedia Appendix 1 shows an example of the coaching portal. The average sitting, standing, and walking parameters for a given period of time are displayed on one page. On the right, the coach can send individual notifications and emails, inspect the individual portal to get more detailed insights into the daily activity behaviors, and create new widgets, such as setting a new goal.

Mobile phone–based workplace sedentary behavior interventions seem to be especially effective in the medium term (3- to 6-month follow-up) if they incorporate several behavior change methods [21]. These include self-monitoring and prompts or cues combined with information about health consequences and information about how to perform the desired behavior. In order to facilitate program sustainability, it is important to tailor the maintenance intervention to the participants who sit the most during their workdays, or, more generally, to those with different motivational profiles, such as a focus on health promotion versus weight loss versus illness prevention [21,57]. The coaches are encouraged to stress the importance of buddy systems and deliver regular short and precise health information in order to stabilize attitudes about sedentary behavior in the target group. Optional email reminders and health blogs help users to be reminded of the importance of reducing sitting time. Analyzing dropouts in the process evaluation and preventing reasons for future dropouts will help to facilitate program sustainability [22].

Intervention Mapping Step 6: Evaluation Plan

We plan to evaluate short-term effectiveness in terms of decreased sitting time, SSSBs, and increased standing and walking time (secondary effects on short-term quality-of-life outcomes [7,58,59]) of UPcomplish (effect evaluation). Furthermore, we will consider whether any program adaptations are needed and what these might be (process evaluations). We will employ a multilevel design with between-subjects and within-subjects factor (measurement moment) comparisons and estimate the intervention's effect in a magnitude of standard deviations (Cohen *d*) to enable the computation of the number needed to treat (number of people that should receive the intervention for one person to change their behavior sufficiently to meet the criteria specified in the intervention goal [38]). The number needed to treat can be used to calculate the cost of the intervention needed for at least 30% of the participants to achieve all three behavioral outcomes.

From May 2019 until January 2020, we had 200 VitaBit monitors at our disposal. We chose a stepped-wedge design (last week of one group is compared with first week of another group) because a control group (VitaBit only) was not possible considering high expected dropout rates and feasibility issues. Splitting up intervention groups into as many groups as possible would reveal a bigger sample size since some groups could provide data for both the baseline and postintervention measurement. Having five different intervention groups was considered the minimum yet doable number of groups where one group can provide data for the two measurements. The five intervention groups, each comprising 40 participants, start with a time lag of 7 weeks. With an anticipated retention rate of 80%, this yields an analyzable sample size of n=192 [60]. With 192 participants, estimation of this effect size is accurate to about a quarter of the standard deviation (see the OSF repository for details and a flowchart illustrating the design).

The process evaluation is informed by qualitative and quantitative output from surveys and behavioral data and will assess both important aspects of the logic model of change and intervention components.

Procedure

Groups of 10 to 15 desk workers from random companies in Germany are recruited via email and personal contacts. Potential groups are randomly assigned to one of the intervention groups and informed about the intervention and the measurements before consent is obtained. Each group receives the 12-week intervention and is requested to complete vitality, performance, and mental health questionnaires at 3 points in time. Participants can refuse participation in the intervention and/or the measurements at any times without giving a reason. The evaluation of this intervention including its consent procedure was approved by the Ethical Review Committee, Psychology and Neuroscience, Maastricht University, the Netherlands (ERCPN- 188_11_02_2018). More details can be found in "The Evaluation of UPcomplish: Sample size planning and procedure" in the OSF directory.

Measures Process Evaluation

All questionnaires can be found in the OSF repository and were translated into German using the back-translation method if no validated German version was available [61]. Table 1 provides an overview of all measurements that are used in the evaluation.

Table 1. Measurements and example items.

Variable	Measurements and indicators	Items	Example item	Point in time
Intervention characteristics				
Acceptability	Taken from a former evaluation [48]	18	"The questions within the recommendations were clear."	End
Fidelity	Messages from automated pool divided by total amount of messages sent by the coach	N/A[a]	N/A	End
Reach	Dropout rate; ratio of participants from the intended target group; dose received	N/A	N/A	End
Determinants				
Attitude	Taken from a former evaluation [48]	6	"Standing and walking around at work is healthy."	Baseline, middle, end
PSN[b]	Taken from a former evaluation [48]	2	"Standing and walking around at work is encouraged by my colleagues."	Baseline, middle, end
PBC[c]	Taken from a former evaluation [48]	4	"I am sure that I can stand and walk around at work, even though I feel bad, tired, tense or depressed."	Baseline, middle, end
Perceived susceptibility	Self-created questions to assess perceived susceptibility to improper sitting habits	2	"My daily sitting time is more than what is recommended."	Baseline, middle, end
Performance objectives				
PO[d] 1.2 Enrollment as VitaBit user	Proportion of successfully enrolled participants among the ones who agreed to participate	N/A	N/A	End
PO 1.3 Registration of sedentary and antagonistic behaviors	Average of days per week that show VitaBit data for at least 6 hours	N/A	N/A	End
PO 1.4 Monitoring of behavior	Number of days missing before the feedback moments	N/A	N/A	End
Action planning, identifying barriers and facilitators, and support	Numbers and quality of responses to coaching questions/requests	N/A	N/A	End
Sedentary behavior and physical activity				
Objectively measured sitting (30-second periods)	VitaBit measurement toolkit [56,62]	N/A	N/A	continuously
Moderate and vigorous physical activity	German version of the International Physical Activity Questionnaire (short form) [63]	max. 6	"During the last 7 days, on how many days did you do vigorous physical activities like heavy lifting, digging, heavy construction, or climbing stairs as part of your work? Think about only those physical activities that you did for at least 10 minutes at a time."	Baseline, middle, end
Secondary outcome: quality-of-life				
Task and contextual performance	Two subscales of the Individual Work Performance Questionnaire [50]	14	"In the past week, I took on extra responsibilities."	Baseline, middle, end
Stress perception	Perceived Stress Scale [51,53]	10	"In the last week, how often have you felt nervous and "'stressed'?"	Baseline, middle, end
Bodily pain	Subscale of the SF[e]-36 health survey [52]	2	"How much bodily pain have you had during the past week?"	Baseline, middle, end
Mental health	Subscale of the SF-36 health survey [52]	5	"How much of the time during the past week have you been a happy person?"	Baseline, middle, end
Vitality	Subscale of the SF-36 health survey [52]	4	"How much of the time during the past week did you have a lot of energy?"	Baseline, middle, end
Covariates: demographic, educational, and job-related variables				

Variable	Measurements and indicators	Items	Example item	Point in time
Gender, age, educational level, height, weight, and job-related variables (eg, team size)	Measured by VitaBit during account creation	8	N/A	Baseline
Job tasks	Taken from a former evaluation [48]	5	"How much on average per day (in %) do you estimate you spend on the following tasks? Phone calls?"	Baseline
Employment status and working times	Self-created questions	2	"How many days do you usually work in a week?"	Baseline

[a]N/A: not applicable.

[b]PSN: perceived social norms.

[c]PBC: perceived behavioral control.

[d]PO: performance objective.

[e]SF: Short Form Health Survey.

Statistical Analyses

Statistical analyses encompass multilevel analyses. For the between-subject comparisons, the outcome variables are centered around baseline company means, and the analyses are nested by calendar week. For the within-subject comparisons, the outcome variables are centered around calendar weeks, and the analyses are nested on the individual level. Analyses are adjusted for possible confounding variables such as company-related variables, gender, or age.

Multilevel linear and logistic regressions are conducted to inspect putative effects of performance objectives and determinants on the continuous primary outcome variables and the dichotomous performance objectives (performed yes/no), respectively.

Discussion

Principal Findings

This paper describes an IM protocol to develop a computer-based intervention aimed at reducing sedentary behavior at work. A tailored intervention was developed to guide participants step by step through a behavioral change process. The support of both colleagues and supervisors was considered and addressed in additional components. A plan to ensure adoption, implementation, and sustainability was drafted. Finally, we developed an evaluation plan for assessing the effects of the intervention and the mechanisms behind these effects.

Strengths and Limitations

Although the IM approach suggests working through all the core processes, not all substeps were performed in our study [64], partly due to the fact that additional research (eg, about the necessity of all behavioral substeps [ie, performance objectives]) was still ongoing. Still, we plan to complete a process evaluation that will investigate mechanisms of effectiveness and provide additional information. A second limitation is that members of the target group and managers of companies who might potentially use the intervention (except those working at VitaBit) were contacted too late to be part of the planning group since they were only contacted as part of the pretest and pilot test. Nevertheless, the interest in reducing sedentary behavior seems to be high, and multiple informal talks during the development process with potential adopters, implementers, and people from the target group have revealed valuable insights.

A benefit of the project was the collaboration between scientific research and information technology practice. To facilitate this collaboration, face-to-face and Skype discussions were used to directly exchange ideas and possibilities. In doing so, we also discovered more challenging aspects of collaboration between health promotion and information technology practice. The usage of technical terms on both sides, different priorities during the development process, and balancing act between tailoring and standardization are examples of the challenges we encountered. However, working together allowed for a quick translation of knowledge about behavioral change into practical applications and provides an example that can be applied to other IM procedures [65,66].

Conclusion

We developed a comprehensive intervention targeting important determinants at two different ecological levels. The development of our intervention was grounded in relevant literature, and multiple theories have been applied. Future evaluation studies should investigate the program effectiveness and further analyze the relevance and utility of single program components.

Acknowledgments

This study was supported by VitaBit Software International BV. We would like to thank VitaBit for the provision of the VitaBit devices and technical support. Furthermore, we want to thank Elena Magazin and Davy Baars for proofreading the paper.

Authors' Contributions

NMB, GAtH, GK, RACR, and GP conceived and designed the study. NMB performed the experiments and analyzed the data. GP und GJYP contributed to the development of materials. NMB, GAtH, GJYP, and GP drafted the manuscript. All authors read, provided feedback on, and approved the manuscript.

Conflicts of Interest

NMB is employed by VitaBit to develop a sedentary behavior intervention. Apart from providing the devices and access to the software and the technical translation of the intervention components into practice, VitaBit had no role in the design of the studies; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Multimedia Appendix 1
Coaching overview of VitaBit portal.
[PNG File , 57 KB - formative_v4i7e14951_app1.png]

References

1. Ryan CG, Dall PM, Granat MH, Grant PM. Sitting patterns at work: objective measurement of adherence to current recommendations. Ergonomics 2011 Jun;54(6):531-538. [doi: 10.1080/00140139.2011.570458]
2. Clemes SA, O'Connell SE, Edwardson CL. Office workers' objectively measured sedentary behavior and physical activity during and outside working hours. J Occup Environ Med 2014 Mar;56(3):298-303. [doi: 10.1097/JOM.0000000000000101] [Medline: 24603203]
3. Biswas A, Oh PI, Faulkner GE, Bajaj RR, Silver MA, Mitchell MS, et al. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. Ann Intern Med 2015 Jan 20;162(2):123-132. [doi: 10.7326/M14-1651] [Medline: 25599350]
4. Wilmot EG, Edwardson CL, Achana FA, Davies MJ, Gorely T, Gray LJ, et al. Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. Diabetologia 2012 Nov;55(11):2895-2905. [doi: 10.1007/s00125-012-2677-z] [Medline: 22890825]
5. van Uffelen JGZ, Wong J, Chau JY, Riphagen I, Gilson ND, Burton NW, et al. Occupational sitting and health risks: a systematic review. Am J Prev Med 2010 Oct;39(4):379-388. [doi: 10.1016/j.amepre.2010.05.024] [Medline: 20837291]
6. Chau JY, Merom D, Chey T, Bauman AE. Cross-sectional associations between occupational and leisure-time sitting, physical activity and obesity in working adults. Prev Med 2012;54(3-4):195-200. [doi: 10.1016/j.ypmed.2011.12.020] [Medline: 22227284]
7. Hamer M, Stamatakis E. Prospective study of sedentary behavior, risk of depression, and cognitive impairment. Med Sci Sports Exerc 2014 Apr;46(4):718-723 [FREE Full text] [doi: 10.1249/MSS.0000000000000156] [Medline: 24121248]
8. Voss M, Carr L, Clark R, Weng T. Revenge of the “sit” II: Does lifestyle impact neuronal and cognitive health through distinct mechanisms associated with sedentary behavior and physical activity? Mental Health Phys Activ 2014 Mar;7(1):9-24. [doi: 10.1016/j.mhpa.2014.01.001]
9. Chau JY, Grunseit A, Midthjell K, Holmen J, Holmen TL, Bauman AE, et al. Sedentary behaviour and risk of mortality from all-causes and cardiometabolic diseases in adults: evidence from the HUNT3 population cohort. Br J Sports Med 2015 Jun;49(11):737-742. [doi: 10.1136/bjsports-2012-091974] [Medline: 23666019]
10. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, SBRN Terminology Consensus Project Participants. Sedentary Behavior Research Network (SBRN) - terminology consensus project process and outcome. Int J Behav Nutr Phys Act 2017 Jun 10;14(1):75 [FREE Full text] [doi: 10.1186/s12966-017-0525-8] [Medline: 28599680]
11. Hamilton MT, Healy GN, Dunstan DW, Zderic TW, Owen N. Too little exercise and too much sitting: inactivity physiology and the need for new recommendations on sedentary behavior. Curr Cardiovasc Risk Rep 2008 Jul;2(4):292-298 [FREE Full text] [doi: 10.1007/s12170-008-0054-8] [Medline: 22905272]
12. Thorp AA, Owen N, Neuhaus M, Dunstan DW. Sedentary behaviors and subsequent health outcomes in adults a systematic review of longitudinal studies, 1996-2011. Am J Prev Med 2011 Aug;41(2):207-215. [doi: 10.1016/j.amepre.2011.05.004] [Medline: 21767729]
13. Ekelund U, Steene-Johannessen J, Brown WJ, Fagerland MW, Owen N, Powell KE, Lancet Physical Activity Series 2 Executive Committe, Lancet Sedentary Behaviour Working Group. Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. Lancet 2016 Sep 24;388(10051):1302-1310. [doi: 10.1016/S0140-6736(16)30370-1] [Medline: 27475271]
14. Pandey A, Salahuddin U, Garg S, Ayers C, Kulinski J, Anand V, et al. Continuous dose-response association between sedentary time and risk for cardiovascular disease: a meta-analysis. JAMA Cardiol 2016 Aug 01;1(5):575-583. [doi: 10.1001/jamacardio.2016.1567] [Medline: 27434872]
15. Bankoski A, Harris TB, McClain JJ, Brychta RJ, Caserotti P, Chen KY, et al. Sedentary activity associated with metabolic syndrome independent of physical activity. Diabetes Care 2011 Feb;34(2):497-503 [FREE Full text] [doi: 10.2337/dc10-0987] [Medline: 21270206]

16. Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, et al. Breaks in sedentary time: beneficial associations with metabolic risk. Diabetes Care 2008 Apr;31(4):661-666. [doi: 10.2337/dc07-2046] [Medline: 18252901]
17. Dunstan DW, Kingwell BA, Larsen R, Healy GN, Cerin E, Hamilton MT, et al. Breaking up prolonged sitting reduces postprandial glucose and insulin responses. Diabetes Care 2012 May;35(5):976-983 [FREE Full text] [doi: 10.2337/dc11-1931] [Medline: 22374636]
18. Owen N, Sugiyama T, Eakin EE, Gardiner PA, Tremblay MS, Sallis JF. Adults' sedentary behavior determinants and interventions. Am J Prev Med 2011 Aug;41(2):189-196. [doi: 10.1016/j.amepre.2011.05.013] [Medline: 21767727]
19. Gupta N, Heiden M, Aadahl M, Korshøj M, Jørgensen MB, Holtermann A. What Is the effect on obesity indicators from replacing prolonged sedentary time with brief sedentary bouts, standing and different types of physical activity during working days? A cross-sectional accelerometer-based study among blue-collar workers. PLoS One 2016;11(5):e0154935 [FREE Full text] [doi: 10.1371/journal.pone.0154935] [Medline: 27187777]
20. Gardner B, Smith L, Lorencatto F, Hamer M, Biddle SJ. How to reduce sitting time? A review of behaviour change strategies used in sedentary behaviour reduction interventions among adults. Health Psychol Rev 2015 Sep 16:1-24. [doi: 10.1080/17437199.2015.1082146] [Medline: 26315814]
21. Stephenson A, McDonough SM, Murphy MH, Nugent CD, Mair JL. Using computer, mobile and wearable technology enhanced interventions to reduce sedentary behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2017 Aug 11;14(1). [doi: 10.1186/s12966-017-0561-4]
22. Bartholomew Eldredge LK, Markham CM, Ruiter RA, Fernández ME, Kok G, Parcel GS. Planning Health Promotion Programs: An Intervention Mapping Approach, 4th Edition. Philadelphia: John Wiley & Sons; 2016.
23. Kok G, Gottlieb NH, Peters GY, Mullen PD, Parcel GS, Ruiter RAC, et al. A taxonomy of behaviour change methods: an Intervention Mapping approach. Health Psychol Rev 2015 Oct 15:1-16. [doi: 10.1080/17437199.2015.1077155] [Medline: 26262912]
24. Coffeng JK, Boot CRL, Duijts SFA, Twisk JWR, Hendriksen IJM. Effectiveness of a worksite social & physical environment intervention on need for recovery, physical activity and relaxation; results of a randomized controlled trial. PLoS One 2014;9(12):e114860 [FREE Full text] [doi: 10.1371/journal.pone.0114860] [Medline: 25542039]
25. Kwak L, Kremers SP, Candel MJ, Visscher TL, Brug J. Changes in skinfold thickness and waist circumference after 12 and 24 months resulting from the NHF-NRG In Balance-project. Int J Behav Nutr Phys Act 2010;7:26 [FREE Full text] [doi: 10.1186/1479-5868-7-26] [Medline: 20370934]
26. McEachan RRC, Lawton RJ, Jackson C, Conner M, Lunt J. Evidence, theory and context: using intervention mapping to develop a worksite physical activity intervention. BMC Public Health 2008 Sep 22;8:326 [FREE Full text] [doi: 10.1186/1471-2458-8-326] [Medline: 18808709]
27. Coffeng JK, Hendriksen IJM, Duijts SF, Proper KI, Boot CRL. The development of the Be Active & Relax. BMC Public Health 2012 Aug 02;12:592 [FREE Full text] [doi: 10.1186/1471-2458-12-592] [Medline: 22852835]
28. Kwak L, Kremers SPJ, Werkman A, Visscher TLS, van Baak MA, Brug J. The NHF-NRG in balance-project: the application of intervention mapping in the development, implementation and evaluation of weight gain prevention at the worksite. Obes Rev 2007 Jul;8(4):347-361. [doi: 10.1111/j.1467-789X.2006.00304.x] [Medline: 17578384]
29. Crutzen R, Peters GY. Evolutionary learning processes as the foundation for behaviour change. Health Psychol Rev 2018 Mar;12(1):43-57. [doi: 10.1080/17437199.2017.1362569] [Medline: 28764599]
30. Schoeppe S, Alley S, Bray NA, Williams SL, Duncan MJ, Vandelanotte C. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
31. Broekhuizen K, Kroeze W, Oenema A, Brug J. A systematic review of randomized controlled trials on the effectiveness of computer-tailored physical activity and dietary behavior promotion programs: an update. Ann Behav Med 2012 Oct;44(2):259-286 [FREE Full text] [doi: 10.1007/s12160-012-9384-3] [Medline: 22767052]
32. Kelders SM, Kok RN, Ossebaard HC. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
33. Gardner B, Smith L, Lorencatto F, Hamer M, Biddle SJ. How to reduce sitting time? A review of behaviour change strategies used in sedentary behaviour reduction interventions among adults. Health Psychol Rev 2015 Sep 16:1-24. [doi: 10.1080/17437199.2015.1082146] [Medline: 26315814]
34. Preventing sedentary behavior at work: An Intervention Mapping approach for data-driven mhealth consulting. Open Science Framework repository. URL: https://osf.io/8vu37/?view_only=5c39c5a2e6184ca68eb1ce4c8fa17bfa [accessed 2020-05-24]
35. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. Lancet Glob Health 2018 Oct;6(10):e1077-e1086 [FREE Full text] [doi: 10.1016/S2214-109X(18)30357-7] [Medline: 30193830]
36. Buckley JP, Hedge A, Yates T, Copeland RJ, Loosemore M, Hamer M, et al. The sedentary office: an expert statement on the growing case for change towards better health and productivity. Br J Sports Med 2015 Nov;49(21):1357-1362. [doi: 10.1136/bjsports-2015-094618] [Medline: 26034192]

37. van der Berg JD, Stehouwer CDA, Bosma H, Willems PJB, Savelberg HHCM, Schram MT, et al. Associations of total amount and patterns of sedentary behaviour with type 2 diabetes and the metabolic syndrome: The Maastricht Study. Diabetologia 2016 Apr;59(4):709-718 [FREE Full text] [doi: 10.1007/s00125-015-3861-8] [Medline: 26831300]
38. Gruijters S, Peters GJ. Gauging the impact of behavior change interventions: a tutorial on the numbers needed to treat. PsyArXiv 2017. [doi: 10.31234/osf.io/2bau7]
39. Schwarzer R. Modeling health behavior change: how to predict and modify the adoption and maintenance of health behaviors. Appl Psychol 2008 Jan;57(1):1-29. [doi: 10.1111/j.1464-0597.2007.00325.x]
40. Sniehotta FF, Presseau J, Hobbs N, Araújo-Soares V. Testing self-regulation interventions to increase walking using factorial randomized N-of-1 trials. Health Psychol 2012 Nov;31(6):733-737. [doi: 10.1037/a0027337] [Medline: 22353025]
41. 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]
42. Giles-Corti B, Donovan RJ. Relative influences of individual, social environmental, and physical environmental correlates of walking. Am J Public Health 2003 Sep;93(9):1583-1589. [doi: 10.2105/ajph.93.9.1583] [Medline: 12948984]
43. Butterfoss F, Kegler M, Francisco V. Mobilizing organizations for health promotion: theories of organizational change. In: Glanz K, Rimer B, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, and Practice, 4th Edition. San Francisco: Jossey-Bass; 2008:335-361.
44. Fishbein M. Predicting and Changing Behavior: The Reasoned Action Approach. Philadelphia: Taylor & Francis; 2011.
45. Hall PA, Fong GT. Temporal self-regulation theory: a neurobiologically informed model for physical activity behavior. Front Hum Neurosci 2015;9:117 [FREE Full text] [doi: 10.3389/fnhum.2015.00117] [Medline: 25859196]
46. Prapavessis H, Gaston A, DeJesus S. The Theory of Planned Behavior as a model for understanding sedentary behavior. Psychol Sport Exercise 2015 Jul;19:23-32. [doi: 10.1016/j.psychsport.2015.02.001]
47. Witte K. Putting the fear back into fear appeals: the extended parallel process model. Comm Monographs 1992 Dec;59(4):329-349. [doi: 10.1080/03637759209376276]
48. De Cocker CK, Cardon G, Vandelanotte C. Theory-driven, web-based, computer-tailored advice to reduce and interrupt sitting at work: development, feasibility and acceptability testing among employees. BMC Public Health 2015;15(1):959 [FREE Full text] [doi: 10.1186/s12889-015-2288-y] [Medline: 26404915]
49. Bardus M, Blake H, Lloyd S, Suggs L. Reasons for participating and not participating in a e-health workplace physical activity intervention. Intl J of Workplace Health Mgt 2014 Nov 04;7(4):229-246. [doi: 10.1108/IJWHM-11-2013-0040]
50. Koopmans L, Bernaards CM, Hildebrandt VH, van Buuren S, van der Beek AJ, de Vet HCW. Improving the Individual Work Performance Questionnaire using Rasch analysis. J Appl Meas 2014;15(2):160-175. [Medline: 24950534]
51. Cohen S. Perceived stress scale. Measuring stress: a guide for health and social scientists 1994:235-283. [doi: 10.1037/t02889-000]
52. Ware JE. SF-36 health survey update. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3130-3139. [Medline: 11124729]
53. Klein EM, Brähler E, Dreier M, Reinecke L, Müller KW, Schmutzer G, et al. The German version of the Perceived Stress Scale—psychometric characteristics in a representative German community sample. BMC Psychiatry 2016 Dec 23;16:159 [FREE Full text] [doi: 10.1186/s12888-016-0875-9] [Medline: 27216151]
54. Hagger MS, Luszczynska A, de Wit J, Benyamini Y, Burkert S, Chamberland P, et al. Implementation intention and planning interventions in Health Psychology: recommendations from the Synergy Expert Group for research and practice. Psychol Health 2016 Jul;31(7):814-839. [doi: 10.1080/08870446.2016.1146719] [Medline: 26892502]
55. McNeill LH, Kreuter MW, Subramanian SV. Social environment and physical activity: a review of concepts and evidence. Soc Sci Med 2006 Aug;63(4):1011-1022. [doi: 10.1016/j.socscimed.2006.03.012] [Medline: 16650513]
56. Berninger NM, Ten Hoor GA, Plasqui G. Validation of the VitaBit Sit-Stand Tracker: detecting sitting, standing, and activity patterns. Sensors (Basel) 2018 Mar 15;18(3) [FREE Full text] [doi: 10.3390/s18030877] [Medline: 29543766]
57. Fukuoka Y, Lindgren TG, Mintz YD, Hooper J, Aswani A. Applying natural language processing to understand motivational profiles for maintaining physical activity after a mobile app and accelerometer-based intervention: the mped randomized controlled trial. JMIR Mhealth Uhealth 2018 Jun 20;6(6):e10042 [FREE Full text] [doi: 10.2196/10042] [Medline: 29925491]
58. Hendriksen IJM, Bernaards CM, Steijn WMP, Hildebrandt VH. Longitudinal relationship between sitting time on a working day and vitality, work performance, presenteeism, and sickness absence. J Occup Environ Med 2016 Aug;58(8):784-789 [FREE Full text] [doi: 10.1097/JOM.0000000000000809] [Medline: 27299213]
59. Dunlop DD, Song J, Arnston EK, Semanik PA, Lee J, Chang RW, et al. Sedentary time in US older adults associated with disability in activities of daily living independent of physical activity. J Phys Act Health 2015 Jan;12(1):93-101 [FREE Full text] [doi: 10.1123/jpah.2013-0311] [Medline: 24510000]
60. Waters LA, Galichet B, Owen N, Eakin E. Who participates in physical activity intervention trials? J Phys Act Health 2011 Jan;8(1):85-103. [Medline: 21297189]
61. Brislin RW. Back-translation for cross-cultural research. J Cross-Cultural Psychol 1970 Sep 01;1(3):185-216. [doi: 10.1177/135910457000100301]
62. Atkin AJ, Gorely T, Clemes SA, Yates T, Edwardson C, Brage S, et al. Methods of measurement in epidemiology: sedentary behaviour. Int J Epidemiol 2012 Oct;41(5):1460-1471 [FREE Full text] [doi: 10.1093/ije/dys118] [Medline: 23045206]

63. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
64. Ruiter R, Crutzen R, Kok G. Core processes for developing theory-and evidence-based interventions. 2018. URL: https://osf.io/pe4hv/download/?version=1&displayName=core_processes_preprint-2018-10-01T09%3A31%3A40.904Z.pdf [accessed 2020-05-21]
65. Smith-Dektor AH, Young SD. Marketing, technology, and medicine: recommendations on how to incorporate psychological principles into new technologies to promote healthy behaviors. J Consum Health Internet 2014 Jul 01;18(3):253-259 [FREE Full text] [doi: 10.1080/15398285.2014.932182] [Medline: 26052262]
66. Montag C, Duke E, Markowetz A. Toward psychoinformatics: computer science meets psychology. Comput Math Methods Med 2016;2016:2983685 [FREE Full text] [doi: 10.1155/2016/2983685] [Medline: 27403204]

Abbreviations

IM: intervention mapping
OSF: Open Science Framework
SSSB: summed squared sitting bouts

Edited by G Eysenbach; submitted 05.06.19; peer-reviewed by M Miyachi, C Thomson; comments to author 03.10.19; revised version received 12.11.19; accepted 23.03.20; published 20.07.20.

Please cite as:
Berninger NM, ten Hoor GA, Plasqui G, Kok G, Peters GJY, Ruiter RAC
Sedentary Work in Desk-Dominated Environments: A Data-Driven Intervention Using Intervention Mapping
JMIR Form Res 2020;4(7):e14951
URL: http://formative.jmir.org/2020/7/e14951/
doi:10.2196/14951
PMID:32706695

Original Paper

A Digital Library for Increasing Awareness About Living Donor Kidney Transplants: Formative Study

Amy D Waterman[1,2], PhD; Emily H Wood[1], MPH; Omesh N Ranasinghe[1], MPH; Amanda Faye Lipsey[2], MFA; Crystal Anderson[1], MPH; Wendy Balliet[3], PhD; Lauren Holland-Carter[3], PhD; Stacey Maurer[3], PhD; Maria Aurora Posadas Salas[3], MD

[1]Division of Nephrology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States

[2]Terasaki Research Institute, Los Angeles, CA, United States

[3]Medical University of South Carolina, Charleston, SC, United States

Corresponding Author:
Amy D Waterman, PhD
Division of Nephrology
David Geffen School of Medicine
University of California, Los Angeles
1018 Westwood Blvd
Los Angeles, CA, 90024
United States
Phone: 1 4243277278
Email: awaterman@mednet.ucla.edu

Abstract

Background: It is not common for people to come across a living kidney donor, let alone consider whether they would ever donate a kidney themselves while they are alive. Narrative storytelling, the sharing of first-person narratives based on lived experience, may be an important way to improve education about living donor kidney transplants (LDKTs). Developing ways to easily standardize and disseminate diverse living donor stories using digital technology could inspire more people to consider becoming living donors and reduce the kidney shortage nationally.

Objective: This paper aimed to describe the development of the Living Donation Storytelling Project, a web-based digital library of living donation narratives from multiple audiences using video capture technology. Specifically, we aimed to describe the theoretical foundation and development of the library, a protocol to capture diverse storytellers, the characteristics and experiences of participating storytellers, and the frequency with which any ethical concerns about the content being shared emerged.

Methods: This study invited kidney transplant recipients who had received LDKTs, living donors, family members, and patients seeking LDKTs to record personal stories using video capture technology by answering a series of guided prompts on their computer or smartphone and answering questions about their filming experience. The digital software automatically spliced responses to open-ended prompts, creating a seamless story available for uploading to a web-based library and posting to social media. Each story was reviewed by a transplant professional for the disclosure of protected health information (PHI), pressuring others to donate, and medical inaccuracies. Disclosures were edited.

Results: This study recruited diverse storytellers through social media, support groups, churches, and transplant programs. Of the 137 storytellers who completed the postsurvey, 105/137 (76.6%) were white and 99/137 (72.2%) were female. They spent 62.5 min, on average, recording their story, with a final median story length of 10 min (00:46 seconds to 32:16 min). A total of 94.8% (130/137) of storytellers were motivated by a desire to educate the public; 78.1% (107/137) were motivated to help more people become living donors; and 75.9% (104/137) were motivated to dispel myths. The ease of using the technology and telling their story varied, with the fear of being on film, emotional difficulty talking about their experiences, and some technological barriers being reported. PHI, most commonly surnames and transplant center names, was present in 62.9% (85/135) of stories and was edited out.

Conclusions: With appropriate sensitivity to ensure diverse recruitment, ethical review of content, and support for storytellers, web-based storytelling platforms may be a cost-effective and convenient way to further engage patients and increase the curiosity of the public in learning more about the possibility of becoming living donors.

(JMIR Form Res 2020;4(7):e17441) doi:10.2196/17441

KEYWORDS

living donor kidney transplant; living donation; health education; informed decision-making; awareness; health literacy; video library; health technology; kidney diseases; diffusion of innovation; digital library; mobile phone

Introduction

Presently, over 740,000 people in the United States are living without functioning kidneys due to end-stage kidney disease (ESKD) [1,2]. In general, patients who can receive a living donor kidney transplant (LDKT) from a family member or friend live longer than those remaining on dialysis or waiting for years for a kidney from someone who has died [1,3]. The demand for kidneys continues to outweigh the supply; while nearly 100,000 patients are currently waiting for a kidney on the national transplant waiting list [4], LDKT rates have declined by 12% over the last decade, generally not exceeding 6500 kidneys annually [1]. There are also significant ethnic/racial disparities in LDKT [1,5-8]; in the last 15 years, Latinx, black, and Asian patients have become even less likely to get an LDKT compared with white patients than they were in the past [5], and they are also less likely to be donors [1].

Although increasing deceased donation rates is limited by practical and medical circumstances surrounding how individuals die, living donation rates are limited only by the number of healthy individuals who become aware, educated, and interested in donating 1 kidney while they are alive. Of the roughly 250 million adults in the United States, only 100,000 more individuals (0.04%) would need to agree to donate 1 kidney to eliminate the entire kidney donor shortage. Education strategies to increase LDKT commonly target patients and families using face-to-face educational sessions [9-14], home-based educational interventions [9,15-17], culturally targeted videos and websites in multiple languages, and decision-making aids [7,18-20]. Although effective, these interventions fail to reach (1) the general public [21,22]; (2) family members and friends who do not come to a transplant center to learn [23]; (3) kidney patients who are scared to ask others to donate on their behalf [19]; and (4) members of specific ethnic/minority groups who have cultural sensitivities to living donation [24], low health literacy [25], or greater medical mistrust of the health care establishment that are unaddressed through general education [26].

To expand the living donor pool, we need to reach beyond the walls of the transplant center to help patients share their interest in LDKT with more individuals and inspire more people who are still unaware of living donation to consider becoming donors. As few people know a living donor personally, we also need to help the general public realize that other people who look like them donate kidneys each year. Innovative strategies to educate and inspire more patients and potential living donors to consider living donation are needed.

Storytelling, the sharing of first-person narratives based on lived experience, is an educational approach that is authentic, emotional, and provides people with the opportunity to learn from others who look like them. Stories have the power to emotionally engage listeners, reach low-literacy audiences, and present complex information in informal and comprehensible ways [25,27-29]. Having opportunities for patients and living donors to share experiences and wisdom with each other is also an important tenet for excellence in patient-centered care [30]. Interventions that use storytelling as a means to produce behavioral change have been successful in increasing cancer screening rates [31], improving adherence to diabetes management behaviors [32], smoking cessation [33], reducing blood pressure [34], and losing weight [35]. Storytelling using digital apps also has the potential to reduce the burden of educating patients placed on providers [25]. As capturing stories using video software can be both expensive and complex, storytelling has only been used minimally, predominately in online communities and discussion forums [36,37] or health testimonials [38].

Developing ways to easily standardize and disseminate diverse living donor stories using digital technology could inspire more people to consider becoming living donors and reduce the kidney shortage nationally. This study aimed to describe the development of the Living Donation Storytelling Project, a web-based digital library of living donation narratives from multiple storyteller types (eg, recipient, donor). There were 4 aims: to describe (1) the theoretical foundation and development of a web-based digital library using video capture technology, (2) a recruitment protocol to capture diverse storytellers, (3) the characteristics and experiences of the participating storytellers, and (4) the frequency with which ethical concerns in the content shared emerged.

Methods

Theoretical Frameworks and Web-Based Library Development

The Narrative Theory supports the use of storytelling as an organic way in which humans naturally process and assign value to information, especially when it is presented by someone who resembles the listener [39-42]. The development of this web-based video library required much more than asking individuals to use their smartphones to film and upload stories. The formal phases of its development included selecting an audience of learners; using theory to determine the best delivery modality for that audience and the features it should have; drafting educational prompts to elicit high educational value; recruiting diverse storytellers to represent the entire transplant and living donation community; screening and editing stories to protect storytellers and eliminate sharing of misinformation; and building a web-based, searchable library of stories and marketing its availability to multiple communities (Multimedia Appendix 1).

On the basis of diffusion of innovation (DOI) theory [43], we chose a platform for recording and sharing stories that allowed

storytellers to select a set of topics that they wanted to share about, then introduced each topic with an open-ended prompt, one at a time, to help them share it easily and clearly. This format satisfied DOI constructs of compatibility, trialability, observability, and relative advantage. The resource was highly compatible—consistent with cultural values and practices—as most people, even those facing socioeconomic challenges, have access to a smartphone [44-46] and are familiar with YouTube-style videos [47]. Our resource was trialable—users could try it out without committing to sharing a final video. After each prompt, the storyteller was able to review what they filmed, re-record it if necessary, and approve that footage before continuing. Observability—the ability to see the product in practice—was addressed with a sample video modeling how digital technology was used and by inviting potential storytellers to view completed stories. This provided a great relative advantage—ease of use and improvement over existing options—compared with having to record and edit your own video as the digital software automatically spliced together responses, creating a seamless story. Only when the storyteller was comfortable with the entire video, was it released for inclusion into the library.

As recommended by the transtheoretical model of behavioral change [48-50], when creating prompts for patients to share about, we considered that stories, especially those with higher emotional valence, have been shown to connect better with audiences who are earlier in their readiness to learn or take action. Therefore, we included prompts to probe for storytellers' best moments after donation and recommendations for people facing similar situations to elicit sharing of more emotional content (eg, "When we learned that the surgery was a success, we all felt..." and "The best advice I could give someone else who is considering being a Living Donor is..."). We also included questions for learners who knew little about transplant and donation to invite them to learn more (eg, "I first considered donating a kidney..."). Storytellers were prompted to talk about the emotional and logistical challenges they faced during the LDKT process and how they overcame them. By doing so, they provided a road map for how the audience, if interested, could follow a similar path.

Social cognitive theory shows that people learn by observing others who look like them and by observing the consequences that others receive as a result of their actions [51,52]. Thus, we recruited real recipients, donors, and allies of many different backgrounds to boost identification with the storytellers. In addition, we built a filtering search function so that users of the library could search for storytellers who matched them in gender or race/ethnicity or were facing a similar situation (eg, needing a kidney, considering living donation; Multimedia Appendix 2). Finally, we developed prompts to enhance sharing in a way that allowed the listener to learn about the consequences of different decisions (eg, "Initially my attitude toward living kidney donation was..."; "Once I learned more about living donation, I considered becoming a living donor because..."; "I had questions and fears of my own, like... I was able resolve my concerns by...").

The resulting Living Donation Storytelling Project web-based digital library [53] includes 5 key features to break open learning and pursuit of living donation in new ways: (1) digital video capture technology to film and seam videos together remotely without the need for an editor; (2) guided prompts to help storytellers select and share about topics most important to them using a mobile phone, tablet, or laptop; (3) a search engine to allow for audiences to locate stories most like them by demographic and story type; (4) the ability to upload content both to the library and social media to help individuals find living donors; and (5) referrals to additional educational content about LDKT, including the location of the nearest transplant center. The library is categorized into donor stories, recipient stories, family/friend stories, and stories from people in need of a living kidney donor. Although the website was launched only after a diverse representation of storytellers could be shown, the library of stories continues to widen as more people film and upload stories. Finally, Google Analytics metrics were enabled to track the usage of the site and its features, as well as views of specific stories.

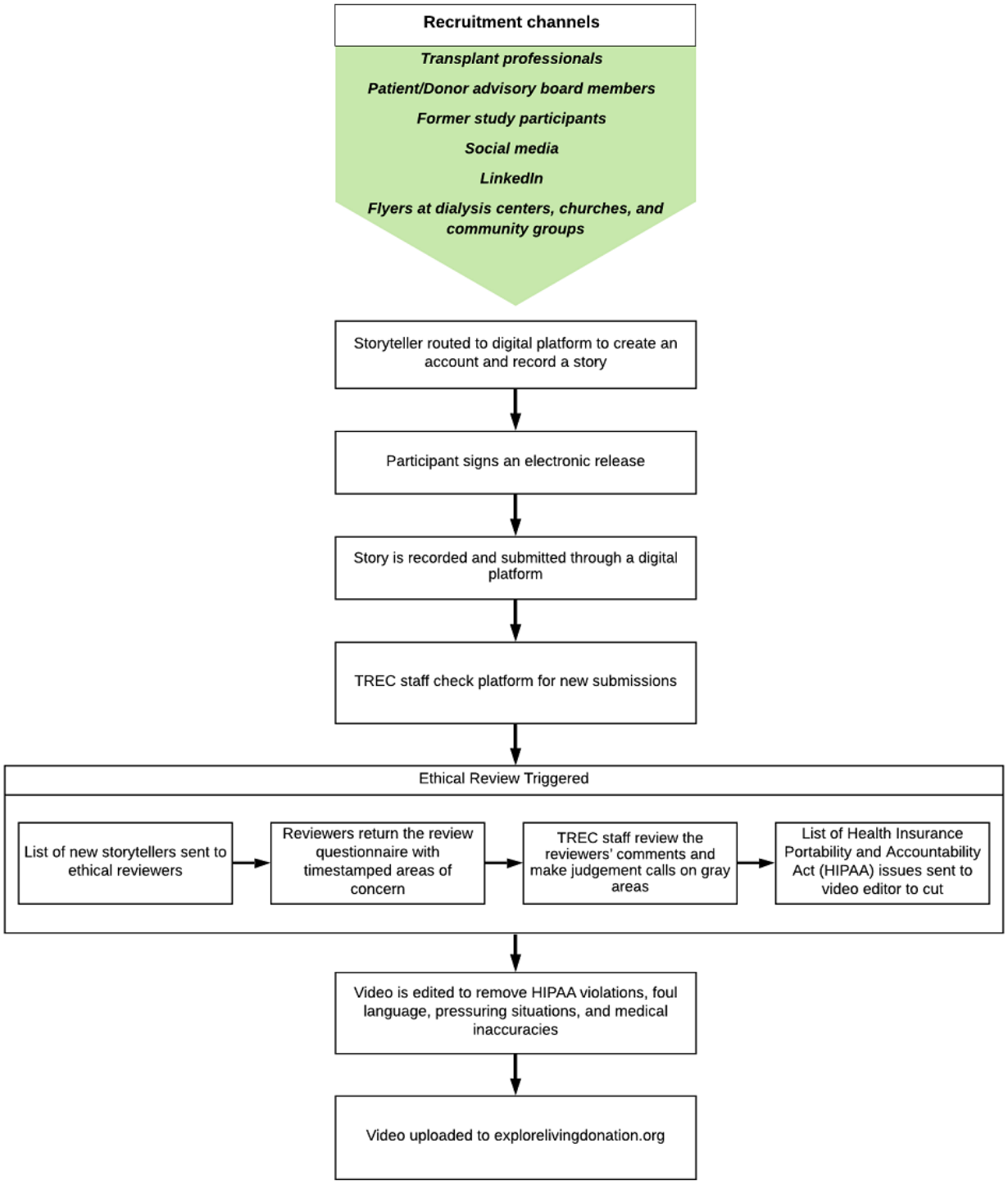
Storyteller Recruitment and Filming

Storytellers were individuals who had previously donated a kidney, recipients of LDKTs, family members or friends, or kidney patients seeking an LDKT. Storytellers were recruited via social media (Facebook, Instagram, Twitter, and LinkedIn), community outreach at support groups and churches, and referrals from kidney professionals and organizations (Figure 1). Once a storyteller was identified, they were given a personal link to privately film their story using a story guide matching their storyteller type.

Prompts were developed by our research team and grouped as story guides based on 6 transplant story types: recipient, donor, family/friend, exploring donation, in need of a kidney, and kidney ally. Depending on the guide, storytellers were offered 5 to 25 open-ended prompts that addressed their decision-making experiences, questions, needs, fears, and hopes, their donation and transplant experiences, how their lives changed after donation or transplant, and recommendations for potential donors.

Storytellers completed a standard media release for use of the stories, received reminders not to disclose any protected health information (PHI), and were supported with questions by a staff member (Multimedia Appendix 3). Storytellers could also use a tool on the website to find their nearest transplant center (Multimedia Appendix 4). Once recorded, the video files were stored on the private servers of the video capture platform. The project underwent the University of California, Los Angeles (UCLA), Institutional Review Board (IRB) review (IRB#18-000516), where exempt status was awarded.

Figure 1. Process flow of initial recruitment of storyteller to a completed story by the Transplant Research & Education Center.



Ethical Review of Story Content and Video Edits

After videos were submitted, transplant professionals (LH, SM, MA, and WB) watched each video and completed an ethical screening worksheet to check for PHI disclosures, about either the storyteller or anyone else, including their last name, addresses, transplant center name, social security numbers or medical record numbers, transplant date, or ESKD diagnosis date. They also screened for medical overgeneralizations or inaccuracies, pressuring language about donation or transplant, or foul language. Any problematic instances were coded by a timestamp and later removed by a video editor. The process flow of the storyteller's path from the initial recruitment to the completed story is outlined in Figure 1.

Storyteller Postsurvey

An opportunity to complete a 32-question postsurvey assessing the storyteller's experience filming their story was offered to those who had already completed and uploaded their stories. Those who completed a story received a voluntary postsurvey link in the email address that they provided. Surveys were collected using the Research Electronic Data Capture (RED

Cap) software, with each storyteller receiving a US $25 gift card after completing their survey. Data were stored on a secure, password-protected UCLA server.

The survey assessed storytellers' demographic characteristics (eg, gender, race/ethnicity, and age), level of education obtained, and type of story completed (eg, donor, recipient, family member, etc). Storytellers were asked what motivated them to share their stories (eg, to dispel myths, to help more people become living donors). On a scale from 1-very difficult to 7-very easy, storytellers were asked how easy or difficult filming and sharing their story was.

Results

Diverse Storyteller Recruitment: Characteristics and Motivation

In total, 412 potential storytellers received an initial introductory email with a story link that was unique to their experience. Of those invited to participate, 34.7% (143/412) storytellers completed stories and, of these, 95.8% (137/143) completed a voluntary postsurvey. Among those who completed a story, 72.2% (99/137) of the storytellers were female, 76.6% (105/137) were white, 60.5% (83/137) were living kidney donors, and 81.0% (111/137) had a college degree or higher. About 8.8% (12/137) of the storytellers were Hispanic and 23.4% (32/137) were nonwhite (Table 1).

Nearly all (130/137, 94.9%) storytellers were motivated by a desire to educate the public about living donation (Table 2).

Table 1. Storyteller characteristics (N=137).

Characteristic	Values
Age (years), mean (SD)	49.6 (12.4)
Gender, n (%)	
Male	38 (27.7)
Female	99 (72.3)
Race, n (%)	
White	105 (76.6)
Black	14 (10.3)
Other	18 (13.1)
Ethnicity, n (%)	
Hispanic	12 (8.7)
Story type, n (%)	
Living kidney donor	83 (60.6)
Kidney recipient	37 (27.0)
Family or friend of kidney recipient	7 (5.1)
Patient on waitlist	3 (2.2)
Family/friend of the patient on the waitlist	2 (1.5)
Other	5 (3.6)
Education, n (%)	
High school diploma or GED[a]	6 (4.4)
Some college or vocational school	19 (14.0)
College or vocational school degree	60 (44.1)
Some professional or graduate school	11 (8.1)
Professional or graduate school degree	40 (29.4)

[a]GED: General Education Development or General Education Diploma.

Table 2. Storyteller motivations, barriers, and disclosure of protected health information (N=137).

Responses	Value
Storyteller motivations, n (%)	
To educate the public about living donation	130 (94.9)
To spread awareness and help others	110 (80.3)
To make a difference in living donor recipient's and donor's lives	109 (79.6)
To help more people become living donors	107 (78.1)
To dispel myths about living donation	104 (75.9)
User experience, n (%)	
Storytellers who found filming a story difficult	34 (24.8)
Storytellers who found filming their story emotionally difficult	29 (21.2)
Disclosure of PHI[a], n (%)[b]	
PHI about themselves	78 (57.8)
Last name	55 (40.7)
Specific transplant center	24 (17.8)
Transplant or donation date, month, and year	25 (18.5)
Geographic specifics of location	11 (8.1)
PHI about another	56 (41.5)
Last name	30 (22.2)
Specific transplant center	15 (11.1)
Transplant or donation date, month, and year	13 (9.6)
Geographic specifics of location	4 (2.9)

[a]PHI: protected health information.

[b]Disclosure of PHI obtained from Ethical Review of Storytellers who completed the storyteller postsurvey (n=135).

Storytelling Experience and Content Shared

On average, storytellers took approximately one hour to review the prompts and prepare and record a story using open-ended prompts within the video capture technology (mean 62.5; SD: 87.8 mins; range 5 mins-12 hours). Completed stories had an average length of 10 min (SD: 6:12 min; range 0:46 seconds - 32 min).

A quarter (34/137, 24.8%) of the storytellers stated that they had difficulty filming their story using the technology and 21.2% (29/137) of the participants found sharing their stories to be emotionally difficult (Table 2). Some reported difficulty navigating the video capture technology, whereas others lacked access to a smartphone, laptop, tablet, or any other device with a camera to record their story. Emotional barriers to sharing their stories included fear of being on film, fear of talking openly about needing a kidney transplant, and vulnerability associated with sharing their donation or transplant experience in a public forum. Although not statistically significant, older storytellers (>50 years of age) had more emotional difficulty sharing their story (37/137, 27.0% vs 20/137, 14.6%; *P*=.12) and filming their story (40/137, 29.2% vs 29/137, 21.2%; *P*=.37) using the technology than the younger storytellers.

Most storytellers answered the majority of the prompts that were available in their story guide. Tables 3 and 4 provide examples of what was shared for the most common prompts selected within various story guides. Recipients shared about how they coped with uncertainties about donation or transplant outcomes, including discussions of their faith, the support they received from their care team, and sources of information that they used. Recipients talked about renewed freedom—the ability to eat and drink what they wanted, travel, swim, and spend time with family. Donors talked about watching their recipient return to improved health, what motivated them to consider LDKT, fears they had for a loved one, fears they had for themselves, and how they resolved these fears, mostly through learning more. Those in need of a kidney, and their supporters, talked about how deserving they were and ways to get in touch with them if interested in being a donor.

Storytellers shared vulnerably about their experiences, both laughing and crying within their stories. Emotionally, people expressed gratitude for the gift of life, worry about whether the kidney would work, and concern for the health of the donor within their stories.

Table 3. Content shared by donors, recipients, and family around open-ended prompts.

Prompts by story guide	Examples of storytelling content shared
Living donor	
The best advice I could give someone else who is thinking about being a living donor is...	"...just go ahead and ask questions. Talk to the transplant team. Let them decide, let the coordinators decide, do their screenings, do their questions. See if you are a likely candidate. You never know. I donated, maybe you can too." (Lisa H)
I ultimately decided to donate a kidney because	"Of course, there was nothing I wouldn't do to save my daughter's life. But, also, because the quicker she got off the [wait] list, the quicker someone else would get an opportunity." (Luther)
The best moment after my surgery was...	"About 4 days later, when I was visiting Lexi again. I saw how much happier and healthier she was [...] just seeing how good she was feeling really made me feel great." (Luther)
Kidney recipient	
My kidney failure began when I was (X) years old. At that time, I was doing (common activities for you before the transplant)...but then I started to notice (changes that affected your daily life)	"Kidney failure began when I was 25 years old, I was at stage 3. My kidney function was at about 40%. When I turned 31, and I became pregnant, that is when I become stage 5. My kidney function went from 40% functioning to about 8% functioning." (Kara)
Living without working kidneys meant that... The first time I had a dialysis treatment was... (explain how it felt)	"It meant that my time was limited. My disease started when I was 27, when I was 45, my kidney function has dropped to 9%. I felt defeated. A lot of the people that I talked to on my first day on dialysis had been coming there for 5 years." (Rochelle)
I found it... (difficult /easy) to talk about living donation with my family and friends, because...	"Found it easy to talk to about living donation with my friends and family. Because of living donation, I am alive today. So, I talk about it openly." (Holly)
Family member	
We learned as a family what kidney failure meant physically for (Recipient). However, for our family it also... (Explain the ways it changed your family's life)...	"…For our family it meant a huge lifestyle change. It was a huge financial burden for our family. It was mentally draining for my parents especially. I was 15 at the time, so, I didn't really grasp Hans's situation." (Drea)
Some of the best resources I used to learn about kidney disease were...	"Living through Hans's situation as a family member. I was 14-15 when his life on dialysis began. It was an in-person, real life, first-hand experience." (Drea)
(Recipient's) dialysis schedule meant that we had to change our current plans because...	"Changed our family's plans for traveling. We like to travel a lot. It changed drastically. We hated going and having him left out. So, we worked around it and there were days that he didn't feel so great. So, dialysis sucks. Period." (Annamarie)

Table 4. Content shared by potential donors and kidney patients around open-ended prompts.

Prompts by story guide	Examples of storytelling content shared
Exploring donation	
I first considered donating a kidney...	"When our family found out that my niece Marie was in kidney failure. She was fairly young, she was about 13-14. She was about to start junior high and she was about start dialysis. That is when I first started to think about being a donor." (Monica)
Initially my attitude toward living kidney donation was...	"I was little, I was a afraid at first. I didn't know what it entailed and until I read about it more. I have two kidneys and I have one to spare." (Kurt)
Once I learned more about living donation, I considered becoming a living donor because...	"Once I learned about kidney donation, I considered to be a donor because my wife had kidney failure." (Kurt)
In need of a kidney	
I could get a kidney from a living donor. So far...(tell what you have done or plan to do to find a donor)	"It is hard for me to find a kidney because; I need a B-, live kidney donor. I've done a lot of social outreach. I've made shirts, made some pamphlets, and gone on social media [to find a donor]." (Kabir)
I started having these symptoms... I was diagnosed with (explain prognosis), (X) months/years ago	"I was having high blood pressure and slowly I was feeling fatigue and I would always have this headache. I was not aware of what was going on with my health. I thought I was just too active. But, slowly the disease took over me and I had stage 4 kidney failure." (Kabir)
To stay alive, I have to go for dialysis, which is... (briefly describe what it is, how often you go, your new quality of life)	"For my dialysis, peritoneal dialysis every day. Most people do couple of hours every day but I do about 12 and a half hours every day [...]" (Kabir)

Ethical Concerns: Disclosure of Protected Health Information or Inaccurate Information

More than half (78/135, 57.8%) of the storytellers disclosed PHI about themselves, most often their last name, specific transplant center name, transplant or donation date, and geographic details about their locations. Some storytellers, 41.5% (56/135), also shared PHI about others involved in their living donation experience (Table 2).

A minority (18/135, 13.3%) of the storytellers shared medical inaccuracies or overgeneralizations. Examples included, "…going on dialysis means your life is over and your caregivers' life is over" or "…1 in 3 actually have kidney disease." Statements that could be considered instances of pressuring were shared by 11.9% (16/135) of the participants. These included statements like "People who do not give to a family member are selfish."

Discussion

Principal Findings

This study evaluated the feasibility of building a web-based digital library and recruiting storytellers to share their experiences. Using community-based participatory research practices and recruitment through social media, we successfully recruited storytellers involved with living donation who were predominately female, white, and motivated to assist others in making choices about donation and transplant. In general, it was more difficult to recruit minority storytellers and males to share their stories. As African Americans and Hispanics have 2.9 and 1.3 times higher ESKD incidence rates, respectively, compared with their white counterparts [1] and the lifetime risk of ESKD is higher in males than in females [1], additional research is needed to explore how to enroll these communities into sharing their stories and help them feel supported when sharing their stories.

Thousands of kidney patients awaiting transplant die each year before a matching kidney is found [1]. The Living Donation Storytelling Project is an important public resource enabling real-life stories about living kidney donation from multiple audiences to be captured using video capture technology and shared through a web-based searchable digital library.

At the start of this project, we were unclear how difficult storytelling would be and what types of content would be shared. More than half of those who were offered a link to record a story did not submit a completed video story, and about one-quarter of participants who completed a story reported either technical or emotional challenges when asked to reflect afterward on the recording process. However, the content shared, including poignant first-person recounting of fears, lessons learned, challenges overcome, and recommendations for others facing these decisions were very powerful. Future research should determine the key topics most commonly shared by storytellers and assess which storytellers and types of content most connect with different audiences. Further examination of the impact of storytelling combined with other, more traditional, educational strategies for increasing the number of living donors coming forward and LDKT rates is also needed.

Transplant professionals who served on our ethical review board concluded that, in general, storytellers act ethically; however, reminders should be sent to ensure privacy and prevent the disclosure of PHI. Clarification of what not to share in a public forum helped reduce disclosure of PHI, as did editing afterward. Additional work is still needed to explore ethical issues, including whether storytellers seeking living donors should be allowed to disclose their contact information in videos shared in the library.

Moving forward, there are many applications of the Living Donation Storytelling Library methodology. Stories can be easily incorporated within the traditional educational process during transplant or donation evaluation, provided as a general

introduction to transplant in dialysis centers, or embedded into educational portals linked to electronic patient medical records. Education delivered through these portals is quickly becoming a part of the standard of care [54-57]. As social media has become a common forum for prospective recipients and donors to seek donors, this library also provides interested patients with an easy way to share their interest in finding a living donor widely with their social media communities [22,58,59]. This library can also be used to amplify the voice of transplant champions and ambassadors to increase public awareness about the cause [60]. Sharing stories through social media may also be an effective way to advance a potential living donor's or transplant patient's stage of readiness for LDKT [61].

Finally, the Living Donation Storytelling Project may be particularly effective for reaching certain groups who are not well-served by existing educational strategies. Specifically, stories may also be a gentler way to introduce the option of living donation for patients who are concerned about harming a loved one, those who cannot read or have difficulty reading [62], those who speak languages other than English [25], and those with higher levels of medical mistrust [63-66]. Stories may also be well-suited for educating patient populations, including Native Americans or First Nations people who have cultures steeped in oral traditions [67-70]. Finally, the digital storytelling library methodology can be applied to health areas outside of living donation, as first-person storytelling has been shown to improve health outcomes for patients who belong to racial or ethnic minorities, have low health literacy, and are of lower socioeconomic status [71].

Conclusions

In summary, watching real-life stories can be reassuring, empowering, and, sometimes, inspiring [25,31,72,73]. With appropriate sensitivity to ensure diverse recruitment, ethical review of content, and support for storytellers while using innovative technology, digital storytelling technologies may be a cost-effective way to further engage patients and increase the curiosity of the public about becoming living donors.

Acknowledgments

The authors acknowledge StoryTap and Bernadette Butler for customizing a digital interface for capturing video stories; the Terasaki Family Foundation, the charitable grant that helped fund some of this work via grant funding; Sarah Brecher, BS, and Natalie Pita, BS, for recruiting and supporting storytellers and editing videos; the many individuals who shared their stories; and transplant centers and institutions who referred storytellers, including Mount Sinai Medical Center and NYKidney.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Steps for building a digital library.
[PDF File (Adobe PDF File), 119 KB - formative_v4i7e17441_app1.pdf]

Multimedia Appendix 2
Living Donation Storytelling Project website search engine.
[PDF File (Adobe PDF File), 95 KB - formative_v4i7e17441_app2.pdf]

Multimedia Appendix 3
Living Donation Storytelling Project storytelling instructions.
[PDF File (Adobe PDF File), 78 KB - formative_v4i7e17441_app3.pdf]

Multimedia Appendix 4
Living Donation Storytelling Project transplant center search tool.
[PDF File (Adobe PDF File), 85 KB - formative_v4i7e17441_app4.pdf]

References

1. Executive Summary: Overview of Kidney Disease in the United States. The United States Renal Data System (USRDS). 2018. URL: https://www.usrds.org/2018/view/Default.aspx [accessed 2020-06-14]
2. US Renal Data System 2019 Annual Data Report: Epidemiology of Kidney Disease in the United States. The United States Renal Data System (USRDS). 2019. URL: https://www.usrds.org/2019/view/USRDS_2019_ES_final.pdf [accessed 2020-06-14]
3. Hart A, Smith JM, Skeans MA, Gustafson SK, Wilk AR, Castro S, et al. OPTN/SRTR 2017 annual data report: kidney. Am J Transplant 2019 Feb;19(Suppl 2):19-123 [FREE Full text] [doi: 10.1111/ajt.15274] [Medline: 30811893]
4. Organ Datasource. Organ Procurement and Transplantation Network. 2019. URL: https://optn.transplant.hrsa.gov/data/organ-datasource/ [accessed 2020-06-15]

5. Purnell TS, Luo X, Cooper LA, Massie AB, Kucirka LM, Henderson ML, et al. Association of race and ethnicity with live donor kidney transplantation in the United States from 1995 to 2014. J Am Med Assoc 2018 Jan 2;319(1):49-61 [FREE Full text] [doi: 10.1001/jama.2017.19152] [Medline: 29297077]
6. Lentine KL, Mandelbrot D. Addressing disparities in living donor kidney transplantation: a call to action. Clin J Am Soc Nephrol 2018 Dec 7;13(12):1909-1911 [FREE Full text] [doi: 10.2215/CJN.06250518] [Medline: 30429153]
7. Kumar K, Holscher CM, Luo X, Wang JG, Anjum S, King EA, et al. Persistent regional and racial disparities in nondirected living kidney donation. Clin Transplant 2017 Dec;31(12):e13135 [FREE Full text] [doi: 10.1111/ctr.13135] [Medline: 29032601]
8. Purnell T, Xu P, Leca N, Hall Y. Racial differences in determinants of live donor kidney transplantation in the United States. Am J Transplant 2013 Jun;13(6):1557-1565 [FREE Full text] [doi: 10.1111/ajt.12258] [Medline: 23669021]
9. Waterman AD, Peipert JD. An explore transplant group randomized controlled education trial to increase dialysis patients' decision-making and pursuit of transplantation. Prog Transplant 2018 Jun;28(2):174-183. [doi: 10.1177/1526924818765815] [Medline: 29699451]
10. Boulware LE, Hill-Briggs F, Kraus ES, Melancon JK, Falcone B, Ephraim PL, et al. Effectiveness of educational and social worker interventions to activate patients' discussion and pursuit of preemptive living donor kidney transplantation: a randomized controlled trial. Am J Kidney Dis 2013 Mar;61(3):476-486 [FREE Full text] [doi: 10.1053/j.ajkd.2012.08.039] [Medline: 23089512]
11. Garonzik-Wang JM, Berger JC, Ros RL, Kucirka LM, Deshpande NA, Boyarsky BJ, et al. Live donor champion: finding live kidney donors by separating the advocate from the patient. Transplantation 2012 Jun 15;93(11):1147-1150 [FREE Full text] [doi: 10.1097/TP.0b013e31824e75a5] [Medline: 22461037]
12. Brown W, McDermott J, Figueiredo AE, Loucaidou M, Galliford J, Papalois V. Invited manuscript poster on renal-related education American society of nephrology, Nov 16-21, 2010. Improving access to kidney transplant information has increased preemptive living kidney donation. Ren Fail 2011;33(7):658-662. [doi: 10.3109/0886022X.2011.589948] [Medline: 21787154]
13. Cardinal H, Durand C, Larrivée S, Verhave J, Pâquet MR, Fortin M. Strategies to increase living kidney donation: a retrospective cohort study. Can J Kidney Health Dis 2015;2:15 [FREE Full text] [doi: 10.1186/s40697-015-0049-6] [Medline: 25926997]
14. Patzer RE, Paul S, Plantinga L, Gander J, Sauls L, Krisher J, Southeastern Kidney Transplant Coalition. A randomized trial to reduce disparities in referral for transplant evaluation. J Am Soc Nephrol 2017 Mar;28(3):935-942 [FREE Full text] [doi: 10.1681/ASN.2016030320] [Medline: 27738125]
15. Ismail S, Luchtenburg A, Timman R, Zuidema WC, Boonstra C, Weimar W, et al. Home-based family intervention increases knowledge, communication and living donation rates: a randomized controlled trial. Am J Transplant 2014 Aug;14(8):1862-1869 [FREE Full text] [doi: 10.1111/ajt.12751] [Medline: 24935081]
16. Massey EK, Gregoor PJ, Nette RW, van den Dorpel MA, van Kooij A, Zietse R, et al. Early home-based group education to support informed decision-making among patients with end-stage renal disease: a multi-centre randomized controlled trial. Nephrol Dial Transplant 2016 May;31(5):823-830. [doi: 10.1093/ndt/gfv322] [Medline: 26330561]
17. Rodrigue JR, Paek MJ, Egbuna O, Waterman AD, Schold JD, Pavlakis M, et al. Making house calls increases living donor inquiries and evaluations for blacks on the kidney transplant waiting list. Transplantation 2014 Nov 15;98(9):979-986 [FREE Full text] [doi: 10.1097/TP.0000000000000165] [Medline: 24825528]
18. Axelrod DA, Kynard-Amerson CS, Wojciechowski D, Jacobs M, Lentine KL, Schnitzler M, et al. Cultural competency of a mobile, customized patient education tool for improving potential kidney transplant recipients' knowledge and decision-making. Clin Transplant 2017 May;31(5):e12944. [doi: 10.1111/ctr.12944] [Medline: 28263408]
19. Gordon EJ, Mullee JO, Ramirez DI, MacLean J, Olivero M, Feinglass J, et al. Hispanic/Latino concerns about living kidney donation: a focus group study. Prog Transplant 2014 Jun;24(2):152-162. [doi: 10.7182/pit2014946] [Medline: 24919732]
20. Patzer RE, McPherson L, Basu M, Mohan S, Wolf M, Chiles M, et al. Effect of the iChoose kidney decision aid in improving knowledge about treatment options among transplant candidates: a randomized controlled trial. Am J Transplant 2018 Aug;18(8):1954-1965 [FREE Full text] [doi: 10.1111/ajt.14693] [Medline: 29446209]
21. Burnapp L, van Assche K, Lennerling A, Slaats D, van Dellen D, Mamode N, et al. Raising awareness of unspecified living kidney donation: an ELPAT view. Clin Kidney J 2020 Apr;13(2):159-165 [FREE Full text] [doi: 10.1093/ckj/sfz067] [Medline: 32296519]
22. Kazley AS, Hamidi B, Balliet W, Baliga P. Social media use among living kidney donors and recipients: survey on current practice and potential. J Med Internet Res 2016 Dec 20;18(12):e328 [FREE Full text] [doi: 10.2196/jmir.6176] [Medline: 27998880]
23. Vedadi A, Bansal A, Yung P, Famure O, Mitchell M, Waterman AD, et al. Ethnic background is associated with no live kidney donor identified at the time of first transplant assessment-an opportunity missed? A single-center retrospective cohort study. Transpl Int 2019 Oct;32(10):1030-1043. [doi: 10.1111/tri.13476] [Medline: 31250484]
24. Sieverdes JC, Nemeth LS, Magwood GS, Baliga PK, Chavin KD, Brunner-Jackson B, et al. Patient-centered mhealth living donor transplant education program for African Americans: development and analysis. JMIR Res Protoc 2015 Aug 10;4(3):e84 [FREE Full text] [doi: 10.2196/resprot.3715] [Medline: 26265532]

25. Robinson JK, Friedewald JJ, Desai A, Gordon EJ. Response across the health-literacy spectrum of kidney transplant recipients to a sun-protection education program delivered on tablet computers: randomized controlled trial. JMIR Cancer 2015 Aug 18;1(2):e8 [FREE Full text] [doi: 10.2196/cancer.4787] [Medline: 28410176]
26. Chronic Kidney Disease (CKD) in the United States. The United States Renal Data System (USRDS). 2016. URL: https://www.usrds.org/2016/view/Default.aspx [accessed 2020-06-15]
27. Schreiner C, Appel M, Isberner M, Richter T. Argument strength and the persuasiveness of stories. Discourse Process 2018;55(4):371-386 [FREE Full text] [doi: 10.1080/0163853X.2016.1257406] [Medline: 29805322]
28. Suzuki WA, Feliú-Mójer MI, Hasson U, Yehuda R, Zarate JM. Dialogues: the science and power of storytelling. J Neurosci 2018 Oct 31;38(44):9468-9470. [doi: 10.1523/jneurosci.1942-18.2018]
29. Briant KJ, Halter A, Marchello N, Escareño M, Thompson B. The power of digital storytelling as a culturally relevant health promotion tool. Health Promot Pract 2016 Nov;17(6):793-801 [FREE Full text] [doi: 10.1177/1524839916658023] [Medline: 27402721]
30. Constand MK, MacDermid JC, dal Bello-Haas V, Law M. Scoping review of patient-centered care approaches in healthcare. BMC Health Serv Res 2014 Jun 19;14:271 [FREE Full text] [doi: 10.1186/1472-6963-14-271] [Medline: 24947822]
31. McQueen A, Kreuter MW, Kalesan B, Alcaraz KI. Understanding narrative effects: the impact of breast cancer survivor stories on message processing, attitudes, and beliefs among African American women. Health Psychol 2011 Nov;30(6):674-682 [FREE Full text] [doi: 10.1037/a0025395] [Medline: 21895370]
32. Peek ME, Harmon SA, Scott SJ, Eder M, Roberson TS, Tang H, et al. Culturally tailoring patient education and communication skills training to empower African-Americans with diabetes. Transl Behav Med 2012 Sep;2(3):296-308 [FREE Full text] [doi: 10.1007/s13142-012-0125-8] [Medline: 24073128]
33. Andrews JO, Felton G, Wewers ME, Waller J, Tingen M. The effect of a multi-component smoking cessation intervention in African American women residing in public housing. Res Nurs Health 2007 Feb;30(1):45-60. [doi: 10.1002/nur.20174] [Medline: 17243107]
34. Nguyen HL, Ha DA, Goldberg RJ, Kiefe CI, Chiriboga G, Ly HN, et al. Culturally adaptive storytelling intervention versus didactic intervention to improve hypertension control in Vietnam- 12 month follow up results: a cluster randomized controlled feasibility trial. PLoS One 2018;13(12):e0209912 [FREE Full text] [doi: 10.1371/journal.pone.0209912] [Medline: 30596749]
35. Manuvinakurike R, Velicer WF, Bickmore TW. Automated indexing of Internet stories for health behavior change: weight loss attitude pilot study. J Med Internet Res 2014 Dec 9;16(12):e285 [FREE Full text] [doi: 10.2196/jmir.3702] [Medline: 25491243]
36. Taylor LA, Bahreman N, Hayat MJ, Hoey F, Rajasekaran G, Segev DL. Living kidney donors and their family caregivers: developing an evidence-based educational and social support website. Prog Transplant 2012 Jun;22(2):119-128 [FREE Full text] [doi: 10.7182/pit2012611] [Medline: 22878067]
37. Wicks P, Sulham KA, Gnanasakthy A. Quality of life in organ transplant recipients participating in an online transplant community. Patient 2014;7(1):73-84 [FREE Full text] [doi: 10.1007/s40271-013-0033-0] [Medline: 24194474]
38. Gordon EJ, Feinglass J, Carney P, Ramirez D, Olivero M, O'Connor K, et al. An interactive, bilingual, culturally targeted website about living kidney donation and transplantation for hispanics: development and formative evaluation. JMIR Res Protoc 2015 Apr 20;4(2):e42 [FREE Full text] [doi: 10.2196/resprot.3838] [Medline: 25896143]
39. Petty R. The importance of cognitive responses in persuasion. Adv Consum Res 1977;4:357-362. [doi: 10.4324/9781315803012]
40. Busselle R, Bilandzic H. Measuring narrative engagement. Media Psychol 2009 Nov 23;12(4):321-347. [doi: 10.1080/15213260903287259]
41. Lee H, Fawcett J, DeMarco R. Storytelling/narrative theory to address health communication with minority populations. Appl Nurs Res 2016 May;30:58-60. [doi: 10.1016/j.apnr.2015.09.004] [Medline: 27091254]
42. Green MC, Brock TC. The role of transportation in the persuasiveness of public narratives. J Pers Soc Psychol 2000 Nov;79(5):701-721. [doi: 10.1037//0022-3514.79.5.701] [Medline: 11079236]
43. Rogers EM. Diffusion of Innovations. New York, USA: Free Press; 1983.
44. Anderson M. Mobile Technology and Home Broadband 2019. Pew Research Center. 2019. URL: https://www.pewresearch.org/internet/2019/06/13/mobile-technology-and-home-broadband-2019/ [accessed 2020-06-15]
45. Mobile Fact Sheet. Pew Research Center. 2019. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2020-06-15]
46. Fox S. Mobile Health. Pew Center for Research. 2011. URL: https://www.pewresearch.org/internet/2011/04/14/mobile-health/ [accessed 2020-06-15]
47. van Kessel P. 10 Facts About Americans and YouTube. Pew Research Center. 2019. URL: https://www.pewresearch.org/fact-tank/2019/12/04/10-facts-about-americans-and-youtube/ [accessed 2020-06-15]
48. Prochaska JO, Velicer WF, Rossi JS, Goldstein MG, Marcus BH, Rakowski W, et al. Stages of change and decisional balance for 12 problem behaviors. Health Psychol 1994 Jan;13(1):39-46. [doi: 10.1037//0278-6133.13.1.39] [Medline: 8168470]

49. Waterman AD, Robbins ML, Paiva AL, Hyland SS. Kidney patients' intention to receive a deceased donor transplant: development of stage of change, decisional balance and self-efficacy measures. J Health Psychol 2010 Apr;15(3):436-445 [FREE Full text] [doi: 10.1177/1359105309351248] [Medline: 20348364]
50. Waterman AD, Robbins ML, Paiva AL, Peipert JD, Davis LA, Hyland SS, et al. Measuring kidney patients' motivation to pursue living donor kidney transplant: development of stage of change, decisional balance and self-efficacy measures. J Health Psychol 2015 Feb;20(2):210-221 [FREE Full text] [doi: 10.1177/1359105313501707] [Medline: 24155194]
51. Bandura A. Social Learning Theory. New York, USA: Prentice-Hall; 1976.
52. Kazdin AE. Vicarious reinforcement and punishment processes in the classroom. In: Strain PS, editor. The Utilization of Classroom Peers as Behavior Change Agents. New York, USA: Academic Press; 1971.
53. Waterman A. Explore Living Donation Storytelling Project. 2019. URL: https://explorelivingdonation.org/ [accessed 2020-06-15]
54. Kim S, Abner E. Predictors affecting personal health information management skills. Inform Health Soc Care 2016;41(3):211-229 [FREE Full text] [doi: 10.3109/17538157.2015.1008490] [Medline: 26268728]
55. Amante DJ, Hogan TP, Pagoto SL, English TM. A systematic review of electronic portal usage among patients with diabetes. Diabetes Technol Ther 2014 Nov;16(11):784-793. [doi: 10.1089/dia.2014.0078] [Medline: 24999599]
56. Osborn CY, Mayberry LS, Wallston KA, Johnson KB, Elasy TA. Understanding patient portal use: implications for medication management. J Med Internet Res 2013 Jul 3;15(7):e133 [FREE Full text] [doi: 10.2196/jmir.2589] [Medline: 23823974]
57. Ramirez M, Wu S, Jin H, Ell K, Gross-Schulman S, Sklaroff LM, et al. Automated remote monitoring of depression: acceptance among low-income patients in diabetes disease management. JMIR Ment Health 2016 Jan 25;3(1):e6 [FREE Full text] [doi: 10.2196/mental.4823] [Medline: 26810139]
58. Chang A, Anderson EE, Turner HT, Shoham D, Hou SH, Grams M. Identifying potential kidney donors using social networking web sites. Clin Transplant 2013;27(3):E320-E326 [FREE Full text] [doi: 10.1111/ctr.12122] [Medline: 23600791]
59. Kumar K, King EA, Muzaale AD, Konel JM, Bramstedt KA, Massie AB, et al. A smartphone app for increasing live organ donation. Am J Transplant 2016 Dec;16(12):3548-3553 [FREE Full text] [doi: 10.1111/ajt.13961] [Medline: 27402293]
60. Shi J, Salmon CT. Identifying opinion leaders to promote organ donation on social media: network study. J Med Internet Res 2018 Jan 9;20(1):e7 [FREE Full text] [doi: 10.2196/jmir.7643] [Medline: 29317384]
61. Ramo DE, Thrul J, Chavez K, Delucchi KL, Prochaska JJ. Feasibility and quit rates of the tobacco status project: a Facebook smoking cessation intervention for young adults. J Med Internet Res 2015 Dec 31;17(12):e291 [FREE Full text] [doi: 10.2196/jmir.5209] [Medline: 26721211]
62. Taylor DM, Bradley JA, Bradley C, Draper H, Dudley C, Fogarty D, ATTOM investigators. Limited health literacy is associated with reduced access to kidney transplantation. Kidney Int 2019 May;95(5):1244-1252. [doi: 10.1016/j.kint.2018.12.021] [Medline: 30952457]
63. Wachterman M, McCarthy E, Marcantonio E, Ersek M. Mistrust, misperceptions, and miscommunication: a qualitative study of preferences about kidney transplantation among African Americans. Transplant Proc 2015 Mar;47(2):240-246 [FREE Full text] [doi: 10.1016/j.transproceed.2015.01.016] [Medline: 25769556]
64. Myaskovsky L, Switzer GE, Crowley-Matoka M, Unruh M, DiMartini AF, Dew MA. Psychosocial factors associated with ethnic differences in transplantation. Curr Opin Organ Transplant 2007;12(2):182-187. [doi: 10.1097/mot.0b013e32805b7192]
65. Waterman AD, Barrett AC, Stanley SL, Rothstein M, Burroughs TE, Brennan DC, et al. Why African-Americans are not pursuing living kidney donation. Transplantation 2006;82(Suppl 2):286. [doi: 10.1097/00007890-200607152-00649]
66. Waterman AD, Beaumont JL. What else can we do to ensure transplant equity for high-risk patients? Clin J Am Soc Nephrol 2018 Apr 6;13(4):529-530 [FREE Full text] [doi: 10.2215/CJN.02120218] [Medline: 29581106]
67. Momper SL, Dennis MK, Mueller-Williams AC. American Indian elders share personal stories of alcohol use with younger tribal members. J Ethn Subst Abuse 2017;16(3):293-313 [FREE Full text] [doi: 10.1080/15332640.2016.1196633] [Medline: 27403635]
68. Montgomery M, Manuelito B, Nass C, Chock T, Buchwald D. The native comic book project: native youth making comics and healthy decisions. J Cancer Educ 2012 Apr;27(1 Suppl):S41-S46 [FREE Full text] [doi: 10.1007/s13187-012-0311-x] [Medline: 22259070]
69. Limaye NP, Rivas-Nieto AC, Carcamo CP, Blas MM. Nuestras Historias- designing a novel digital story intervention through participatory methods to improve maternal and child health in the Peruvian Amazon. PLoS One 2018;13(11):e0205673 [FREE Full text] [doi: 10.1371/journal.pone.0205673] [Medline: 30396200]
70. Vallesi S, Wood L, Dimer L, Zada M. 'In their own voice'-incorporating underlying social determinants into aboriginal health promotion programs. Int J Environ Res Public Health 2018 Jul 18;15(7):1514 [FREE Full text] [doi: 10.3390/ijerph15071514] [Medline: 30021953]
71. Lipsey AF, Waterman AD, Wood EH, Balliet W. Evaluation of first-person storytelling on changing health-related attitudes, knowledge, behaviors, and outcomes: a scoping review. Patient Educ Couns 2020 Apr 20:- epub ahead of print. [doi: 10.1016/j.pec.2020.04.014] [Medline: 32359877]

72. Patten CA, Windsor RA, Renner CC, Enoch C, Hochreiter A, Nevak C, et al. Feasibility of a tobacco cessation intervention for pregnant Alaska native women. Nicotine Tob Res 2010 Feb;12(2):79-87 [FREE Full text] [doi: 10.1093/ntr/ntp180] [Medline: 20018946]
73. Sheppard VB, Williams KP, Harrison TM, Jennings Y, Lucas W, Stephen J, et al. Development of decision-support intervention for Black women with breast cancer. Psychooncology 2010 Jan;19(1):62-70 [FREE Full text] [doi: 10.1002/pon.1530] [Medline: 19267384]

Abbreviations

DOI: diffusion of innovation
ESKD: end-stage kidney disease
IRB: Institutional Review Board
LDKT: living donor kidney transplant
PHI: protected health information
UCLA: University of California, Los Angeles

Edited by G Eysenbach; submitted 12.12.19; peer-reviewed by CY Lin, H Barahimi; comments to author 16.03.20; revised version received 08.05.20; accepted 14.05.20; published 21.07.20.

Please cite as:
Waterman AD, Wood EH, Ranasinghe ON, Faye Lipsey A, Anderson C, Balliet W, Holland-Carter L, Maurer S, Aurora Posadas Salas M
A Digital Library for Increasing Awareness About Living Donor Kidney Transplants: Formative Study
JMIR Form Res 2020;4(7):e17441
URL: https://formative.jmir.org/2020/7/e17441
doi:10.2196/17441
PMID:32480362

Original Paper

A Novel Educational Control Group Mobile App for Meditation Interventions: Single-Group Feasibility Trial

Jennifer Huberty[1], PhD; Ryan Eckert[2], MSc; Megan Puzia[3], MSc; Breanne Laird[1], BSc; Linda Larkey[4], PhD; Ruben Mesa[2], MD

[1]College of Health Solutions, Arizona State University, Phoenix, AZ, United States

[2]Mays Cancer Center, University of Texas Health San Antonio MD Anderson, San Antonio, TX, United States

[3]Behavioral Research and Analytics, LLC, Salt Lake City, UT, United States

[4]Edson College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

Corresponding Author:
Jennifer Huberty, PhD
College of Health Solutions
Arizona State University
500 North 3rd Street
Phoenix, AZ, 85004
United States
Phone: 1 402 301 1304
Email: jhuberty@asu.edu

Abstract

Background: Smartphone ownership is becoming ubiquitous among US adults, making the delivery of health interventions via a mobile app (ie, mobile health [mHealth]) attractive to many researchers and clinicians. Meditation interventions have become popular and have been delivered to study participants via mobile apps to improve a range of health outcomes in both healthy adults and those with chronic diseases. However, these meditation mHealth interventions have been limited by a lack of high-quality control groups. More specifically, these studies have lacked consistency in their use of active, time-matched, and attention-matched control groups.

Objective: The purpose of this study is to beta test a novel health education podcast control condition delivered via a smartphone app that would be a strong comparator to be used in future studies of app-based meditation interventions.

Methods: Patients with myeloproliferative neoplasm (MPN) cancer were recruited nationally. Upon enrollment, participants were informed to download the investigator-developed health education podcast app onto their mobile phone and listen to ~60 min/week of cancer-related educational podcasts for 12 weeks. The benchmarks for feasibility included ≥70% of participants completing ≥70% of the prescribed 60 min/week of podcasts, ≥70% of participants reporting that they were satisfied with the intervention, and ≥70% of participants reporting that they enjoyed the health education podcasts.

Results: A total of 96 patients with MPN were enrolled in the study; however, 19 never began the intervention. Of the 77 patients who participated in the intervention, 39 completed the entire study (ie, sustained participation through the follow-up period). Participation averaged 103.2 (SD 29.5) min/week. For 83.3% (10/12) of the weeks, at least 70% of participants completed at least 70% of their total prescribed use. Almost half of participants reported that they enjoyed the health education podcasts (19/39, 48.7%) and were satisfied with the intervention (17/39, 43.6%). There were no significant changes in cancer-related outcomes from baseline to postintervention.

Conclusions: A 12-week, health education podcast mobile app was demanded but not accepted in a sample of patients with cancer. Using the mobile app was not associated with significant changes in cancer-related symptoms. Based on findings from this study, a health education podcast mobile app may be a feasible option as a time- and attention-matched control group for efficacy trials with more extensive formative research for the content of the podcasts and its acceptability by the specific population.

Trial Registration: ClinicalTrials.gov NCT03907774; https://clinicaltrials.gov/ct2/show/NCT03907774

(JMIR Form Res 2020;4(7):e19364) doi:10.2196/19364

KEYWORDS

feasibility; smartphone; mHealth; digital health; cancer; beta test

Introduction

Smartphone ownership is becoming ubiquitous among adults in the United States (81% in 2019) [1]. Using mobile devices to support health and wellness (ie, mobile health [mHealth]) [2] may be a promising approach to help individuals prevent or manage chronic conditions and improve health outcomes [3]. However, many studies of mHealth interventions, particularly mobile app interventions, have been substantially limited by the lack of high-quality comparators (ie, control conditions). The most common types of control conditions include usual care (ie, usual care for the critical condition), wait-list control (ie, usual care and will later receive the intervention), and active control (ie, control group receives an activity or intervention that controls for some aspect of attention, time, or expectation) [4]. Active control groups can be more effective than wait-list control groups [5]. However, there is a lack of research on active control groups in app-based interventions [6,7]. Recent reviews and meta-analyses have called for improvements in the design of control groups within randomized control trials that evaluate the efficacy of mobile app interventions [4,8], and the National Institutes of Health (NIH) has recommended that careful selection of a comparator be designed to reflect the *primary purpose* of the study [9]. It is clear that there is a need for active *time- and attention-matched* comparators. That is, control conditions must aim to not only match the mode of activity or delivery of the intervention but also match the time and attention that is spent on the intervention. Without active control conditions that match interventions with regard to time and attention, studies of mHealth interventions are unavoidably confounded by differences in participant engagement [9]. There is a need to design and explore the feasibility of active time- and attention-matched control groups for mHealth studies.

Meditation apps have become quite popular in recent years [4,10] and have been used to improve mental and physical health in a range of healthy [7,11-15] and health-compromised populations [16,17]. However, few mobile app meditation studies have used active, time-matched, and attention-matched comparators [7,13,14,18] even when the study primarily aimed to determine effectiveness or efficacy. The most common comparators for these studies have been wait-list control groups, usual care, and educational handouts [12,15]. For example, in our work using a meditation app to reduce symptom burden in patients with hematological cancer, we used an educational handout with information about managing fatigue in cancer as our control group. Participation in the control group was not associated with improvements in health or cancer-related symptoms, suggesting that the cancer-related educational content may be reasonable for a control condition; however, this group did not match the engagement level of our intervention participants. This type of comparator could be improved by modifying it to mirror the basic functionality, look, and feel of the intervention group's meditation app (ie, active); match the time that the intervention group spends participating in meditation (ie, time-matched); and match the attention and basic mode of delivery that the intervention group requires to meditate (ie, attention-matched).

We sought to develop and beta-test an appropriate comparator app for interventions using a mobile meditation app. We developed the app with the ability for content to be modified, added, and updated, and be used across various populations participating in mobile meditation interventions. To further our progressive line of research, we chose to conduct the beta test in patients with hematological cancer (specifically, myeloproliferative neoplasm [MPN]) due to our ongoing work with this population and our partnerships with foundations in which to recruit patients with cancer for our beta test. Therefore, the purpose of this study is to beta test a novel health education podcast control condition delivered via a mobile app that would be a strong comparator to be used in future studies of app-based meditation interventions. We hypothesized that implementing the health education podcasts in a sample of patients with MPN would be feasible (ie, demanded, accepted) and that using the podcasts would not be associated with significant improvements in health, cancer-related symptoms (ie, depression, anxiety, pain intensity, and sleep disturbance), or total symptom burden. Our benchmarks for success were ≥70% of participants completing ≥70% of the prescribed 60 min/week of podcasts, ≥70% satisfied with the intervention, and ≥70% enjoying the health education podcasts.

Methods

This study was approved by the Institutional Review Board at Arizona State University.

Recruitment and Enrollment

Guided by Bowen and colleagues' [19] recommendations for designing feasibility studies, we aimed to enroll 100 participants in the study [19]. Because these early trials are used, in part, to estimate effect size and power for future trials, we did not expect to be fully powered to detect changes in primary study outcomes. Participants were recruited nationally via internet-based strategies, including social media (ie, Facebook, Twitter), social networking sites, and online and email listservs. All recruitment methods were approved by the Arizona State University Institutional Review Board. The study was advertised as a mobile app health education intervention. We recruited self-reported patients with hematological cancer (ie, patients with MPN) because we have a progressive line of work involving patients with MPN using mindfulness approaches to reduce symptom burden (eg, yoga, meditation). However, the app was developed to be applied in app-based meditation interventions across various populations. Interested participants completed an eligibility link on REDCap (Vanderbilt University). Inclusion criteria were as follows: (1) had a diagnosis of MPN (ie, polycythemia vera, myelofibrosis, essential thrombocythemia) identified by a treating physician, (2) had access to a smartphone on a regular basis, (3) had access to reliable home internet, (4) could read and understand English, and (5) were 18 years or older. The exclusion criteria were as follows: (1) planned change in pharmacologic intervention (ie, new drug, bone marrow transplant) during the study interval (ie, 12 weeks) and (2) resided outside of the United States. If eligible, the patients were emailed a link to a video explaining the informed consent and study procedures. If interested, patients

were asked to respond to the email indicating that they reviewed the video and had the opportunity to ask questions. Patients then completed an electronic informed consent delivered via REDCap prior to participation. If ineligible, patients were sent an email notification thanking them for their interest in the study and to respond if they were interested in being notified about future studies.

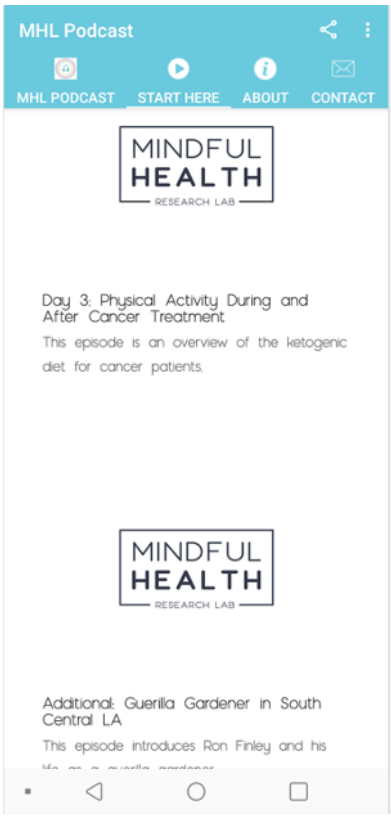
Research Design

The study was a single-group, *beta test study* of a new control intervention to be used as a comparator in future studies involving mindfulness and meditation mobile apps. A total 96 participants were enrolled in the study and assigned to a health education podcast group.

Podcast Control App Development

The concept for the podcast control app was developed by a PhD-level mHealth researcher with expertise in the development of national, digitally delivered interventions to improve physical and mental health. The goal was to develop an app that could deliver education information in the same context that a consumer-based mindfulness meditation app delivers content (ie, log onto an app, click on the content, listen). See Figure 1 for a screenshot of the Mindful Health Lab (MHL) podcast control app. The podcast control app was developed to match time (ie, 60-70 min/week) and attention (ie, same context of delivery and same functionality) of some mobile app meditation interventions [6,7,15]. We used the mobile app Calm as our model because of our long-standing partnership and research being conducted with the app [15,16]. The app was also developed so that content could be changed and tailored to any population in future studies. For example, if the study team was conducting a study in college students with the Calm app, the podcast health education content could be modified to be specific to college students. In the case of this study, we used patients with cancer. Thus, content was tailored to health education for patients with cancer. The app was designed to have the same general features as the Calm app (eg, reminders to listen to podcasts, ability to share use on social media, ability to track time spent listening to podcasts) but without the same branding. We did not include similar branding as Calm to keep control participants blinded to the app that would be used in the intervention group.

Figure 1. Mindful Health Lab (MHL) podcast app screenshot.



To gather content for the app, the research team searched publicly available podcasts related to health education through credible government and higher education websites and podcasts that could be used with our target population for our beta test (ie, patients with cancer). The selected content was then uploaded to an app created by a developer. The podcast content was only used for noncommercial purposes with credit to the source of the content included in the app. The audio was not changed or altered in any way and only those enrolled in the study had access to the content. The type of specific educational content included a variety of topic areas including nutrition, physical activity, time, and stress management, as well as general wellness and life-related topics (eg, cultivating happiness, organization practices).

Podcast Control App Prescription

Participants were asked to listen to the health education podcasts on their smartphone for approximately 60 min/week for 12 weeks. Podcasts were arranged by week and by day. There were approximately two to three podcasts prescribed per week,

averaging about 22 minutes per podcast. Additionally, there were one to three podcasts offered per week for participants to complete if they wanted to listen to more than 60 min/week. Although navigation through the podcast prescription was suggested by week, participants were not restricted from skipping around within the weeks. Time spent listening to the podcasts was collected by the app and downloaded by the research team.

Outcomes

The a priori benchmarks for feasibility were based on Bowen and colleagues [20] feasibility criteria and included ≥70% of participants completing ≥70% of the prescribed 60 min/week (42 min/week) of podcasts (ie, demand), ≥70% of participants reporting that they were satisfied with the intervention, and ≥70% of participants reporting that they enjoyed the health education podcasts (ie, acceptability). These specific benchmarks have been used successfully in other recent feasibility studies [16,21].

Questionnaires were administered at baseline (week 0), midintervention (week 6), and postintervention (week 12). These questionnaires included demographics (baseline only), satisfaction-related questions developed by the researchers (week 12 only), NIH Patient-Reported Outcomes Measurement Information System (PROMIS) outcomes (global health, pain intensity, anxiety, depression, and sleep disturbance), and the MPN Symptom Assessment Form Total Symptom Score (MPN-SAF TSS). The satisfaction questionnaire asked questions related to enjoyment, satisfaction, recommendation to others, etc (see Textbox 1 for satisfaction survey questions and responses). Answers were either a yes or no format, or a 5-point Likert scale. The NIH PROMIS is a valid and reliable tool for the measurement of symptoms among patients with cancer [22-25]. The MPN-SAF TSS is a valid and reliable way of assessing total symptom burden among patients with MPN [5]. All participants were provided with a US $25 digital gift card for completion of all questionnaires.

Textbox 1. Postintervention satisfaction survey questions.

1. On a scale of 1 to 5 (1=did not enjoy at all, 5=very much enjoyed), how would you rate your overall enjoyment of listening to the podcasts?
2. On a scale of 1 to 5 (1=not at all satisfied, 5=very much satisfied), how would you rate your overall satisfaction with the podcasts?
3. Would you recommend that other patients with myeloproliferative neoplasm (MPN) listen to the 12-week podcast prescription?
   - Yes
   - No
4. Do you feel like you learned something about your MPN, or about cancer in general, that you did not know before starting the study?
   - Yes
   - No
5. Have you made any changes to your normal daily activities because of something that you learned in the podcast?
   - Yes
   - No
6. Did you experience any limitations while trying to access the podcasts?
   - No, none
   - Bad or slow internet connection
   - Hard to hear or view the podcasts
   - Smartphone broken
   - Other, please describe: [free response textbox]
7. Do you have anything else you would like to share with us in regard to your participation in the MPN podcast study?
   - [Free response textbox]

Statistical Analysis

All analyses were conducted using SPSS 26.0 (IBM Corp). Descriptive statistics were used to characterize participants' app-use patterns over time, and frequency data from the satisfaction survey were used to describe participants' perceptions of the podcasts and the prescription schedule. Changes in PROMIS outcomes and MPN-SAF TSS were assessed using multivariate analyses of variance.

Results

Recruitment and Enrollment

Initially, 96 patients with MPN were enrolled in the study; however, 19 (19.8%) never began the intervention without disclosing their reasoning and without responding to contact attempts (see Figure 2). Of the 80 patients who participated in the intervention, 39 (48.8%) completed the entire study (ie, sustained participation through the follow-up period). There were 28 participants lost to follow-up that never responded to the three contact attempts after initial enrollment and 16 that

dropped out and provided the research team with a specific reason for discontinuing the study. The reasons for dropout included content not being MPN-specific (n=10), lack of time (n=3), internet connectivity issues (n=1), or illness or hospitalization (n=2).

Enrollment by week is presented in Figure 3. Analyses included data from all participants who were enrolled in the study through the point that they stopped using the app, defined as the point at which participants did not engage with the app for any subsequent weeks during the intervention period. For example, if a participant used the app during weeks 1, 2, 3, and 8, they were included in adherence analyses for weeks 1-8 (with use during weeks 4-7 calculated as 0 minutes) but were not included in analyses after the point when they stopped using the app (weeks 9-12), at which point they were considered to have discontinued study participation.

As shown in Table 1, the sample was predominately White, non-Hispanic, and female. The average age was 56.1 (SD 10.9) years. Most participants had earned a higher education degree, were married, and had an annual household income more than US $61,000. Study noncompleters were demographically similar to those who completed the study.

Figure 2. Study consort diagram. MPN: myeloproliferative neoplasm.

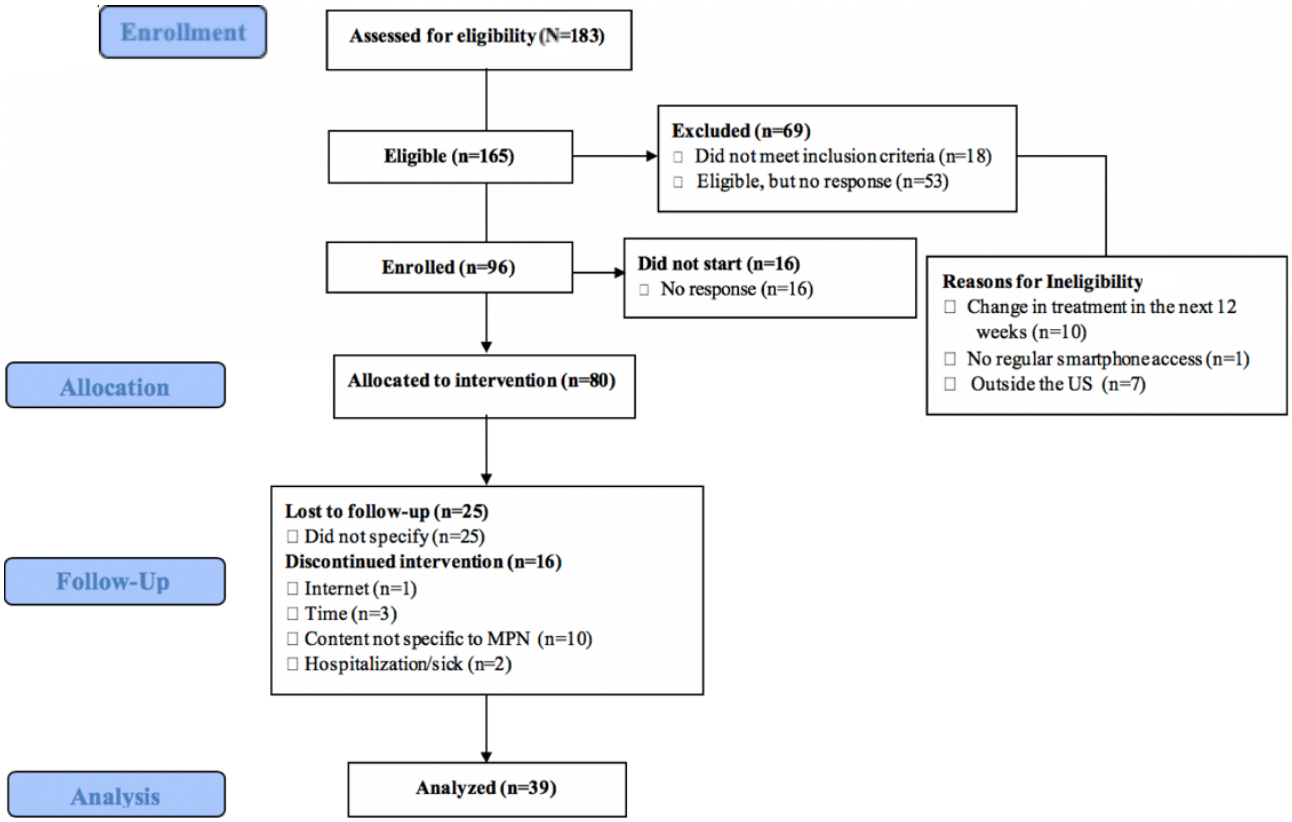


Figure 3. Participation and attrition by week during the intervention period.

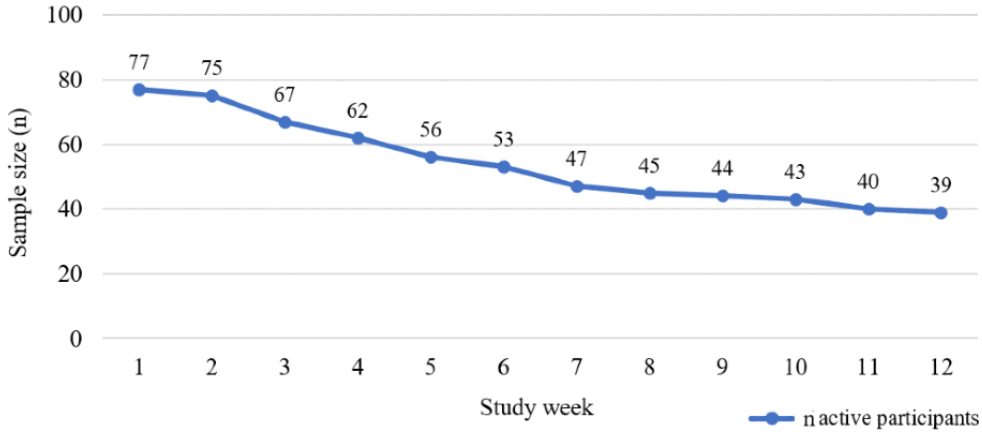


Table 1. Demographic characteristics of the sample.

Category	Completed study (n=39), n (%)	Did not complete study (n=38), n (%)
Gender		
Female	37 (94.9)	33 (86.8)
Male	2 (5.1)	5 (13.2)
Race		
White	38 (97.4)	35 (92.1)
Black/African American	0 (0.0)	1 (2.6)
Asian	1 (2.6)	1 (2.6)
American Indian	0 (0.0)	2 (5.3)
Other	1 (2.6)	1 (2.6)
Ethnicity[a]		
Non-Hispanic	38 (97.4)	37 (100.0)
Hispanic	1 (2.6)	0 (0.0)
Education		
High school/GED[b]	2 (5.1)	2 (5.3)
Some college	3 (7.7)	10 (26.3)
Associate's degree	1 (2.6)	5 (13.5)
Bachelor's degree	17 (43.6)	9 (23.7)
Graduate's degree	16 (41.0)	12 (31.6)
Marital status		
Single	2 (5.1)	3 (7.9)
Partnered	1 (2.6)	2 (5.3)
Married	33 (84.6)	28 (73.7)
Divorced	3 (7.7)	3 (7.9)
Widowed	0 (0.0)	2 (5.3)
Income (US $)[a]		
<20,000	3 (7.7)	5 (13.5)
21,000-40,000	3 (7.7)	5 (13.5)
41,000-60,000	5 (12.8)	7 (18.9)
>61,000	28 (71.8)	20 (54.1)
Chronic conditions		
Anxiety	10 (25.6)	12 (31.6)
Hypertension	8 (20.5)	8 (21.1)
Arthritis/rheumatic disease	7 (17.9)	5 (13.2)
Depression	6 (15.4)	8 (21.1)
Asthma	3 (7.7)	3 (7.9)
Hypercholesterolemia	3 (7.7)	4 (10.5)
PTSD[c]	2 (5.1)	1 (2.6)
Diabetes	1 (2.6)	3 (7.9)
Heart disease	1 (2.6)	0 (0.0)
Other	9 (23.1)	4 (10.5)
None	1 (2.6)	8 (21.1)

[a]Due to nonresponse, n=37 for ethnicity and income among those who did not complete the study. Percentages reflect percent of valid responses.

[b]GED: General Educational Development.

[c]PTSD: posttraumatic stress disorder.

Outcomes

Use

On average, participants listened to the health education podcasts for 103.2 (SD 29.5) minutes per week (see Figure 4), translating to an average of 4.9 (SD 0.9) completed podcast sessions each week (see Figure 5). For 83.3% (10/12) of the weeks, at least 70% of participants completed at least 70% of their total prescribed use (ie, 42 min/week; see Figure 6).

Figure 4. Average minutes using podcast app by week.

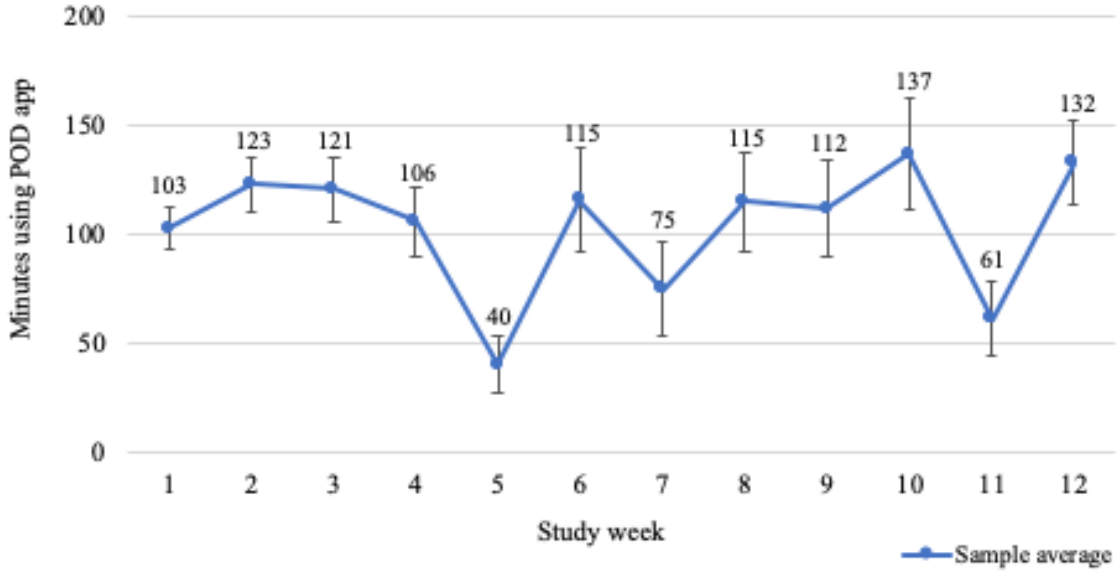


Figure 5. Average podcast sessions completed by week.

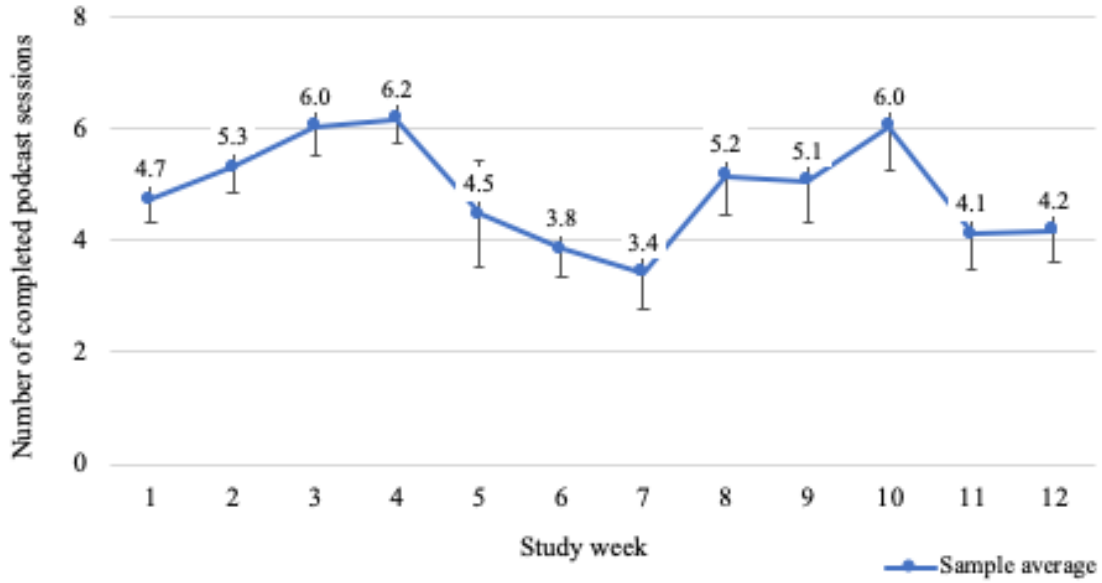
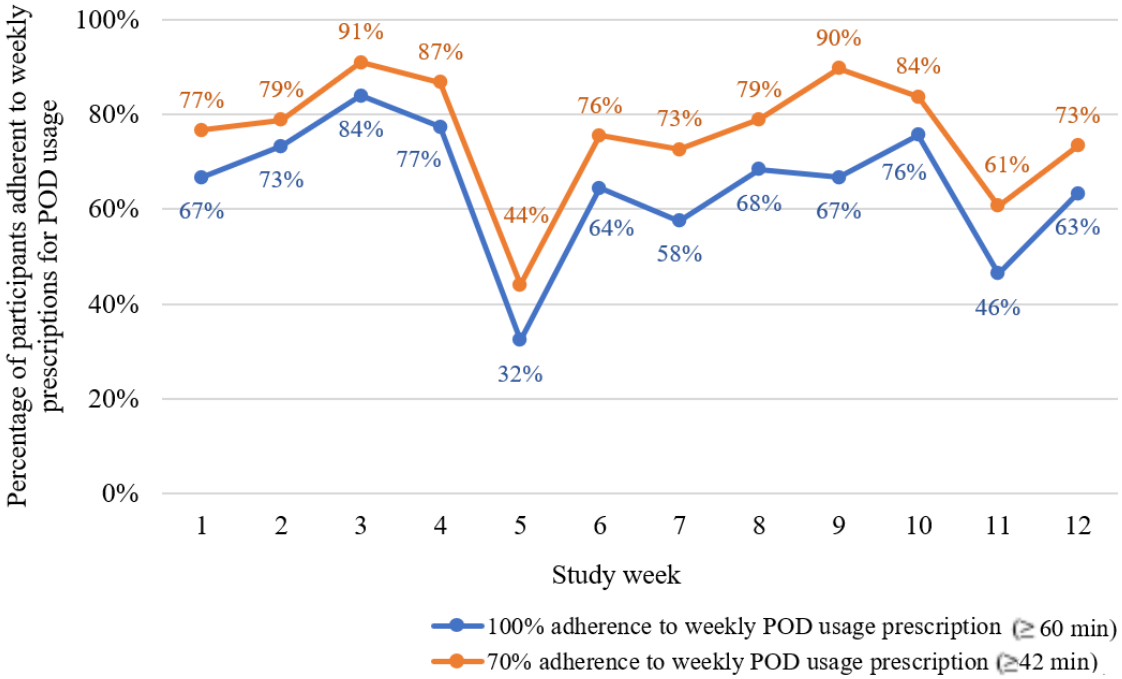


Figure 6. Percent adherence to prescribed podcast app use by week.



Satisfaction

Because we were unable to obtain satisfaction survey data from those who discontinued the study, satisfaction survey data were available only for study completers (n=39). Almost half of participants reported that, overall, they enjoyed the health education podcasts (n=19, 48.7%) and that they were satisfied with the intervention (n=17, 43.6%). Approximately half indicated that they learned something about their MPN or cancer in general from the podcasts (n=21, 53.8%), and 51.3% (n=20) reported that they made changes to their normal daily activities because of something that they learned in the podcasts. More than half of participants (n=22, 56.4%) indicated that they would recommend the 12-week podcast prescription to other patients with MPN.

Over half of the 39 participants (n=20, 51.3%) experienced some form of logistical limitation during the intervention. Specifically, 4 (10.3%) participants reported that they had difficulty hearing or viewing the podcasts (eg, too small on the screen), and 1 (2.6%) reported bad or slow internet connection. However, most frequently, participants indicated that they had "other" problems that were not available as survey response options (n=15, 38.5%). Of the 15 participants who provided open-ended responses describing their "other" difficulties, the most common were general problems with app functionality (n=7, 46.7%), problems accessing specific podcasts (n=5, 33.3%), or problems accessing any app content (ie, could not engage with any podcast; n=3, 20.0%). When asked if there was anything else they would like to share with us, the majority (21/34, 62%) that responded to this question made recommendations for changes to the podcast content, either to make it more specific to MPN or because they did not enjoy certain podcasts.

Changes in Health and Cancer-Related Symptoms

As shown in Table 2, using the health education podcasts was not associated with significant changes in global health, specific cancer-related symptoms, or MPN-SAF TSS.

Table 2. Changes in health and cancer-related symptoms during and after the health education podcast intervention (n=37).

Outcome	Baseline, mean (SD)	Week 6, mean (SD)	Week 12, mean (SD)	*F* test (*df*)[a]	*P* value[a]	Partial η^2
Global health				0.29 (1,36)	.88	0.033
Physical health	33.24 (7.22)	33.08 (7.50)	33.62 (7.75)	0.46 (1,36)	.50	0.012
Mental health	40.01 (7.65)	39.76 (8.29)	39.63 (8.53)	0.27 (1,36)	.61	0.007
Cancer-related symptoms				0.53 (1,36)	.83	0.127
Anxiety	51.25 (8.00)	52.99 (7.51)	52.52 (7.29)	1.54 (1,36)	.22	0.041
Depression	49.46 (7.54)	48.81 (6.83)	49.71 (7.11)	0.04 (1,36)	.84	0.001
Pain intensity	41.77 (7.44)	41.56 (8.13)	42.26 (8.42)	0.23 (1,36)	.64	0.006
Sleep disturbance	53.2 (8.62)	53.43 (8.55)	53.05 (8.56)	0.02 (1,36)	.90	<0.001
MPN-SAF TSS[b]	24.97 (15.44)	22.58 (15.63)	22.71 (16.00)	1.99 (1,36)	.17	0.051

[a]*F* test and *P* values assume symptom change to be a linear trend over time. Examination of change as a quadratic function produced similar results. For 2 participants, complete data were unavailable; they were excluded from the multivariate analyses.

[b]MPN-SAF TSS: Myeloproliferative Neoplasm Symptom Assessment Form Total Symptom Score.

Discussion

Principal Results

The purpose of this study was to beta test a health education podcast control group delivered via a smartphone app to inform the development of a comparator that could be used in future app-based meditation intervention studies. Feasibility benchmarks were met for demand (ie, for 10 of the 12 weeks at least 70% of participants listened to at least 42 min/week of podcasts); although, the attrition should be noted as an important qualifier when considering engagement more broadly. Acceptability benchmarks were not met, as less than half of participants enjoyed the podcasts and were satisfied with the intervention. As expected, there were no changes in global health, cancer-related symptoms, or total symptom burden (ie, MPN-SAF TSS) over the 12 weeks.

Comparison With Prior Work

Use benchmarks were set at ≥70% of participants listening to ≥70% (ie, 42 min/week) of the prescribed total weekly podcasts. In nearly all weeks (ie, 10 out of the 12 weeks) this benchmark was met. This is encouraging as a control group in which participants adhere to the study prescriptions allows researchers to control for nonintervention treatment effects such as time and attention [4]. Furthermore, participants averaged 103 min/week of podcast listening or viewing. Other studies that have tested mobile-app control groups have not reported weekly participation data [13,26]. In our research studies in which we used a commercially available meditation app in college students for 8 weeks [15] and to patients with MPN cancer for 4 weeks [16], average weekly participation in meditation on the app was ~38 min/week and ~71 min/week, respectively. Our findings are promising because the use of the podcast control app was comparable to the use levels in our intervention studies, supporting its feasibility for a control group with the same time and attention as a consumer-based mobile app meditation intervention.

Satisfaction benchmarks were not met, indicating that participants' overall satisfaction with the app was lacking. Less than half of participants indicated that they enjoyed the health education podcasts and less than half indicated that they were satisfied with the intervention. This is likely due to the podcast topics and technical issues and difficulties that came up during the intervention related to the functionality and usability of the app. The majority of participants that responded to an open-ended question to provide additional feedback about the study reported that they thought the app content should be modified to include more MPN-specific education or higher quality podcasts. The podcasts selected for the control app were intended to provide more generic cancer-related health education and not MPN-specific education. A better understanding of the potential users' content preferences with a more user-centered approach is necessary for future iterations of the control app, especially because the app was designed to be able to change health-education podcasts per target population being studied. A user-centered approach when developing products or software may help the user feel more at ease and make engaging with the content more intuitive, potentially improving adherence and enjoyment [27].

Over half of participants reported experiencing trouble accessing podcasts or content, or general problems with the functionality of the app. The app was not developed to be of commercial quality but was rather developed to be a "shell" of a design to be improved upon in subsequent developments after beta testing. Therefore, it is not surprising that such technical difficulties related to app functionality and accessibility were experienced. Specifically, there was a decrease in participation during week five, coinciding with participants reporting technical difficulties to the research staff. It is well known that user experience is a critical component to the success of mHealth apps and that apps must appeal to the motivations of the user [28]. For example, the unified theory of acceptance and use of technology (UTAUT) suggests that users' expectations of how an app will perform and how much effort it takes to use the app will influence their intentions and behaviors [29-31]. Future iterations of this app will use the current findings and a model-driven approach (eg, the UTAUT) to inform the

development of the next version of the app to be used as a comparator in a randomized control trial.

Importantly, there were no significant changes in health or cancer-related symptoms from baseline to postintervention. This is despite more than half of the participants reporting that they made changes to their normal daily activities because of something they learned from the podcasts. This is encouraging and indicates that this control group design is an appropriate time- and attention-matched condition that is not associated with any meaningful change in study outcomes. Education-based control groups have been used successfully as comparators in a range of smartphone-based interventions across different populations, typically without having significant effects on psychological or physical outcomes [13,16,26]. This indicates that an education-based control group could be appropriate for use in efficacy trials without producing changes in primary outcomes that are physical or psychological in nature. Future iterations of the app will be evaluated as a comparator in studies with other populations.

Limitations

This study is not without its limitations. First, attrition during the study was high, such that only 51% (41/80) of participants completed the full 12-week intervention. Given that the podcast app use (minutes listened per week) was high even when including use of participants who did not complete the study, future research should collect more nuanced data on participant satisfaction to improve podcast app user experiences and reduce attrition rates. For example, the lack of a postintervention interview did not allow for deeper qualitative analyses into what participants liked and did not like related to the app. This qualitative data could have been a useful addition for gathering deeper insights into user satisfaction and informing future development of the app. Second, the sample was predominantly White, non-Hispanic, and female. This is not representative of the general population of patients with MPN and, more importantly for the purpose of this beta test, does not capture feedback that reflects the experiences of different genders, races, and ethnicities. Future feasibility research should aim to include more diverse samples to gather more representative feedback. Third, there was no comparative group, and this must be considered when analyzing the results. Fourth, offering a US $25 digital gift card for completion of all questionnaires could contribute to higher engagement rates. Researchers may need to account for an impact on their participant engagement if they do not provide similar incentivization. Finally, the use of cancer-specific educational content would limit the app as currently developed to use in cancer studies only. However, the content of the app can easily be adjusted to fit the needs of different populations while maintaining the integrity and functionality of the app to match the Calm meditation app.

Future Research

Although this app was originally designed to be a comparative control app for app-based mindfulness meditation interventions involving Calm, the app is not reliant on the use of Calm and may be used as a control app in future app-based interventions. Currently, this app is not available to other researchers because it was only developed as a template to then create an improved comparator app for future mindfulness meditation app-based interventions based on the data collected. Once the app can be determined feasible, future interventions will be developed to assess app-based mindfulness meditation interventions as compared to the control app to assess many aspects of health and well-being across various populations.

Conclusion

In summary, a 12-week mobile app health education podcast met the demand benchmark but not the acceptability benchmark for feasibility in a sample of patients with hematological cancer. Using the mobile app health education podcast was not associated with significant changes in cancer-related symptoms. Participants reported dissatisfaction with content and technical or functionality difficulties, which will be addressed in the future development of the health education podcast app. Based on findings from this study, a mobile app health education podcast may be a feasible option as a time- and attention-matched comparator condition for efficacy trials with more extensive formative research for the content of the podcasts and its acceptability by the specific population.

Acknowledgments

This study was funded with philanthropic funds.

Conflicts of Interest

JH and LL both report being members of the Scientific Advisory Board for Calm and serve in paid consulting roles. Neither earn additional financial compensation from the sale of the app. The remaining authors report no conflicts of interest.

References

1. Mobile fact sheet. Pew Research Center. 2019. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2019-12-31]
2. mHealth: New horizons of health through mobile technology. World Health Organization. 2011. URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf? [accessed 2018-12-31]
3. Robbins R, Krebs P, Jagannathan R, Jean-Louis G, Duncan DT. Health app use among US mobile phone users: analysis of trends by chronic disease status. JMIR Mhealth Uhealth 2017 Dec 19;5(12):e197 [FREE Full text] [doi: 10.2196/mhealth.7832] [Medline: 29258981]

4. Dorsey ER, Yvonne Chan YF, McConnell MV, Shaw SY, Trister AD, Friend SH. The use of smartphones for health research. Acad Med 2017 Feb;92(2):157-160. [doi: 10.1097/ACM.0000000000001205] [Medline: 27119325]
5. Kinser PA, Robins JL. Control group design: enhancing rigor in research of mind-body therapies for depression. Evid Based Complement Alternat Med 2013;2013:140467. [doi: 10.1155/2013/140467] [Medline: 23662111]
6. Hart T, Fann JR, Novack TA. The dilemma of the control condition in experience-based cognitive and behavioural treatment research. Neuropsychol Rehabil 2008 Jan;18(1):1-21. [doi: 10.1080/09602010601082359] [Medline: 17852761]
7. Duraimani SL. A cross-sectional and longitudinal study of the effects of a mindfulness meditation mobile application platform on reducing stress and anxiety. Int J Yoga 2019;12(3):226-232 [FREE Full text] [doi: 10.4103/ijoy.IJOY_56_18] [Medline: 31543631]
8. Flett JAM, Hayne H, Riordan BC, Thompson LM, Conner TS. Mobile mindfulness meditation: a randomised controlled trial of the effect of two popular apps on mental health. Mindfulness 2018 Oct 31;10(5):863-876. [doi: 10.1007/s12671-018-1050-9]
9. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can smartphone apps increase physical activity? systematic review and meta-analysis. J Med Internet Res 2019 Mar 19;21(3):e12053 [FREE Full text] [doi: 10.2196/12053] [Medline: 30888321]
10. Freedland KE, King AC, Ambrosius WT, Mayo-Wilson E, Mohr DC, Czajkowski SM, National Institutes of Health Office of Behavioral and Social Sciences Research Expert Panel on Comparator Selection in Behavioral and Social Science Clinical Trials. The selection of comparators for randomized controlled trials of health-related behavioral interventions: recommendations of an NIH expert panel. J Clin Epidemiol 2019 Jun;110:74-81 [FREE Full text] [doi: 10.1016/j.jclinepi.2019.02.011] [Medline: 30826377]
11. Lunde P, Nilsson BB, Bergland A, Kværner KJ, Bye A. The effectiveness of smartphone apps for lifestyle improvement in noncommunicable diseases: systematic review and meta-analyses. J Med Internet Res 2018 May 04;20(5):e162 [FREE Full text] [doi: 10.2196/jmir.9751] [Medline: 29728346]
12. Bostock S, Crosswell AD, Prather AA, Steptoe A. Mindfulness on-the-go: effects of a mindfulness meditation app on work stress and well-being. J Occup Health Psychol 2019 Feb;24(1):127-138 [FREE Full text] [doi: 10.1037/ocp0000118] [Medline: 29723001]
13. Champion L, Economides M, Chandler C. The efficacy of a brief app-based mindfulness intervention on psychosocial outcomes in healthy adults: a pilot randomised controlled trial. PLoS One 2018;13(12):e0209482. [doi: 10.1371/journal.pone.0209482] [Medline: 30596696]
14. Economides M, Martman J, Bell MJ, Sanderson B. Improvements in stress, affect, and irritability following brief use of a mindfulness-based smartphone app: a randomized controlled trial. Mindfulness (N Y) 2018;9(5):1584-1593 [FREE Full text] [doi: 10.1007/s12671-018-0905-4] [Medline: 30294390]
15. Carissoli C, Villani D, Riva G. Does a meditation protocol supported by a mobile application help people reduce stress? Suggestions from a controlled pragmatic trial. Cyberpsychol Behav Soc Netw 2015 Jan;18(1):46-53. [doi: 10.1089/cyber.2014.0062] [Medline: 25584730]
16. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the mindfulness meditation mobile app "Calm" to reduce stress among college students: randomized controlled trial. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14273 [FREE Full text] [doi: 10.2196/14273] [Medline: 31237569]
17. Huberty J, Eckert R, Larkey L, Kurka J, Rodríguez De Jesús SA, Yoo W, et al. Smartphone-based meditation for myeloproliferative neoplasm patients: feasibility study to inform future trials. JMIR Form Res 2019 Apr 29;3(2):e12662 [FREE Full text] [doi: 10.2196/12662] [Medline: 31033443]
18. Adams ZW, Sieverdes JC, Brunner-Jackson B, Mueller M, Chandler J, Diaz V, et al. Meditation smartphone application effects on prehypertensive adults' blood pressure: dose-response feasibility trial. Health Psychol 2018 Sep;37(9):850-860 [FREE Full text] [doi: 10.1037/hea0000584] [Medline: 30010353]
19. 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]
20. Huberty J, Eckert R, Gowin K, Mitchell J, Dueck AC, Ginos BF, et al. Feasibility study of online yoga for symptom management in patients with myeloproliferative neoplasms. Haematologica 2017 Oct;102(10):e384-e388 [FREE Full text] [doi: 10.3324/haematol.2017.168583] [Medline: 28596279]
21. Cessna JM, Jim HSL, Sutton SK, Asvat Y, Small BJ, Salsman JM, et al. Evaluation of the psychometric properties of the PROMIS Cancer Fatigue Short Form with cancer patients. J Psychosom Res 2016 Feb;81:9-13 [FREE Full text] [doi: 10.1016/j.jpsychores.2015.12.002] [Medline: 26800633]
22. Flynn KE, Lin L, Cyranowski JM, Reeve BB, Reese JB, Jeffery DD, et al. Development of the NIH PROMIS ® Sexual Function and Satisfaction measures in patients with cancer. J Sex Med 2013 Feb;10 Suppl 1:43-52 [FREE Full text] [doi: 10.1111/j.1743-6109.2012.02995.x] [Medline: 23387911]
23. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. Qual Life Res 2009 Sep;18(7):873-880 [FREE Full text] [doi: 10.1007/s11136-009-9496-9] [Medline: 19543809]

24. Quach CW, Langer MM, Chen RC, Thissen D, Usinger DS, Emerson MA, et al. Reliability and validity of PROMIS measures administered by telephone interview in a longitudinal localized prostate cancer study. Qual Life Res 2016 Nov;25(11):2811-2823 [FREE Full text] [doi: 10.1007/s11136-016-1325-3] [Medline: 27240448]
25. Scherber R, Dueck AC, Johansson P, Barbui T, Barosi G, Vannucchi AM, et al. The Myeloproliferative Neoplasm Symptom Assessment Form (MPN-SAF): international prospective validation and reliability trial in 402 patients. Blood 2011 Jul 14;118(2):401-408. [doi: 10.1182/blood-2011-01-328955] [Medline: 21536863]
26. Bush NE, Smolenski DJ, Denneson LM, Williams HB, Thomas EK, Dobscha SK. A Virtual Hope Box: randomized controlled trial of a smartphone app for emotional regulation and coping with distress. Psychiatr Serv 2017 Apr 01;68(4):330-336. [doi: 10.1176/appi.ps.201600283] [Medline: 27842473]
27. Goodwin K. Designing for the Digital Age: How to Create Human-Centered Products and Services. Hoboken, New Jersey: John Wiley & Sons; 2011.
28. Liew MS, Zhang J, See J, Ong YL. Usability challenges for health and wellness mobile apps: mixed-methods study among mHealth experts and consumers. JMIR Mhealth Uhealth 2019 Jan 30;7(1):e12160 [FREE Full text] [doi: 10.2196/12160] [Medline: 30698528]
29. 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 reviews. J Med Internet Res 2018 Jun 06;20(6):e10120 [FREE Full text] [doi: 10.2196/10120] [Medline: 29875087]
30. Di Sorbo AP, Alexandru C, Shimagaki J, Visaggio C, Canfora G, Gall H. What would users change in my app? Summarizing app reviews for recommending software changes. In: Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering. 2016 Presented at: 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering; 2016; Seattle, Washington p. 499-510. [doi: 10.1145/2950290.2950299]
31. Frie K, Hartmann-Boyce J, Jebb S, Albury C, Nourse R, Aveyard P. Insights from Google Play Store user reviews for the development of weight loss apps: mixed-method analysis. JMIR Mhealth Uhealth 2017 Dec 22;5(12):e203 [FREE Full text] [doi: 10.2196/mhealth.8791] [Medline: 29273575]

Abbreviations

mHealth: mobile health
MPN: myeloproliferative neoplasm
MPN-SAF TSS: Myeloproliferative Neoplasm Symptom Assessment Form Total Symptom Score
NIH: National Institutes of Health
PROMIS: Patient-Reported Outcomes Measurement Information System
UTAUT: unified theory of acceptance and use of technology

Edited by G Eysenbach; submitted 14.04.20; peer-reviewed by M Economides, K Frie; comments to author 11.05.20; revised version received 27.05.20; accepted 14.06.20; published 21.07.20.

Please cite as:
Huberty J, Eckert R, Puzia M, Laird B, Larkey L, Mesa R
A Novel Educational Control Group Mobile App for Meditation Interventions: Single-Group Feasibility Trial
JMIR Form Res 2020;4(7):e19364
URL: http://formative.jmir.org/2020/7/e19364/
doi:10.2196/19364
PMID:32706719

Original Paper

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

Mark D Attridge[1], MA, DPhil; Russell C Morfitt[2], DPhil; David J Roseborough[3], DPhil; Edward R Jones[2], DPhil

[1]Attridge Consulting, Inc, Minneapolis, MN, United States

[2]Learn to Live, Inc, Minneapolis, MN, United States

[3]School of Social Work, University of St Thomas, St Paul, MN, United States

Corresponding Author:
Mark D Attridge, MA, DPhil
Attridge Consulting, Inc
1129 Cedar Lake Rd S
Minneapolis, MN
United States
Phone: 1 612 889 2398
Email: mark@attridgeconsulting.com

Abstract

Background: The growing behavioral health needs of college students have resulted in counseling centers reporting difficulties in meeting student demand.

Objective: This study aims to test the real-world voluntary use by college students of 4 digital, self-directed mental health modules based on a cognitive behavioral therapy clinical model. The findings were also compared with those of employee users.

Methods: Archival operational data from Learn to Live were extracted for student users at 4 colleges and universities in the Midwest region of the United States (N=951). The inclusion criteria were having clinical symptoms at established levels of moderate or higher severity and the use of 2 or more of the 8 lessons of a program within a 6-month period. Unique users in each program included 347 for depression; 325 for stress, anxiety, and worry; 203 for social anxiety; and 76 for insomnia. Paired *t* tests (two-tailed) compared the average level of change over time on a standardized measure of clinical symptoms appropriate to each program. Cohen *d* statistical effect sizes were calculated for each program. Potential moderator factors (age, gender, preliminary comprehensive assessment, number of lessons, duration, live coach support, and live teammate support) were tested together in repeated measures analysis of variance models with covariates in the full sample. Follow-up survey data (n=136) were also collected to explore user satisfaction and outcomes. Select data from another study of the same 4 programs by employee users meeting the same criteria (N=707) were examined for comparison.

Results: The percentage of users who improved to a clinical status of no longer being at risk after program use was as follows: stress, anxiety, and worry program (149/325, 45.8%); insomnia program (33/76, 43.4%), depression program (124/347, 35.7%); and social anxiety program (45/203, 22.2%). Significant improvements (all $P<.001$) over time were found in the mean scores for the clinical measures for each program: stress, anxiety, and worry ($t_{324}=16.21$; $d=1.25$); insomnia ($t_{75}=6.85$; $d=1.10$); depression ($t_{346}=12.71$; $d=0.91$); and social anxiety ($t_{202}=8.33$; $d=0.80$). Tests of the moderating factors across programs indicated that greater improvement was strongly associated with the use of more lessons and it also differed by program, by gender (males demonstrated more improvement than females), and by the use of live support (particularly coaching). Analyses of survey data found high satisfaction, improved academic outcomes, and successful integration into the university counseling ecosystem. The operational profile and outcomes of the college students were also similar to those of employee users of the same programs from our other study of employee users. Thus, this study provides a replication.

Conclusions: Self-directed internet-based cognitive behavioral therapy mental health modules are promising as a supplement to traditional in-person counseling services provided by college counseling centers.

(JMIR Form Res 2020;4(7):e17712) doi:10.2196/17712

KEYWORDS

anxiety; cognitive behavioral; college; depression; insomnia; mental health; social phobia; stress

Introduction

Background

Adjusting to college life is a significant challenge for many students [1], as they could struggle with exposure to new forms of stress and new social demands. Students navigating these challenges may experience loneliness, may lack a sense of campus belonging, and can be vulnerable to experiences such as the first depressive episode [2]. The onset of many mental health conditions coincides with adolescence and young adulthood [3].

A 2019 national survey of 43,140 college undergraduate and graduate students in the United States found that 42% of females and 34% of males met the criteria for moderate or severe psychological distress [4]. The same study also documented the past year prevalence rates for diagnosed mental health conditions and determined that 28% of females and 13% of males had depression and that 22% of females and 15% of males had an anxiety disorder (including social anxiety). Graduate students may face additional stressors and risks for mental health issues [5]. For example, in the same recent national survey in 2019, 33% of graduate students [6] received psychological health services in the past year compared with 25% of undergraduate students [7].

Although not a mental health disorder, sleep issues also pose risks for young adults attending colleges and universities [8]. For example, 50% of college students in the United States reported getting less than the recommended 7 hours of sleep during nights of the school week [4]. Although sleep is important, insufficient sleep is of particular risk for students with bipolar disorder who have an acute sensitivity to circadian and sleep-wake cycle disruption (what Frank [9] calls “sleep disrupters”). This realization has led sleep researchers Walker and van der Helm [10] to refer to good quality sleep as “overnight therapy.” Some colleges now offer resources specifically to address student sleep problems, such as the College Sleep Center at the University of St. Thomas [11].

Experiencing mental health problems can also adversely impact success at school. A recent national survey [4] found that college students reported the following common behavioral health issues had negatively affected their academic performance in the past year: stress (38%), anxiety (28%), sleep difficulties (23%), and depression (22%). Other studies have also shown that depression among college students is linked to poor academic performance [12] and to dropping out of school [13].

Technology-Based Mental Health Resources

The growing behavioral health needs of college and university students have resulted in college counseling centers reporting difficulties in meeting the increased student demand [14,15]. This context provided an opportunity to test if internet-based cognitive behavioral therapy (iCBT) tools might assist students in self-treating some of the most commonly occurring mental health conditions [16]. A related question was whether self-directed technology support services can supplement in-person clinical relationships and encourage the appropriate use of on-campus counseling services.

New types of technology-based resources feature self-directed digital tools that are accessible from a website or mobile device. Most of these tools are asynchronous and do not involve live interaction with a mental health professional, although some are supplemented with live coaching. Many of these computerized tools are based on principles and clinical strategies derived from cognitive behavioral therapy (CBT) [17]. The general advantages of such tools include providing users with greater access to therapeutic support, flexibility in accessing support anytime from anywhere with internet access, and a significantly lower cost than in-person services. The privacy of technology tools can also help offset social stigma and related barriers to help-seeking that confront young adults [18]. According to recent reviews of the literature, at least 89 studies have explored the use of technology-based tools for college students interested in seeking support for a range of mental health, stress, and behavioral concerns [19-21].

These reviews offer support for iCBT tools in general for being able to improve clinical outcomes, but they also call for further research to be done in real-world settings that go beyond the use of college students as convenience samples in academic research pilot studies for testing the general efficacy of tools. Few studies have used an applied study design with archival data to examine the naturally occurring experiences of voluntary users of commercially available iCBT tools for mental health issues. This emphasis on experimental over applied real-life contexts for research in this area is reflected in the literature in general and especially among investigations using college students as participants.

Interventions: The Learn to Live Suite of Self-Directed Web-Based Tools

This applied archival study was conducted with operational data from voluntary, registered users of Learn to Live, a suite of digital CBT-based self-directed programs for behavioral health issues, including anxiety, depression, social anxiety, and insomnia. The programs were hosted on a single dedicated website [22] and were accessible from any internet-capable device. Participants entered a code specific to their college or university on the website to get access to the service. In addition to a wide range of educational content on the website, the user had the option to begin their online experience by taking a brief assessment covering 5 domains: anxiety, stress, depression, social anxiety, and insomnia. Standardized and validated clinical assessment tools were used for each domain. After reviewing the assessment results, participants then had the opportunity to enroll in 1 of 4 CBT-based programs. Alternatively, users could directly start a program without first taking a comprehensive assessment. The very small percentage of users who were identified as at-risk, based on self-harm questions on the preliminary assessment, were referred to a crisis center for immediate support.

Each program consisted of 8 lessons that contained a brief assessment of clinical severity (repeated every lesson; using the same self-report assessment tool from the comprehensive assessment), videos, animations, and web-based application of CBT tools. Each set of lessons was designed to be completed in order from 1 to 8, and a prior lesson had to be completed before the user could progress to the next lesson in the program. Homework and practice with the tools was optional but recommended between lessons. Key elements of each lesson for each of the 4 programs are listed in Table 1.

In addition, users could opt to receive individualized coaching, and if selected, could choose the preferred channel of communication, either email, text, or telephone, with the coach. The coaches were employed by the service provider, and every coach had at least a master's level education in psychology or counseling.

Users could also select a person from their personal life to communicate with during the program use, serving in a supportive role called a *teammate*. These live supports are offered with the goal of adding a relational component beyond self-help as a form of activating natural social support [23]. Coaching support from real people while using iCBT programs has also been demonstrated to improve outcomes among college users of technology-based tools for behavioral health issues [21,24,25].

Support for these 4 programs from Learn to Live was obtained in an earlier study [26] that involved a sample of 707 employee users who worked with multiple employers (also located in the Midwest region of the United States). The employee study used the same study design, the same archival data collection processes, and the same general time frame as this study of college student users. However, the employee study examined the experiences of users at both subclinical and clinical statuses at the start of program use for symptom severity level. Relevant data from only the clinical status group of the employee users were reanalyzed and presented in this study for comparative purposes with the college student users who were all starting out at the clinical status level of severity.

Table 1. Intervention elements and assessment for each lesson for the 4 Learn to Live iCBT programs.

Lesson	Intervention elements	Assessment
Stress, anxiety, and worry		
1	PMR[a] + better life goals + stress and anxiety tracker	GAD-7[b]
2	STEPP[c] model + mini thought inspection + ANTs[d]	GAD-7
3	Full thought inspection + 12 assumptions + precautions	GAD-7
4	Flaw-facing + worry-facing + perfectionism	GAD-7
5	Active problem solving + time snapshot + reflection moment	GAD-7
6	Present awareness + worry time	GAD-7
7	Assertiveness + boundaries	GAD-7
8	Lessons learned toolbox	GAD-7
Depression		
1	Depression profile + my better life goals + activity log	PHQ-9[e]
2	Testable hypothesis + identify 30-minute exercise + precautions	PHQ-9
3	STEPP model + mini thought inspection + ANTs	PHQ-9
4	Forgiveness scripts + thought inspection + 12 assumptions	PHQ-9
5	Active problem solving + learned helplessness	PHQ-9
6	Alternatives to dwelling + sleep enhancement form	PHQ-9
7	Assertiveness + boundaries	PHQ-9
8	Lessons learned toolbox	PHQ-9
Social anxiety		
1	Social anxiety profile + social life goals	SPIN-17[f]
2	STEPP model + identifying thoughts	SPIN-17
3	Thought inspection + hot thought	SPIN-17
4	Self-defense tactics checklist + find out for myself	SPIN-17
5	Full thought inspection + ANTs	SPIN-17
6	Fear facing trials list + fear facing log + fear facing menu	SPIN-17
7	Fear facing debrief + avoidance	SPIN-17
8	Lessons learned toolbox	SPIN-17
Insomnia sleep		
1	Sleep tracker	MOS-Sleep-6[g]
2	Alternatives to lying awake + sleep drive + sleep scheduling + recommended bedtime	MOS-Sleep-6
3	Sleep helpers form + sleep barriers	MOS-Sleep-6
4	PMR + guided imagery + worry notebook	MOS-Sleep-6
5	STEPP model + mini thought inspection + ANTs	MOS-Sleep-6
6	Thought inspection + deep sleep	MOS-Sleep-6
7	Present awareness + worry time	MOS-Sleep-6
8	Lessons learned toolbox	MOS-Sleep-6

[a]PMR: progressive muscle relaxation.

[b]GAD-7: Generalized Anxiety Disorder 7-item scale.

[c]STEPP: situation-thought-emotion-performance-precautions.

[d]ANTs: automatic negative thoughts.

[e]PHQ-9: Patient Health Questionnaire 9-item scale.

[f]SPIN-17: Social Phobia Inventory 17-item scale.

[g]MOS-Sleep-6: Medical Outcomes Study Sleep 6-item scale.

Objectives of the Study

This study featured a longitudinal, repeated measures research design with archival data for college student users of iCBT programs offered by the same commercial provider that supported 4 different clinical issues. The goals of this study were to obtain empirical answers to the following research questions (RQs):

- RQ1: What is the profile among college students based on demographic factors and utilization factors for the 4 Learn to Live web-based iCBT programs?
- RQ2: Are each of the 4 programs effective in reducing the level of clinical symptoms after program use?
- RQ3: Is the extent of improvement in clinical symptoms after program use moderated by demographic factors of the user (age or gender) or by operational factors (ie, use of preliminary comprehensive assessment, number of lessons, time period of use, use of live support from a coach or from a teammate, or use of multiple programs)?
- RQ4: Are these moderator effects (if any) similar, when tested within each of the 4 programs?
- RQ5: From the survey data collected following the intervention, what can be learned about the different sources of promotion of the services, the impact of use on college-related outcomes and attitudes, different aspects of the user experience with counseling, and the level of overall user satisfaction with the service?
- RQ6: How does the user profile and clinical outcomes of the 4 programs for college student users in this study compare with that of employee users from a different study?

Methods

Archival Data

The data for the study were from postsecondary students using the *Learn to Live* service. Users were made aware of the service as a benefit open to all students through a variety of on-campus digital and interpersonal promotional practices. There was no direct cost to the participants in this study, as access to the website with the programs was sponsored by each of the schools. Students participated voluntarily and were not paid for using the web-based tools. The study period spanned 3 years (from October 2016 through October 2019). The insomnia program was added to the service suite in November 2017 and therefore was available only for the most recent 2 years of the full study period.

Ethical Considerations

The privacy of users was protected by having all program and survey data deidentified before being shared with the independent consultant who conducted all of the analyses. As this was an applied study of archival anonymized data collected from routine use of the service, additional informed consent from individual participants beyond their initial consent agreement in terms of use was not required. Project approval from a college or a university internal review board was also not required. This context is similar to other applied studies of commercial web-based programs [26,27]. The use and analysis of archival operational data in this manner is consistent with the published ethical guidelines of the American Psychological Association [28].

Inclusion Criteria and Participants

The following 5 criteria were established to select users appropriate to the study goals: (1) users had to be from a customer of *Learn To Live* in the higher education market segment (ie, a college or university), (2) users were required to be at a sufficient level of clinical severity (*clinical status*) at the start of program use, (3) users needed to engage in at least 2 (or more) of the 8 lessons of a program, and (4) users should not have completed all 8 lessons of a program in a single day nor should the time period of use from the start date (either the comprehensive assessment or the first lesson) to the date of the last lesson have exceeded 6 months. Application of these criteria yielded a sample of 951 unique users.

The final criterion was that if a student had used multiple iCBT programs during the 3-year period, then the experience data from only one program were included. Of the sample that met all 4 of the above inclusion criteria, most participants (806/951, 84.8%) used only 1 program, with the remaining 15.2% (145/951) having used multiple programs. Using more than 1 program was defined as using at least 1 lesson of 2 or more different programs during the study period. More specifically, 122 (12.8%) students had used 2 programs, 19 (1.9%) students had used 3 programs, and 4 (0.4%) students had used all 4 programs. The average number of lessons *per program* varied between these groups with a decreasing number of lessons per any one program as the total number of programs used increased: Users of 1 program had an average of 3.44 lessons per program; users of 2 programs used had an average of 2.25 lessons per program; users of 3 programs used had an average of 1.80 lessons per program; and users of all 4 of the programs had an average of only 1.68 lessons per program. Note that at least one of the programs used by an individual who had used 2 or more programs, had to have a minimum of 2 lessons used.

The choice of which program's data to use for each student with multiprogram status was based on several criteria. Listed in order of importance, these criteria included: (1) clinical status on the symptom assessment at the start, (2) a higher number of lessons used (two minimum), (3) earlier start date in the study period, and (4) a longer time period for program use. For example, if a student had met the first set of criteria for use of 2 programs and had a score above the clinical score cutoff for the depression program but also a score that was below the clinical score cutoff for the insomnia program, then the data for the former was retained but not of the latter. To continue, if a student was above the cutoff scores for the severity symptom measures for both programs, then the program that had more lessons was retained. For example, if the depression program had 5 lessons and the insomnia program had 3 lessons, then the depression program was retained. Application of these criteria resulted in the following final mix of participants with data from only one program used: Of the 951 total unique users in the study, 347 (36.5%) users for the depression program, 325

(34.2%) users for the stress, anxiety, and worry program, 203 (21.3%) users for the social anxiety program, and 76 (8.0%) users for the insomnia program.

Clinical Symptom Measures

Each of these measures of clinical symptom severity is a published, reliable, and validated scale from the scientific literature. Within each program, the symptom measure was repeated in every lesson. These 4 measures were aggregated in the comprehensive assessment (along with a fifth measure of perceived stress).

Anxiety

The generalized anxiety disorder 7-item scale was used to assess symptoms of anxiety [29]. This is one of the most widely used screening and outcome tools available for anxiety and has been shown, in past research, to have adequate levels of reliability and validity [30,31]. Sample items include the following: (1) *feeling nervous, anxious or on edge* and (2) *not being able to stop or control worrying*. The instructions state: "Over the last 2 weeks, how often have you been bothered by any of the following problems?" Items are rated on a 0 to 3 scale. Ratings on the items were summed and scores were categorized into levels of severity: low 0-4, mild 5-9, moderate 10-14, and severe 15-21. The clinical status for anxiety was defined as moderate or higher (score of 10+). The severity mix for general anxiety in the sample was 52.6% moderate (171/325) and 47.4% severe (154/325).

Depression

The patient health questionnaire 9-item scale was used to assess symptoms of depression [32]. This scale has been used in hundreds of research studies and has well-established validity and reliability [33]. The instructions state: "Over the last 2 weeks, how often have you been bothered by any of the following problems?" Sample items included the following: *Little interest or pleasure in doing things* and *feeling down, depressed, or hopeless*. Items are rated on a 0-3 scale. Scores on the 9 items were summed and then categorized into levels of severity: minimal 0-4, mild 5-9, moderate 10-14, moderately severe 15-19, and severe 20-27. The clinical status for depression was defined as moderate or higher (score of 10+). The severity mix for depression in the sample was 32.3% moderate (112/347), 44.7% moderately severe (155/347), and 23.0% severe (80/347).

Social Anxiety

The social phobia inventory (SPIN) was used to assess symptoms of social anxiety [34]. Past research has shown SPIN to have adequate levels of reliability and validity [35]. The instructions state: "Select the answer that best describes how much the following problems have bothered you during the past week." Scores on the 17 items are rated on a 0-4 scale. Scores were summed and categorized into 5 levels of severity: minimal 0-18, mild 19-30, moderate 31-40, severe 41-50, and very severe 51-68. The clinical status for social anxiety was defined as moderate or higher (score of 31+). The severity mix for social anxiety in the study sample was 36.0% moderate (73/203), 42.4% severe (86/203), and 21.6% very severe (44/203).

Insomnia

To assess symptoms of sleep disturbance and insomnia, the sleep scale from the medical outcomes study (MOS) developed by the Rand Corporation [36]. The MOS sleep scale has proved to have adequate levels of reliability and validity [37]. The 6-item short version used item numbers 4, 5, 7, 8, 9, and 12 from the original full 12-item scale. The instructions for the measure state: "How often during the past week did you…?" The items included the following: (4) *get enough sleep to feel rested upon waking in the morning?* (5) *awaken short of breath or with a headache?* (7) *have trouble falling asleep?* (8) *awaken during your sleep time and have trouble falling asleep again?* (9) *have trouble staying awake during the day?* and (12) *get the amount of sleep you needed?* The recall period was slightly modified for use by Learn to Live, such that in the preliminary comprehensive assessment, the instructions for this scale used the *past 4 weeks* reference time period, whereas in each lesson of the program, the instructions had a reference time period of the *past week*. The 6 items were rated on a scale of 1-6. The ratings were weighted (1=0, 2=20, 3=40, 4=60, 5=80, and 6=100) and summed. The total score was then categorized into 4 levels of severity: minimal 0-29, mild 30-43, moderate 44-60, and severe 61-100. The clinical status of insomnia was defined as moderate or higher (score of 44+). The severity mix of insomnia in the study sample was 55.3% moderate (42/76) and 44.7% severe (34/76).

Sources of First and Last Scores on Outcome Measures

The data source for the score at the start of program use was most often the student's score from the preliminary comprehensive assessment, which had been completed by almost 9 out of every 10 participants (829/951, 87.2%). For those who did not complete the comprehensive assessment, the score for the start of program use was taken from the symptom assessment done as part of the first lesson. The data source for the last score at the end of program use was taken from the student's score on the last lesson used, which varied within person from lesson 2 to lesson 8. Each program had a full range of lessons represented from 2 to 8.

Follow-Up Survey

All registered users of the *Learn to Live* services were sent an email and invited to complete a self-report survey about their experiences. Modest financial incentives were provided to students who participated in a follow-up survey. Note that offering incentives for survey completion was a routine component of business operations and not a procedure unique to the research study. The specific questions and response options and the findings are presented later in this paper. A total of 136 users completed a survey during the valid follow-up period, defined as at least 1 month but not more than 6 months after the date of the last use of the program (between 31 and 183 days after the date of the last lesson used). About 1 out of every 8 students in the study (136/951, 14.3%) completed a valid follow-up survey.

Preliminary tests were performed by comparing the survey sample with the others in the total sample to determine the representativeness of the survey group. Specific statistical results

for each of the measures compared are shown in Multimedia Appendix 1. The survey group was similar to the nonsurvey group on factors of age, completing the comprehensive assessment, and using a teammate. However, the survey group differed from the nonsurvey group in that it had more females, used more lessons, had a longer timer period of use, had more who used a coach, and had more who had tried more than one of the programs. Although those who participated in the follow-up survey process had greater engagement with the program, both groups had nearly identical clinical outcomes. Thus, the survey sample was considered representative of each program and the overall program use experience, despite some differences between the respondents and nonrespondents.

Data Analysis

All analyses were conducted using SPSS version 25 [38]. Descriptive and inferential tests were performed as appropriate to the data and research questions. Details on the specific analyses performed are presented in the Results section.

Results

Part 1: Profile of College Student Users of iCBT Programs

This part of the paper describes the profile of college students based on demographic factors and their use of the 4 programs.

Demographic Profile of the Total Sample

The sample included 951 students from 4 colleges and universities, all located in the Midwest region of the United States. In the total sample, the average age of the users was 23.38 years (SD=6.54) and ranged from 15 to 62 years, although most users were in the 18 to 21 age group, typically of the college undergraduate experience. Females comprised the vast majority of program users at 74.2% (706/951), males were 23.6% (224/951), and 2.2% (21/951) of users self-identified as *gender diverse*.

Other background factors of race and year in college were only collected in the follow-up survey (described later in the paper, *n*=136). For race, 80.1% (109/136) identified as white, 8.1% (11/136) as Asian, 3.7% (5/136) as black or African American, 1.5% (2/136) as Hispanic or Latino, and another 6.6% (9/136) as *Other* or *no answer*. Most of the users who completed the survey were undergraduate students (85/136, 62.5%), although slightly more than a third were graduate students (51/136, 37.5%). Among the 85 undergraduates, the mix of school class was 15 freshman, 31 sophomores, 18 juniors, and 21 seniors. This profile indicates that these iCBT programs appealed to both undergraduate and graduate students who had a profile for age, gender, and race, which was consistent with the larger college student population in the United States [4].

Profile of Program Utilization in the Total Sample

This profile of program use is for the total sample (*N*=951) across the 4 programs. The average number of lessons used was 3.63 (SD=2.08) out of the 8 possible. The mean number of days of use from the start to the last lesson was 41.55 days (SD=43.77), with a median of 25 days and a range of 1 to 183 days. The period of time between the use of each lesson averaged 11.65 days (SD=13.14).

As expected, the number of lessons used and the duration of the period of use were strongly positively correlated (r=0.47; P<.001). The number of lessons used was not related to gender (r=−0.02; P=.54) but was somewhat correlated with age (r=0.14; P<.001), in that older students used more lessons. The duration of use was not related to gender (r=−0.01; P=.98) but was somewhat correlated with age (r=0.11; P<.001), in that older students had a longer period of use.

About 1 in every 5 students chose to involve a *coach* from the program staff for ongoing support during use of the program (209/951, 22.0%). The use of a coach was strongly associated with participating in a greater number of lessons (1.49 more lessons on average than when not using a coach; r=0.30; P<.001) and with greater duration of program use (37.13 more days of use on average than when not using a coach; r=0.35; P<.001). In contrast, the use of a coach was only weakly associated with demographic factors of older age (r=0.08; P=.02) and was unrelated to gender ($X^2_{1,951}$=0.01; P=.76).

About 1 in every 8 students chose to engage a *teammate* (a personal friend/family member) for ongoing support (124/951, 13.0%). The use of a teammate was not associated with the number of lessons used (r=0.05; P=.12), the duration of use (r=0.01; P=.72), or age of the user (r=−0.01; P=.81). However, the use of a teammate was strongly associated with gender ($X^2_{1,951}$=8.2; P<.001), such that females (102/706, 14.4%) were twice as likely to have used a teammate than were males (16/224, 7.1%).

Comparison of the Four Programs Based on User Demographic and Operational Factors

The demographic factors of age, gender, and operational use factors were also compared between the 4 iCBT programs. The results revealed significant differences between programs on 7 of the 9 factors tested (Table 2).

Table 2. Comparison of user demographic and operational factors: The Learn to Live program (N=951).

Factor	iCBT[a] program				Test of differences		*P* value
	Stress, anxiety, and worry	Depression	Social anxiety	Insomnia (sleep)	*F* test (*df*)	Chi-square (*df*)	
Number of users	325	347	203	76	N/A[b]	N/A	N/A
User demographic factors							
User age (years)					3.6 (3,947)	N/A	.01
Median	21	21	20	22			
Mean (SD)	23.24 (5.51)	23.21 (6.78)	23.00 (6.38)	25.72 (9.06)			
User gender, n (%)					N/A	25.5 (6,951)	<.001
Female	360 (80.0)	259 (74.6)	129 (63.5)	58 (76.4)			
Male	63 (19.4)	76 (21.9)	70 (34.5)	15 (29.7)			
Gender diverse	2 (0.6)	12 (3.5)	4 (2.0)	3 (3.9)			
Operational factors of program use							
Comprehensive assessment: Yes, n (%)	286 (88.0)	309 (89.0)	165 (81.3)	69 (90.8)	N/A	8.5 (1,951)	.04
Lessons used, n (%)					2.0 (3,947)	N/A	.12
2	152 (46.8)	152 (43.8)	83 (40.9)	39 (51.3)			
3	52 (16.0)	69 (19.9)	52 (25.6)	9 (11.8)			
4	41 (12.6)	27 (7.8)	31 (15.3)	7 (9.2)			
5	24 (7.4)	20 (5.8)	10 (4.9)	3 (3.9)			
6	11 (3.4)	13 (3.7)	11 (5.4)	2 (2.6)			
7	7 (2.2)	8 (2.3)	5 (2.5)	1 (1.3)			
8	38 (11.7)	58 (16.7)	11 (5.4)	15 (19.7)			
Mean (SD)	3.58 (2.03)	3.80 (2.23)	3.37 (1.69)	3.78 (2.37)			
Number of days					6.0 (3,947)		<.001
Median	30	28	22	16			
Mean (SD)	46.54 (46.82)	44.17 (44.77)	33.34 (38.30)	30.07 (33.83)			
Range	1-182	1-183	1-163	1-143			
Coach used: Yes, n (%)	84 (25.8)	79 (22.8)	33 (16.3)	13 (17.1)	N/A	7.9 (3,951)	.05
Teammate used: Yes, n (%)	42 (12.9)	77 (22.2)	2 (1.0)	3 (3.9)	N/A	57.2 (3,951)	<.001
Multi-user: Yes, n (%)	36 (11.1)	64 (18.4)	28 (13.8)	17 (22.4)	N/A	10.4 (3,951)	.02
Survey at follow-up: Yes, n (%)	37 (11.4)	54 (15.6)	34 (16.7)	11 (14.5)	N/A	3.7 (3,951)	.30

[a]iCBT: internet-based cognitive behavioral therapy.

[b]N/A: not applicable.

Demographics Compared by Program

Both demographic characteristics of users differed significantly between programs. The average age of the students in the insomnia program was about 3 years older than the average age of the students in each of the other 3 programs. Although females were the majority of users in every program, the social anxiety program had relatively fewer women among users (64%) than the gender mix in the other 3 programs (range 75%-80%; female).

Operational Factors Compared by Program

The 4 programs were similar in the average number of lessons used per person. The percentage of users who completed a valid

survey at follow-up after use was also similar across the 4 programs. However, significant differences between programs were present in 5 other utilization factors (Table 2). The percentage of users who had completed the comprehensive assessment was lower in the social anxiety program than in the other 3 programs (81% vs 88%-91%). The average period of use differed between programs, with the stress, anxiety, and worry and depression programs both having longer average periods of use (roughly 12 days more) than the social anxiety and insomnia programs. The stress, anxiety, and worry program and the depression program both had a higher coaching use rate than the social anxiety and insomnia programs (26%, 23% vs 16%, 17%, respectively). The depression and stress, anxiety, and worry programs both had a much higher use of teammates for support than did the insomnia and social anxiety programs (22%, 13% vs 4%, 1%, respectively). The percentage of students who used multiple programs during the 3-year study period differed significantly between programs (range 11%-22%).

Part 2: Improvement in Clinical Symptoms by Program

This part of the results examines the primary outcomes of the study for change in the clinical symptoms among users in each program separately. Changes from before to after program use in the level of severity of clinical symptoms were empirically examined in two ways. The first approach was more clinically focused and determined how many cases changed from being in clinical status at the start (100% of users by design) to no longer being in clinical status after use. The second approach compared the average levels of symptom severity across all cases in each program before and after use. Both approaches were performed separately for each of the 4 programs. See Table 3 for details of the findings.

Table 3. Users at clinical status at start and last use and average level of clinical symptoms at start and at last use: by the Learn to Live program.

Characteristics	Internet-based cognitive behavioral therapy program				
	Stress, anxiety, and worry	Depression	Social anxiety	Insomnia (sleep)	All
Users, n	325	347	203	76	951
Measure	GAD-7[a]	PHQ-9[b]	SPIN-17[c]	MOS-Sleep-6[d]	N/A[e]
Score range	0-21	0-27	0-51	0-100	N/A
Clinical level score	10+	10+	31+	44+	N/A
Clinical status at start, %	100	100	100	100	N/A
Test 1 of reduction in symptom severity at individual user level					
No change: Stayed clinical at last use, n (%)	176 (54.2)	223 (64.3)	158 (77.8)	46 (56.4)	N/A
Changed from clinical at start to subclinical at last use, n (%)	149 (45.8)	124 (35.7)	45 (22.2)	33 (43.4)	Unweighted average: 36.8%
Test 2 of reduction in symptom severity level average across all users					
Start, mean (SD)	14.50 (3.09)	16.44 (4.12)	43.89 (8.55)	60.46 (11.32)	N/A
Last use, mean (SD)	10.19 (5.02)	12.34 (6.41)	37.06 (13.31)	47.03 (18.43)	N/A
Improvement, %	29.7	24.9	15.6	22.2	Unweighted average: 23.1%
Paired *t* test (*df*)[f]	16.21 (324)	12.71 (346)	8.33 (202)	6.85 (75)	N/A
Paired correlation, *r*	0.38	0.41	0.50	0.42	N/A
Effect size, Cohen *d*	1.25	0.91	0.80	1.10	N/A
Effect size level	Large	Large	Large	Large	N/A

[a]GAD-7: Generalized Anxiety Disorder 7-item scale.

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

[c]SPIN-17: Social Phobia Inventory 17-item scale.

[d]MOS-Sleep-6: Medical Outcomes Study Sleep 6-item scale.

[e]N/A: not applicable.

[f]All *P* values <.001.

Change in Clinical Status by Program

For each program, the number of cases that were still above the established scale score cutoffs for *clinical* status at the last lesson used was examined. The results showed that the percentage of users who changed from clinical status at the start to become lower and subclinical in their level of severity after use was 46% for stress, anxiety, and worry, 43% for insomnia, 36% for

depression, and 22% for social anxiety. Averaged across the 4 programs (unweighted by different sample sizes), 37% of the students changed from clinical to non–clinical status after use.

Change in Average Level of Symptom Severity by Program

Paired *t* test (two-tailed probability) and statistical effect size (Cohen *d*) analyses were performed for each of the 4 programs to test the extent of change in clinical symptom severity on average across all users from before to after use. The results indicated that each program had a significant improvement in severity. The average improvement (as a percentage of score reduction) for each program was as follows: 30% for stress, anxiety, and worry, 25% for depression, 22% for insomnia, and 16% for social anxiety. Each of these results had a statistical effect size that was considered as *large* (*d* range 0.80-1.25). Averaged across the 4 programs (unweighted by different sample sizes), there was a 23% reduction in the level of severity of clinical symptoms after use.

Part 3: Multivariate Tests of Moderating Factors of the Extent of Improvement in Clinical Symptoms After Program Use in Total Sample

The findings in Part 1 revealed many differences between the 4 programs in terms of user characteristics and how the programs were used. Overall, there were also some correlations between many of the operational factors. This context indicated that the overall results found in Part 2 for clinical improvement specific to each program may be influenced by user demographics and operational factors. Given the different profiles of the 4 programs, it made sense then for us to explore changes over time in the clinical outcomes using other more sophisticated tests that take into account the joint influences of the user demographic and program use characteristics. This part of the results examined the impact of relevant potential moderator factors on the change in clinical outcomes in the full sample, considering all programs together.

As each of the programs had different outcome measures and a standardized measure was needed for analyses involving all programs in the same tests in the total sample, we used the difference scores for the change in symptom severity measures for each program (Multimedia Appendix 2). These difference scores when converted into percentages of change from first to last use ranged from 100% to −100% and had a near normal distribution of variance for each program such that some users had increased severity, some had little or no change, and some had decreased severity. The variation between the 4 programs in the extent of change in clinical outcome severity was statistically tested when controlling for all of the other user demographic and operational factors. The same tests also explored whether each of the demographic and operational factors were moderating the results for the extent of improvement after use, when controlling for the shared effects of the program used and all other factors. For example, was there a greater reduction in clinical symptoms for students who used more lessons in a program, all other factors being the same?

A repeated measures analysis of variance model with covariates (ANCOVA) was conducted with the dependent measure of the difference score for the amount of change in symptom severity from start to last use. The factors included the program (4 groups), the covariates of the demographic factors of user age (number of years) and gender (male, female, and gender diverse), the operational factors of use of the comprehensive assessment (yes/no), the number of lessons used (2-8), the time period of use (number of days), the use of a coach (yes/no), the use of a teammate (yes/no), and if the student had used more than one program (yes/no). Each continuous variable was used in the statistical tests as covariates with their full variance. The results for each factor are shown in the second column of Table 4.

However, subgroups of the continuous variables were also created for descriptive purposes to better understand the results using the estimated scores. This recoding process was done for the user, the number of lessons, the user's age, and the duration of use. The number of lessons used was recoded into 3 categories: (1) lessons used 2, (2) lessons used 3-7, or (3) all 8 lessons completed. The age of the student was recoded into 2 categories: (1) those who were within the traditional undergraduate age range of 21 years or younger (*n*=534; mean=19.41 years, SD=1.11; median=19) and (2) those who were older and in the age range of 22 to 62 years (*n*=417; mean=28.46 years, SD=7.07; median=27).

The time period of use was recoded into 3 categories, based on dividing the sample into thirds (which was done separately within each program): (1) shortest time period, (2) middle time period, or (3) longest time period. As expected, these 3 groups differed in the total number of days of program use: (1) shortest time period (*n*=299; mean=5.20 days, SD=4.11; median=4 days), (2) middle time period (*n*=330; mean=26.13 days, SD=11.17; median=24 days), and (3) longest time period (*n*=322; mean=91.11 days, SD=39.83; median=79 days). Essentially, this yielded 3 groups that used the program for time periods that lasted approximately 1 week, 4 weeks, and 13 weeks.

Adjusted mean scores on the dependent measure of the percentage change in clinical symptoms were then calculated for each subgroup within a factor. This process yielded estimated scores for the clinical change outcome measure for subgroups of a particular factor, while at the same time controlling for the influence of all of the other factors. The detailed descriptive results for adjusted percentage change levels for each subgroup are presented in the third and fourth columns of Table 4.

Table 4. Main effect tests of each factor on the outcome of reduction in clinical symptoms when controlling for other factors: total sample (N=951).

Factor	Results with all factors in the model[a]		Improvement from start to last use in clinical symptoms by subgroups of factor (%)	Sample, n	Relative odds for pairings of subgroups with most difference
	F test (*df*)	*P* value			
Overall	11.37 (11,950)	<.001			
All program average			23.9	951	N/A[b]
Comprehensive assessment	<1 (1,950)	.81			Yes 1.04 X greater than No
Yes			22.9	829	
No			22.0	122	
iCBT[c] program used	7.42 (3,950)	<.001			Stress, anxiety, and worry program 1.83 X greater than social anxiety program
Stress, anxiety, and worry			29.3	325	
Depression			24.1	347	
Insomnia			21.2	76	
Social anxiety			16.0	203	
Age of user[d] (years)	3.01 (1,950)	.08			Older 1.14 X greater than college age
Older age of 22+ years			25.7	417	
College age			22.5	534	
Gender of user	3.42 (2,950)	.03			Male 1.40 X greater than average of female and gender diverse
Male			28.7	224	
Female			21.0	706	
Gender diverse			19.9	21	
Number of lessons used[d]	43.56 (1,950)	<.001			All 8 lessons 2.41 X greater than 2 lessons
All 8 lessons			43.1	122	
3-7 lessons			24.5	403	
2 lessons			17.9	426	
Duration (days of use)[d]	<1 (1,950)	.66			Longer 1.13 X greater than shorter
Longer: 13 weeks			24.6	322	
Middle: 3 weeks			25.3	330	
Shorter: 1 week			21.7	299	
Coach used	2.90 (1,950)	.09			Coach yes 1.20 X greater than no
Yes			27.5	209	
No			22.9	742	
Teammate used (*SAW or DEP program users only*)	2.11 (1,671)	.15			Teammate 1.20 X greater than no
Yes			30.8	119	
No			25.7	553	
Live support combined use	4.53 (2,950)	.01			Both 1.70 X greater than neither
Both coach and teammate			37.7	33	
One live support			26.4	267	

Factor	Results with all factors in the model[a]		Improvement from start to last use in clinical symptoms by subgroups of factor (%)	Sample, n	Relative odds for pairings of subgroups with most difference
	F test (*df*)	*P* value			
Neither used			22.2	651	
Multiple program user	1.05 (1,950)	.31			Multiuser 0.86 X less than no
Used 2+ programs			20.1	145	
No			23.5	806	

[a]Mean percentages of change for subgroup estimated after statistical adjustment for all other factors listed. Test for teammate use was smaller as only 2 programs had enough participants with the use of a teammate to qualify (N=672).

[b]N/A: not applicable.

[c]iCBT: internet-based cognitive behavioral therapy.

[d]Tested in an analysis of variance model as a covariate using the full range of continuous data available. However, the subcategories shown have estimated mean scores for descriptive purposes only.

Outcomes Overall

Overall, when adjusted for the program and all other factors, the typical college student user experienced a 24% reduction in the severity of clinical symptoms. This overall level of reduction in clinical outcomes was moderated to a significant degree ($P<.05$) by 3 factors. These 3 factors included the program topic, the number of lessons used, and the gender of the user.

Outcome Differences by Program

The programs significantly differed from each other in terms of the change in outcome level, all other factors being equal. The stress, anxiety, & worry program had the highest level of improvement in clinical symptoms per average user (29%), followed by the depression program (24%), the insomnia program (21%), and the social anxiety program (16%) as the lowest. Thus, the program effectiveness differed between the 4 programs by a range of 16% to 29%.

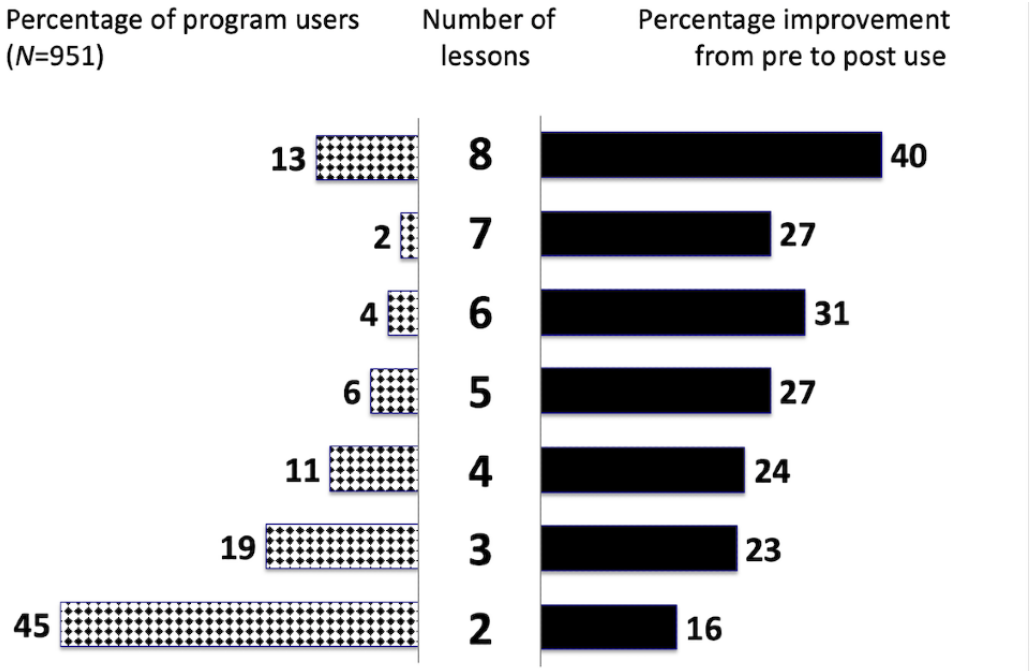
A comparison of these adjusted results from the multivariate tests with the unadjusted results (Table 3) indicated very small differences between the 2 test methods (ie, change in the percentage of users at clinical status after use vs. change in the average level of symptom severity for all students after use). For the stress, anxiety, and worry program, 29.7% versus 29.3% equals a raw difference of 0.4 and a relative difference of 1.5%. For the depression program, 24.9% versus 24.1% equals a raw difference of 0.8 and a relative difference of 3.2%. For the insomnia program, 22.2% versus 21.2% equals a raw difference of 1.0 and a relative difference of 4.5%. For the social anxiety program, 15.6% versus 16.0% equals a raw difference of 0.4 and a relative difference of 2.6%. Thus, the overall results specific to each iCBT program did not change significantly, when taking into consideration the influences from the set of other relevant user and use context factors as well. This small difference between the testing methods also suggests that the other factors were acting in similar ways within each of the programs in how they affected the outcomes. However, this interpretation is examined directly in Part 4 of the results.

Outcome Differences Moderated by the Number of Lessons Used

The largest difference among the covariates was found for the operational factor of the number of lessons used. After controlling for which program was used and for all of the other operational and demographic factors, the students who completed the full program had an average improvement of 43%, compared with a 25% improvement among the students who used between 3 and 7 lessons, and only an 18% improvement for the students who used just 2 lessons. This was a relative difference such that the clinical success rate was 2.4 times higher for program completers than for those who participated minimally in doing only the first 2 lessons. This factor represents the level of dosage of the full content and intervention material. As this was the largest effect of all of the moderator factors tested, it was explored in more detail at each of the 7 specific levels of the number of lessons used (Multimedia Appendix 3).

The general pattern of how the number of lessons used was associated with clinical change and distribution in the total sample is presented in Figure 1. This graphic shows 2 bar graphs side by side and sharing a center column for the number of lessons used from a low of 2 to a high of 8. The bar chart on the left side shows how many students in the total sample used each total number of lessons (as a percentage). This figure reveals that most of the students (45%) had used 2 lessons, with the percentage of students in the middle levels of the number of lessons (from 3 to 7) being between 2% and 19% of all users, and only 13% of students who had completed all 8 lessons. The other bar chart on the right side shows the average percentage of improvement over time in clinical symptoms at each level of number of lessons used. The lowest level of improvement at 16% was for the students who used the fewest lessons (2), whereas the highest level of improvement at 40% was for the students who had used the most lessons (8), with the levels of improvement being in between these low and high ranges for the other students who had completed a number of lessons in between these extreme groups (3-7 lessons used). Thus, the impact on the extent of change in clinical symptoms was greater in an almost linear fashion as the number of lessons increased.

Figure 1. Results in total sample across internet-based cognitive behavioral therapy programs for moderating effect of greater number of lessons used (center) associated with greater improvement in clinical symptom severity (right side) and percentage of users at each level of lessons (left side).



Outcomes Moderated by the Demographic Characteristics of the User

The gender of the user was significant as a covariate when controlling for all other factors (Table 4). Males had a higher level of improvement (29%) than did either females or gender diverse students (21% and 20%, respectively). When tested more specifically, with just 2 groups of males compared with the combination of females and gender diverse students as one group, controlling for all other factors, the result was even stronger for males, having a greater improvement by 1.4 ($F_{1,950}$=12.95; *P*<.001). Older students had slightly more clinical improvement after use than did the traditional college-age students (26% vs 23%; *P*=.08).

Outcomes Moderated by the Use of Live Support

The use of live support only approached significance as a moderator when controlling for other factors (Table 4). Students who used a live coach from the service for support had slightly more clinical improvement after use than did the students who did not (28% vs 23%; *P*=.09). When tested combing the data from the only in the 2 programs with enough users of a teammate to provide a fair test (ie, the stress, anxiety, and worry and the depression programs), the students who had used a friend or family member for support had slightly more clinical improvement after use than the students who did not (31% vs 26%; *P*=.15).

However, when the 2 live support options were categorized into a combined variable of 3 groups (users of both coaching and teammates: 3.5% of all users; users of either kind of support: 28.0%; and nonusers of live supports: 68.5%), the findings were significant (*P*<.001) when controlling for all other covariate factors. These results indicated that the users of both live supports had the most improvement (38%), followed by the users of one of the live supports (26%), and with the non-users of live supports users having the least amount of improvement (22%).

Outcomes Moderated by Other Factors

Other operational factors did not significantly impact the extent of improvement in clinical symptoms once all other factors were considered (Table 4). These factors with no moderator effects included the duration of use, optional use of the comprehensive assessment before starting the first lesson, and whether or not the student had used multiple iCBT programs.

An explanation for why the duration of use was not significant is that it was positively correlated with the much stronger factor of the number of lessons used and thus it had minimal impact when the number of lessons factor was also included in the same analysis. The lack of effect for the comprehensive assessment is perhaps due to the timing of completing the comprehensive assessment, which is often done on the same day as starting the first lesson. Thus, the first lesson and the comprehensive assessment tend to have very similar scores on the clinical symptom measure used as the starting score.

Using more of the lessons on outcomes was critical for improvement. The lack of effect for the factor of multiple program user status makes sense when the average number of lessons *per program used* was lower by about 1 lesson among multi-program users than for single-program users (n=145, mean=2.14, SD=2.57 vs n=806, mean=3.44, SD=1.9, respectively).

In addition, both the comprehensive assessment and multi-program user factors had imbalanced distributions, with the vast majority of users being in 1 of the 2 categories (n=829 vs n= 122; n=806 vs n=145, respectively). Nonequivalence of sample sizes between groups can also affect the fidelity of the comparison tests.

Part 4: Multivariate Tests Within Each Program of Moderating Factors of the Extent of Improvement in Clinical Symptoms After Program Use

Finally, additional repeated measures analysis of variance model tests with covariates were conducted to determine if the demographic and operational factors had moderator effects that were similar within each of the specific programs concerning their influence on the changes in the average level of symptom severity. The insomnia program, however, was excluded because it had too few users to allow us to conduct reliable tests of scores in the various subgroups of the factors. The analysis sample was also slightly smaller in each program tested, as the gender diverse students were excluded as they represented too small a group within each program for a reliable test (gender diverse: n=2 for stress, anxiety, and worry program; n=12 for depression program; n=4 for social anxiety program).

Stress, Anxiety, and Worry Program: Moderator Tests

For the stress, anxiety, & worry program (*N*=323), the results of the statistical analysis found that 2 of the 8 factors were significant as moderators of clinical symptom reduction after use (Multimedia Appendix 4). The effect for the lessons used was the largest of all the factors tested. Students who had used 2 lessons had the lowest level of clinical improvement at 23%, students in the middle levels of the number of lessons used (from 3 to 7) had a 31% improvement, and students who had completed all 8 lessons had the highest level of improvement with 46% reduction in symptoms. This was a relative difference such that the clinical success rate for the group of program completers was 1.9 times higher than the group that used just 2 lessons. The gender of the user was also significant when controlling for all other factors. Males in this program had a higher level of improvement than females (40% vs 27%, respectively). This was a relative difference such that the clinical success rate for males was 1.5 times higher than that for females. Other factors were not significant as moderators of improvement in the stress, anxiety, and worry program.

Depression Program: Moderator Tests

For the depression program (*N*=335), the results of the analysis found that 2 of the 8 factors were significant as moderators of clinical symptom reduction after use (Multimedia Appendix 4). The effect for the lessons used was the largest of all the factors tested. Students who had used 2 lessons had the lowest level of clinical improvement at 18%, students in the middle levels of the number of lessons used had a 24% improvement, and the students who had completed all 8 lessons had a 48% reduction in symptoms. This was a relative difference such that the clinical success rate for the group of program completers was 2.8 times higher than the group that used just 2 lessons. The use of a teammate among the depression program participants had a higher level of improvement than those who did not use a teammate (32% vs 23%, respectively). In addition, both demographic factors (age and gender) approached significance when controlling for all other factors among users of the depression program ($P<.10$). Males in the depression program had a slightly higher level of improvement than females (31% vs 23%, respectively). Older age students in the depression program had a slightly higher level of improvement than younger students (29% vs 22%, respectively). Other factors were not significant as moderators of improvement after use in the depression program.

Social Anxiety Program: Moderator Tests

For the social anxiety program (*N*=199), only one factor was significant in the statistical analysis as a moderator of clinical symptom reduction after use (Multimedia Appendix 4). The gender of the user was significant when controlling for all other factors. Males who addressed social anxiety had a higher level of improvement than females (21% vs 13%, respectively). The effect for the lessons used was in the same direction as in the other programs, yet failed to reach significance, and all of the other factors were not significant as moderators of improvement in the social anxiety program.

In summary, the findings for tests of moderator effects conducted separately within each program had a pattern of results that was mostly similar to the same tests conducted in the full sample across programs. The number of lessons used had the same pattern of findings in each program and was highly significant in 2 of the 3 programs specifically, although not for the social anxiety program (perhaps due to the smaller sample size). For gender, males improved more than females in each program. The effect of age approached significance in the depression and social anxiety programs but not in the stress, anxiety, and worry program. The effect of the use of teammate was found to be significant for the depression program but not for the stress, anxiety, and worry program. Coaching contributed to slightly better outcomes, but was not significant for any of the 3 programs tested. Other operational factors did not affect the outcomes within each program.

Part 5: Survey Outcomes

For the survey data, a descriptive approach guided the data analysis to explore user responses concerning the promotion of the services, experience with counseling, impact on college–related outcomes and attitudes, and overall user satisfaction with the service. The items and responses are presented in Table 5. Students reported becoming aware of the services from a variety of sources within the college environment. The most effective promotional channel was campus email (45%), followed by on-campus printed signage (27%), campus health center (19%), on-campus digital signage (15%), the website of the college/university (13%), and a friend or classmate (4%).

Other questions on the survey asked about the impact of the service use concerning 9 different academic related outcomes. The first set of questions examined positive academic behaviors. The results found that users were more likely after use to ask for help needed for classes (40%), to participate in class (29%), to complete assignments (15%), to make presentations in class (15%), and to go to class on time (12%). When combined, about 8 in every 10 students had one or more of these positive academic outcomes. A second theme involved how various adverse academic outcomes were reduced or avoided after use of the service. Some students reported that they were less likely to drop out of school (20%), were less likely to transfer away (15%), were less at risk of dropping a course (14%), or were

able to avoid academic probation (10%). Considered together, about 1 in every 3 students who used the service had avoided experiencing at least one of these adverse outcomes. Overall, almost 9 in every 10 student users were successful in one or more of the full set of 9 academic outcomes.

Table 5. Results of items on follow-up survey: data averaged across all programs (N=136).

Item	Users, n (%)
Awareness of service: How did you learn about Learn to Live?	
Campus email [a]	61 (44.9)
On-campus printed signage	27 (27.2)
On-campus digital signage	20 (14.7)
College website	18 (13.2)
On-campus health clinic	26 (19.1)
A friend or classmate	6 (4.4)
Academic outcomes: How has Learn to Live affected your college experience (check all that apply)?	
Positive outcomes experienced	
I am more likely to ask for help I need for classes	55 (40.4)
I am more likely to participate in class	39 (28.7)
I am more likely to complete assignments	21 (15.4)
I am more likely to make presentations in class	20 (14.7)
I am more likely to go to class on time	16 (11.8)
Sum: Any one of the above 5 positive outcomes (yes)	108 (79.4)
Adverse outcomes avoided	
I am less at risk of dropping a course	19 (14.0)
I have been able to avoid academic probation	13 (9.6)
I am less likely to drop out of college/university	27 (19.9)
I am less likely to transfer away from this college/university	20 (14.7)
Sum: Any one of the above 4 adverse outcomes avoided (yes)	50 (36.8)
Sum: Any one of the above 9 outcomes (yes)	118 (86.7)
Attitude toward college: Do you now have a more favorable attitude toward your college because they provide Learn to Live as a free benefit?	
Yes	103 (79.4)
No	33 (20.6)
Satisfaction: Overall, how satisfied were you with the Learn to Live experience?	
Very satisfied	28 (27.9)
Somewhat satisfied	81 (59.6)
Somewhat dissatisfied	14 (10.3)
Very dissatisfied	1 (0.7)
Don't know	2 (1.5)

[a]Italics indicate specific response options to the question.

The results of another item indicated that about 8 out of every 10 students had a more favorable attitude toward the school because the school provided the web-based counseling service. Improved academic–related outcomes from the use of the program may have contributed to a better overall opinion about the school. Finally, a key finding of the survey was that almost 9 out of every 10 users were satisfied with their experience with the web-based program.

The results of another item on the survey indicated that about half of the students (74/136, 54.5%), all of whom had met the clinical symptom threshold at the start of use, reported either a current or a past year experience with in-person counseling. Thus, about half of the students who used these digital mental health support tools already had experience with counseling from a live person. In addition, about a third of the users (52/136, 38.3%) were actively engaged in other in-person therapy at the time of their use of the iCBT program. This

finding indicates that the use of the self-directed web-based support tool was adjunctive to in-person counseling for about 1 in every 3 of these students. In addition, among the subgroups of those who reported being currently involved with other therapies, half (26/52, 50.0%) reported that they were *getting more out of it* since adding the web-based resources. Thus, the use of the iCBT tools had a positive effect on their ongoing therapy experience.

The findings presented earlier on the importance of doing more of the lessons to achieve better outcome improvement were informed by responses on the survey for an item that asked why a student did *not* use all of the lessons in a program. This was answered by 74 of the 136 survey respondents. Not having enough time to participate more often was the most common response for the subgroup of users (37/74, 50.0%). About 1 in 6 of these noncompleters indicated that they had improved enough to feel that they could stop before using more lessons (12/74, 16.2%). Another 1 in every 6 of these students did not complete the program because they were not able to relate to the content or did not find the program was helpful (11/74, 14.9%). Taken together, these comments suggest that using fewer than the full 8 lessons was influenced more by a variety of nonsystematic factors unique to the students involved than it was to the program content or functionality.

Part 6: Comparison of College Student and Employee Users—User Profiles and Outcomes

This part of the results examines the replication of this study with another similar study of employee users of the same iCBT resources. The primary findings of this study are similar to those found in a recent study of the same 4 programs when used by employees [26]. The sample that was starting program use above the clinical severity score thresholds was a total of 707 people. The sample sizes, demographics, and program use characteristics for the college and employee user samples are listed for each of the 4 programs in Multimedia Appendix 5. To simplify the comparison, the characteristics and results for the college users and the employee users were examined as simple averages across the 4 programs (ie, unweighted by the sample size differences in programs within each user sample, calculated as scores on a factor added up for the 4 programs and then divided by 4). These full study averages are presented in Multimedia Appendices 5 and 6.

For the demographic characteristics, the two samples were similar for gender (78% females in the employee sample) but, as expected, were quite different in age. The college sample users with an average age in the mid 20s were about 15 years younger than the average age of almost 40 years old for the employee user sample.

The two samples were quite similar in terms of the operational factors (see details in Multimedia Appendix 5). Both samples had most of the participants at the clinical level using the stress, anxiety, and worry or the depression programs more than the other two programs. However, the insomnia program had been used by a higher percentage of employee users than college users (18% vs 8%, respectively). The rate of completion of the comprehensive assessment before starting the program was the same at 87% for both of the samples. The average number of lessons used in the depression program was similar for the two samples, but the employee sample had a slightly higher number of lessons used in the other 3 programs compared with the student sample. Across all 4 programs, the average number of lessons used per program participant was similar, but slightly more for the employee users than the college users (yet this difference between the group averages was quite small - at less than one-half of 1 lesson). The average period of time used in each program was also similar for the two samples in each program. Comparison of the average across all 4 program for the period of use was also very similar, at only 1 day difference between the college and employee users (both at about 40 days).

The use of live support also had some modest group differences within each program (see details in Multimedia Appendix 5). The percentage of users with coach support in each program was greater for employees than it was for college students. The average across programs was 32% with coaches for employees compared to 21% with coaches for college students. The percentage of users with teammate support in each program was more similar for the two groups. The average across programs was 11% with teammates for employees and 10% with teammates for college students.

By design, all users in both samples were above the cutoff scores for clinical status with moderate or higher levels of symptom severity. Depending on the measure, there were 2 or 3 levels of severity within the clinical status range. The data showed that the college and employee samples were also largely similar in the percentage of users at each severity level (Figure 2; with the statistical details in Multimedia Appendix 7).

Considered together, the college student and employee groups were similar in most of the key factors. A summary of the user group averages is displayed in Figure 3. The source statistics for Figure 3 are provided in Multimedia Appendices 5 and 6. This high level of similarity offered a fair context to compare the clinical outcomes for each iCBT program for college and employee users.

The first outcome was how many of the users had changed from being at clinical status at the start of use to being below the threshold and no longer at clinical status after use. On average across the 4 programs (unweighted mean), 36.8% of the college users had this outcome compared to 46.8% of the employee users.

Next, the outcome of the average amount of change across all users in each sample was compared. The relative level of reduction in clinical symptoms for the typical user (as a percentage change from the mean score at the start to the mean score at the last use) was generally similar in each of the iCBT programs for the two samples. However, the employee users had a higher level of symptom improvement than the college users within each program. When averaged across the 4 programs (unweighted mean), the employee sample had an average of 28.9% reduction in symptom severity per person. This same metric was a 23.1% reduction for the college users. The results for clinical change by each program are presented in Multimedia Appendix 6.

Perhaps the levels of the moderating factors of the number of lessons used and the use of coaching, which were both somewhat higher for employee users compared with college users, contributed to the slightly better clinical results for the employee user sample. Both these factors were also significant moderators of clinical improvement when tested in the study of employees involving both clinical and subclinical users [26]. In summary, the comparison of the college user sample to the employee user sample revealed mostly similar profiles of user characteristics, operational use activity, and the primary clinical outcome results.

Figure 2. Comparison of percentage of college users and employee users of the same internet-based cognitive behavioral therapy programs at different levels of symptom severity at pre: by clinical assessment measure specific to each program.

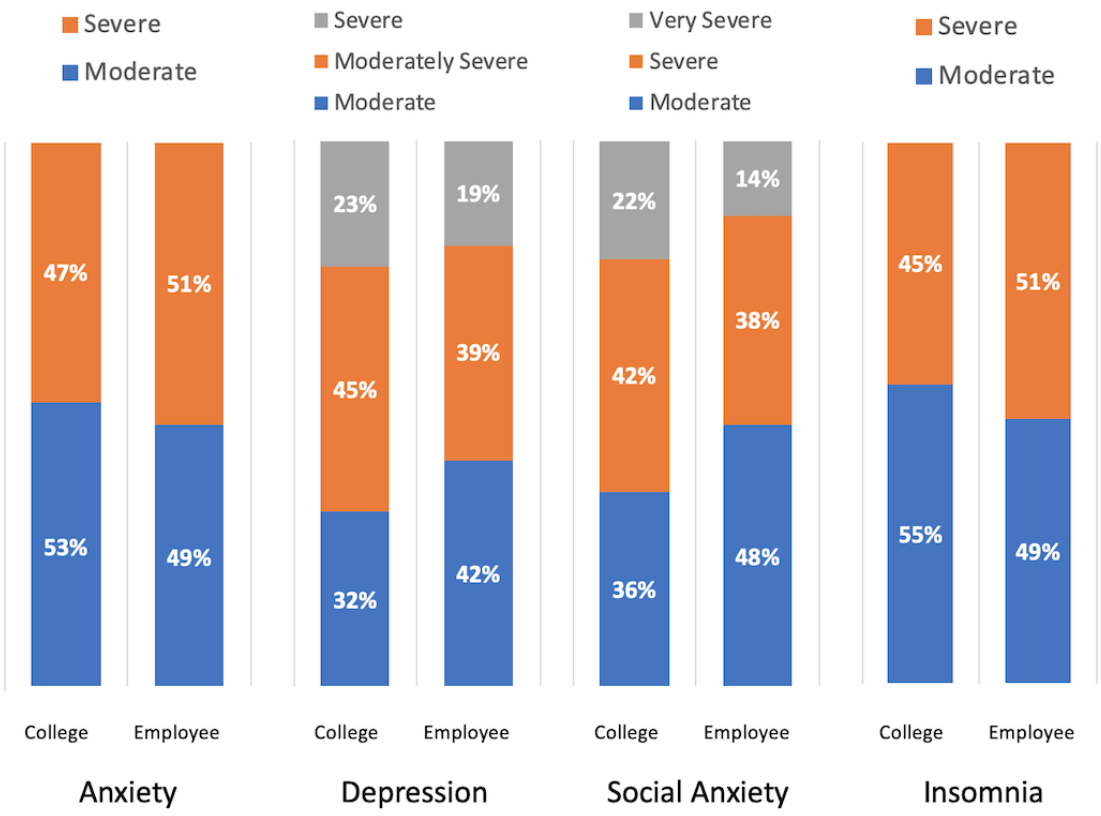
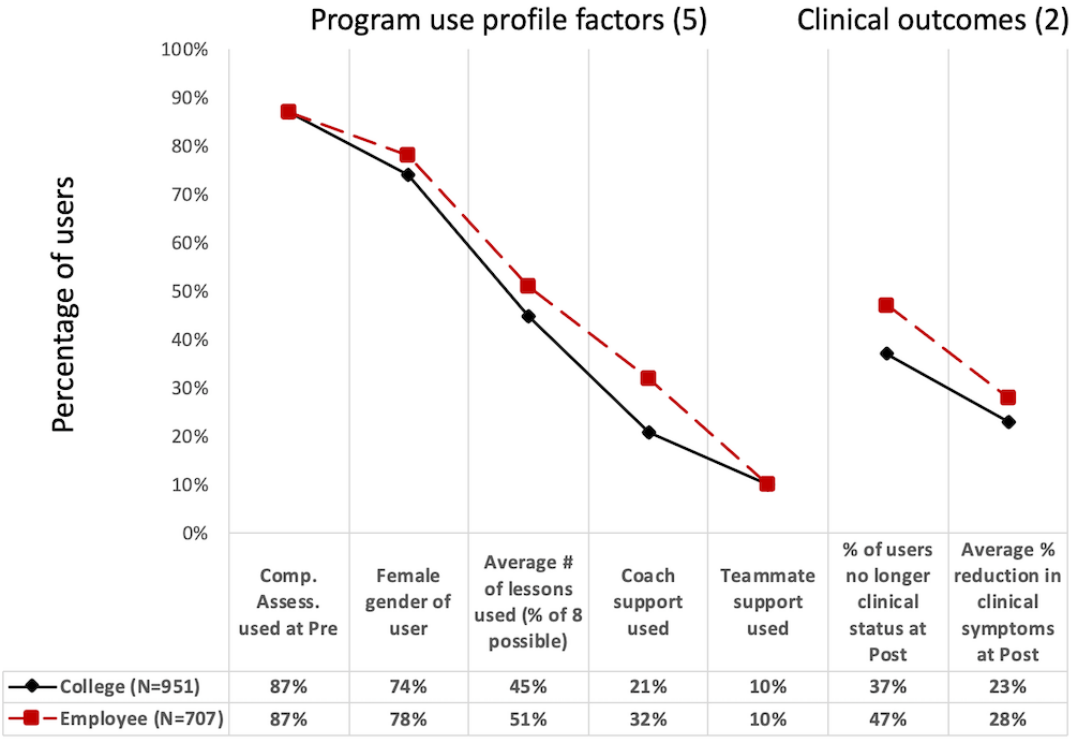


Figure 3. Comparison of college users and employee users of the same internet-based cognitive behavioral therapy programs on context factors and on clinical outcomes: both groups averaged across the 4 programs.



	Comp. Assess. used at Pre	Female gender of user	Average # of lessons used (% of 8 possible)	Coach support used	Teammate support used	% of users no longer clinical status at Post	Average % reduction in clinical symptoms at Post
College (N=951)	87%	74%	45%	21%	10%	37%	23%
Employee (N=707)	87%	78%	51%	32%	10%	47%	28%

Discussion

Principal Findings

This study examined how different technology-assisted support tools can reduce symptoms associated with anxiety, depression, social anxiety, and insomnia. All 4 of these programs had significant reductions in clinical symptom severity after use, and each program had results of a large statistical effect. When averaged across the programs and also controlling for relevant demographic and operational factors, the typical user had a 24% reduction in symptom severity. At the individual level, about 1 in every 3 users changed from being in clinical status at the start of use to being at a nonclinical level of severity after use. Our findings replicate the generally positive results found in other studies of college student users of behavioral health digital tools on clinical symptoms [19-21,25].

Therapeutic *dosage* emerged as a particularly potent moderating variable, with much better results for those who participated fully in each program by completing all 8 lessons compared to those who participated minimally. Among program completers, the extent of improvement was *more than double* that of those who used only 2 lessons. This beneficial effect of adherence to the intervention (through completing more of the intended lessons) is a finding consistent with other research on adherence conducted in tests of similar kinds of iCBT tools for mental health issues [24,39-41].

Better clinical improvement among male users than among female users emerged as a surprise finding, even though less than a third of all users identified as male. One explanation could be the anonymity that this computerized medium affords and that there is no requirement to talk about one's mental health challenges with another person. In this way, male users (especially those with social anxiety issues) may have felt less socially exposed by not having to talk with a counselor to better understand the nature of social anxiety. Male users may also be drawn more to the CBT oriented approach than to emotion-focused coping strategies, which are often emphasized in traditional talk therapy and which male college students tend not to prefer [42]. Gender differences in use rates and the effectiveness of technology-based support tools are also being explored in other research [43-45].

The use of optional live support was associated with an increased level of engagement in the key aspects of participation and yet was only weakly associated with better improvement in clinical symptoms, once these other more impactful factors were also considered in the same tests. In the total sample, however, the combination of using both live supports at the same time was significantly greater than using just one type of support or none at all. Coaching tended to be used more often and had a larger effect on the clinical outcome improvement than the options of live peer support from friends or family. The survey also asked participants to comment on their experience with using a coach. Two themes emerged from these qualitative comments as to why it was helpful. First, the coach provided accountability for engaging in more lessons and over a longer period of time. Second, the coach provided a caring, personalized form of support. The positive role of live coach supports found in this study has also been found in past research on technology tools for depression [46].

Optional peer supports of a friend or family member as teammates were used far less often than coaching support (and almost not at all in two programs). Note that coaches were also used more than teammates in the study of employee users of

the same programs with the same pattern of very low use of teammates for the social anxiety and insomnia programs (see details in Multimedia Appendix 5). Teammate use was not associated with increased operational engagement. However, when a teammate was used, support from a personal contact had a small positive impact on the clinical change outcomes for the users in the depression program (but not in the other program with enough users of teammates to test). One potential explanation could be the preference among some college students to use social support from peers (ie, teammates in this study) when mental health issues are framed more as enhancing well-being and normal life challenges rather than as assisting in the treatment of clinical issues [47].

The survey data from the same users also revealed high levels of satisfaction and more favorable attitudes toward college or university. Program participation was also positively associated with at least one school performance outcome for 9 out of 10 students. About half of the students who used the web-based programs also had past year or concurrent experience in utilizing a face-to-face counselor. Moreover, one-fifth of the student users were referred to the digital services by the university counseling center, reflecting the acceptance of the digital services by the counseling centers and some existing integration of in-person and digital services. The benefits of using both traditional in-person and technological clinical supports have emerging research support [48].

The results of this study replicate the findings obtained from a recent study of employee users of the same set of web-based programs from Learn to Live [26]. The results of this study are similar to the findings in recent studies from other commercial providers of digital technologies designed to support common mental health and insomnia concerns among adults [25,27,39,49,50].

Limitations

As in all applied research projects, there are certain limitations to this study. It was conducted using samples of college and university student users from multiple schools who had voluntarily used one particular commercial service. It is unknown whether these findings with this suite of tools can be replicated in other contexts of college student users. There were no comparison groups to assess the relative benefit from these iCBT tools compared to a matched group of nonusers. It is likely that some level of symptom reduction would occur naturally over time or form other causal forces not measured in this study. If so, whether the improvements over time in each program would have large size statistical effects if comparison groups of nonusers had been included in the study design is unknown. In addition, the causal mechanisms of how these web-based tools contribute to clinical improvements need further examination under more rigorous experimental study design conditions.

Conclusions

This paper adds to the sophistication of the existing literature by comparing 4 distinct clinical topics, all of which shared the same digital platform and similar interactive website tools. Considered together, the variability in the number of lessons and in the total time period of use lends evidence to the flexibility that iCBT programs have to accommodate individual student preferences concerning when, how often, and how much to use the lessons. These iCBT tools appear to be associated with improved academic functioning and for some students, even the ability to stay in school. Additional study of outcomes beyond the typical focus on clinical symptoms in future studies would potentially add to the overall value proposition for web-based tools in support of the mental health of college students.

Acknowledgments

The authors wish to thank Brandon Densmore of Learn to Live for his assistance in preparing the raw data used in this study and Erica Schmit, Learn to Live, for her assistance in managing the project. Funding for Dr Attridge was provided by Learn to Live for this research.

Conflicts of Interest

MA was hired by Learn to Live as an independent consultant for this research project. ER and RM are senior executives employed by the Learn to Live company. DR has no conflicts of interest.

Multimedia Appendix 1
Comparison of survey and non-survey subsamples on demographic, utilization and clinical outcomes: Total sample across programs.
[DOCX File , 77 KB - formative_v4i7e17712_app1.docx]

Multimedia Appendix 2
Psychometric profile of difference scores of change in severity of clinical symptoms from Pre to Post use: by Learn To Live program.
[DOCX File , 77 KB - formative_v4i7e17712_app2.docx]

Multimedia Appendix 3

Source statistics for Figure 1. Percentage of program users and improvement in clinical outcome, both by number of lessons.
[DOCX File , 75 KB - formative_v4i7e17712_app3.docx]

Multimedia Appendix 4
ANCOVA tests of moderator effect for demographic and operational factors on outcome of percentage reduction in clinical symptoms from pre to post (when also controlling for other factors): for three programs.
[DOCX File , 83 KB - formative_v4i7e17712_app4.docx]

Multimedia Appendix 5
Comparison of college student users (this study) and employee users (other study) on age, gender, and operational use factors: by Learn To Live program.
[DOCX File , 78 KB - formative_v4i7e17712_app5.docx]

Multimedia Appendix 6
Comparison of college student users (this study) and employee users (other study) on clinical symptom severity groups and change in severity from Pre to Post use: by Learn To Live program.
[DOCX File , 77 KB - formative_v4i7e17712_app6.docx]

Multimedia Appendix 7
Source statistics for Figure 2 - Severity levels at start of use by program.
[DOCX File , 80 KB - formative_v4i7e17712_app7.docx]

References

1. Douce L, Keeling R. A Strategic Primer on College Student Mental Health. The National Association of Student Personnel Administrators. 2014. URL: https://www.naspa.org/images/uploads/main/Campus_Mental_Health_Primer_web_final.pdf [accessed 2020-07-09]
2. Steger MF, Kashdan TB. Depression and everyday social activity, belonging, and well-being. J Couns Psychol 2009 Apr;56(2):289-300. [FREE Full text] [doi: 10.1037/a0015416] [Medline: 20428460]
3. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. Arch Gen Psychiatry 2005 Jun;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
4. Reference Group Data Report. American College Health Association (ACHA). 2019. URL: https://www.acha.org/documents/ncha/NCHA-III_FALL_2019_REFERENCE_GROUP_DATA_REPORT.pdf [accessed 2020-06-16]
5. Flynn AM, Li Y, Sánchez B. The mental health status of law students: implications for college counselors. J Coll Couns 2019 Apr 4;22(1):2-12. [doi: 10.1002/jocc.12110]
6. Graduate/Professional Student Reference Group Executive Summary. American College Health Association (ACHA). 2019. URL: https://www.acha.org/documents/ncha/NCHA-III_Fall_2019_Graduate_Reference_Group_Executive_Summary.pdf [accessed 2020-06-16]
7. Undergraduate Student Reference Group Executive Summary. American College Health Association (ACHA). 2019. URL: https://www.acha.org/documents/ncha/NCHA-III_Fall_2019_Undergraduate_Reference_Group_Executive_Summary.pdf [accessed 2020-06-30]
8. Prichard J. The Importance of College Sleep: Connections to Mental and Physical Health, Retention, Academic Success Depression on College Campuses. In: Society for Adolescent Medicine Conference. 2019 Presented at: SAHM'19; March 10-13, 2019; Ann Arbor, MI p. -.
9. Frank E. Treating Bipolar Disorder: A Clinician's Guide to Interpersonal and Social Rhythm Therapy. New York, USA: Guilford Press; 2005.
10. Walker MP, van der Helm E. Overnight therapy? The role of sleep in emotional brain processing. Psychol Bull 2009 Sep;135(5):731-48. [FREE Full text] [doi: 10.1037/a0016570] [Medline: 19702380]
11. Center for College Sleep. University of St Thomas – Minnesota. 2019. URL: https://www.stthomas.edu/collegesleep/ [accessed 2020-07-01]
12. Hysenbegasi A, Hass SL, Rowland CR. The impact of depression on the academic productivity of university students. J Ment Health Policy Econ 2005 Sep;8(3):145-151. [Medline: 16278502]
13. Eisenberg D, Golberstein E, Hunt J. Mental health and academic success in college. B E J Econom Anal Policy 2009;9(1):-[FREE Full text] [doi: 10.2202/1935-1682.2191]
14. Fu M, Cheng AW. College counseling services: meeting today's demands. Psychol Serv 2017 Nov;14(4):403-406. [doi: 10.1037/ser0000219] [Medline: 29120198]
15. New J. Balancing Response and Treatment. Inside Higher Ed. 2017. URL: https://www.insidehighered.com/news/2017/01/13/colleges-struggle-provide-ongoing-treatment-demands-mental-health-services-increases [accessed 2020-07-01]

16. Rickwood D, Bradford S. The role of self-help in the treatment of mild anxiety disorders in young people: an evidence-based review. Psychol Res Behav Manag 2012;5:25-36. [FREE Full text] [doi: 10.2147/PRBM.S23357] [Medline: 22427736]
17. Andersson G. Using the internet to provide cognitive behaviour therapy. Behav Res Ther 2009 Mar;47(3):175-180. [doi: 10.1016/j.brat.2009.01.010] [Medline: 19230862]
18. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010 Dec 30;10:113 [FREE Full text] [doi: 10.1186/1471-244X-10-113] [Medline: 21192795]
19. Regehr C, Glancy D, Pitts A. Interventions to reduce stress in university students: a review and meta-analysis. J Affect Disord 2013 May 15;148(1):1-11. [doi: 10.1016/j.jad.2012.11.026] [Medline: 23246209]
20. Davies EB, Morriss R, Glazebrook C. Computer-delivered and web-based interventions to improve depression, anxiety, and psychological well-being of university students: a systematic review and meta-analysis. J Med Internet Res 2014 May 16;16(5):e130 [FREE Full text] [doi: 10.2196/jmir.3142] [Medline: 24836465]
21. Lattie EG, Adkins EC, Winquist N, Stiles-Shields C, Wafford QE, Graham AK. Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review. J Med Internet Res 2019 Jul 22;21(7):e12869 [FREE Full text] [doi: 10.2196/12869] [Medline: 31333198]
22. Learn to Live. 2019. URL: http://www.learntolive.com [accessed 2020-06-16]
23. Chou S, Chang C. The relationship between academic stress and health status: the moderating role of social support. In: Rønningsbakk L, Wu LT, Sandnes F, editors. Innovative Technologies and Learning. Switzerland: Springer International Publishing; 2019.
24. Johansson R, Andersson G. Internet-based psychological treatments for depression. Expert Rev Neurother 2012 Jul;12(7):861-9; quiz 870. [doi: 10.1586/ern.12.63] [Medline: 22853793]
25. Palacios JE, Richards D, Palmer R, Coudray C, Hofmann SG, Palmieri PA, et al. Supported internet-delivered cognitive behavioral therapy programs for depression, anxiety, and stress in university students: open, non-randomised trial of acceptability, effectiveness, and satisfaction. JMIR Ment Health 2018 Dec 14;5(4):e11467. [FREE Full text] [doi: 10.2196/11467] [Medline: 30552076]
26. Attridge M. Internet-based cognitive-behavioral therapy for employees with anxiety, depression, social phobia, or insomnia: clinical and work outcomes. Sage Open 2020 Mar 23;10(1):215824402091439 [FREE Full text] [doi: 10.1177/2158244020914398]
27. Carpenter J, Crutchley P, Zilca RD, Schwartz HA, Smith LK, Cobb AM, et al. Seeing the 'big' picture: big data methods for exploring relationships between usage, language, and outcome in internet intervention data. J Med Internet Res 2016 Aug 31;18(8):e241 [FREE Full text] [doi: 10.2196/jmir.5725] [Medline: 27580524]
28. APA Ethical Guidelines for Research. American Psychological Association. 2017. URL: http://www.sandplay.org/pdf/APA_Ethical_Guidelines_for_Research.pdf [accessed 2020-05-26]
29. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
30. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 6;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
31. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. Med Care 2008 Mar;46(3):266-274. [doi: 10.1097/MLR.0b013e318160d093] [Medline: 18388841]
32. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-13. [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
33. Martin A, Rief W, Klaiberg A, Braehler E. Validity of the brief patient health questionnaire mood scale (PHQ-9) in the general population. Gen Hosp Psychiatry 2006;28(1):71-77. [doi: 10.1016/j.genhosppsych.2005.07.003] [Medline: 16377369]
34. Connor KM, Davidson JR, Churchill LE, Sherwood A, Foa E, Weisler RH. Psychometric properties of the social phobia inventory (SPIN). New self-rating scale. Br J Psychiatry 2000 Apr;176:379-386. [doi: 10.1192/bjp.176.4.379] [Medline: 10827888]
35. Antony MM, Coons MJ, McCabe RE, Ashbaugh A, Swinson RP. Psychometric properties of the social phobia inventory: further evaluation. Behav Res Ther 2006 Aug;44(8):1177-1185. [doi: 10.1016/j.brat.2005.08.013] [Medline: 16257387]
36. Hays R, Stewart A. Sleep measures. In: Stewart AL, Ware JE, editors. Measuring Functioning and Well-being: The Medical Outcomes Study Approach. Durham, NC: Duke University Press; 1992:235-259.
37. Stewart AL, Hays RD, Ware JE. The MOS short-form general health survey. Reliability and validity in a patient population. Med Care 1988 Jul;26(7):724-735. [doi: 10.1097/00005650-198807000-00007] [Medline: 3393032]
38. IBM SPSS Statistics for Macintosh. IBM Corporation. 2017. URL: https://spss-mac.en.softonic.com/mac [accessed 2020-07-09]
39. Dryman MT, McTeague LM, Olino TM, Heimberg RG. Evaluation of an open-access CBT-based internet program for social anxiety: patterns of use, retention, and outcomes. J Consult Clin Psychol 2017 Oct;85(10):988-999. [doi: 10.1037/ccp0000232] [Medline: 28650193]

40. Karin E, Dear BF, Heller GZ, Crane MF, Titov N. 'Wish you were here': examining characteristics, outcomes, and statistical solutions for missing cases in web-based psychotherapeutic trials. JMIR Ment Health 2018 Apr 19;5(2):e22 [FREE Full text] [doi: 10.2196/mental.8363] [Medline: 29674311]
41. Karyotaki E, Kleiboer A, Smit F, Turner DT, Pastor AM, Andersson G, et al. Predictors of treatment dropout in self-guided web-based interventions for depression: an 'individual patient data' meta-analysis. Psychol Med 2015 Oct;45(13):2717-2726. [doi: 10.1017/S0033291715000665] [Medline: 25881626]
42. Brougham RR, Zail CM, Mendoza CM, Miller JR. Stress, sex differences, and coping strategies among college students. Curr Psychol 2009 Feb 11;28(2):85-97. [doi: 10.1007/s12144-009-9047-0]
43. Ellis LA, Collin P, Hurley PJ, Davenport TA, Burns JM, Hickie IB. Young men's attitudes and behaviour in relation to mental health and technology: implications for the development of online mental health services. BMC Psychiatry 2013 Apr 20;13:119 [FREE Full text] [doi: 10.1186/1471-244X-13-119] [Medline: 23601273]
44. Wang J, Lam RW, Ho K, Attridge M, Lashewicz BM, Patten SB, et al. Preferred features of e-mental health programs for prevention of major depression in male workers: results from a Canadian national survey. J Med Internet Res 2016 Jun 6;18(6):e132 [FREE Full text] [doi: 10.2196/jmir.5685] [Medline: 27267782]
45. Wang J, Patten SB, Lam RW, Attridge M, Ho K, Schmitz N, et al. The effects of an e-mental health program and job coaching on the risk of major depression and productivity in Canadian male workers: protocol for a randomized controlled trial. JMIR Res Protoc 2016 Nov 15;5(4):e218 [FREE Full text] [doi: 10.2196/resprot.6350] [Medline: 27847352]
46. Melling B, Houguet-Pincham T. Online peer support for individuals with depression: a summary of current research and future considerations. Psychiatr Rehabil J 2011;34(3):252-254. [doi: 10.2975/34.3.2011.252.254] [Medline: 21208865]
47. Park SY, Andalibi N, Zou Y, Ambulkar S, Huh-Yoo J. Understanding students' mental well-being challenges on a university campus: interview study. JMIR Form Res 2020 Mar 5;4(3):e15962 [FREE Full text] [doi: 10.2196/15962] [Medline: 32134393]
48. Erbe D, Eichert H, Riper H, Ebert DD. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. J Med Internet Res 2017 Sep 15;19(9):e306 [FREE Full text] [doi: 10.2196/jmir.6588] [Medline: 28916506]
49. Schneider J, Foroushani PS, Grime P, Thornicroft G. Acceptability of online self-help to people with depression: users' views of MoodGYM versus informational websites. J Med Internet Res 2014 Mar 28;16(3):e90 [FREE Full text] [doi: 10.2196/jmir.2871] [Medline: 24681717]
50. Schladweiler K, Hirsch A, Jones E, Snow L. Real-world outcomes associated with a digital self-care behavioral health platform. Ann Clin Res Trials(2) 2017;1:- [FREE Full text]

Abbreviations

CBT: cognitive behavioral therapy
GAD-7: generalized anxiety disorder 7-item scale
iCBT: internet-based cognitive behavioral therapy
MOS: medical outcomes study
MOS-Sleep-6: medical outcomes study sleep 6-item scale
PHQ-9: patient health questionnaire 9-item scale
SPIN: social phobia inventory
SPIN-17: social phobia inventory 17-item scale

Edited by G Eysenbach; submitted 06.01.20; peer-reviewed by J Wang; comments to author 17.03.20; revised version received 29.05.20; accepted 14.06.20; published 23.07.20.

Please cite as:
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;4(7):e17712
URL: http://formative.jmir.org/2020/7/e17712/
doi:10.2196/17712
PMID:32706662

Original Paper

Continuous 7-Month Internet of Things–Based Monitoring of Health Parameters of Pregnant and Postpartum Women: Prospective Observational Feasibility Study

Johanna Saarikko[1], RM, MHSc; Hannakaisa Niela-Vilen[1], RM, PhD; Eeva Ekholm[2,3], MD, PhD; Lotta Hamari[1,4], PT, PhD; Iman Azimi[5], PhD; Pasi Liljeberg[5], PhD; Amir M Rahmani[6], MBA, PhD; Eliisa Löyttyniemi[7], MSc; Anna Axelin[1], RN, PhD

[1]Department of Nursing Science, University of Turku, Turku, Finland

[2]Department of Obstetrics and Gynecology, Turku University Hospital, Turku, Finland

[3]Faculty of Medicine, University of Turku, Turku, Finland

[4]Faculty of Communication Sciences, University of Tampere, Tampere, Finland

[5]Department of Future Technologies, University of Turku, Turku, Finland

[6]School of Nursing and Department of Computer Science, University of California, Irvine, CA, United States

[7]Department of Biostatistics, University of Turku, Turku, Finland

Corresponding Author:
Amir M Rahmani, MBA, PhD
School of Nursing and Department of Computer Science
University of California
106D Berk Hall
Irvine, CA, 92697-3959
United States
Phone: 1 949 824 3590
Email: a.rahmani@uci.edu

Abstract

Background: Monitoring during pregnancy is vital to ensure the mother's and infant's health. Remote continuous monitoring provides health care professionals with significant opportunities to observe health-related parameters in their patients and to detect any pathological signs at an early stage of pregnancy, and may thus partially replace traditional appointments.

Objective: This study aimed to evaluate the feasibility of continuously monitoring the health parameters (physical activity, sleep, and heart rate) of nulliparous women throughout pregnancy and until 1 month postpartum, with a smart wristband and an Internet of Things (IoT)–based monitoring system.

Methods: This prospective observational feasibility study used a convenience sample of 20 nulliparous women from the Hospital District of Southwest Finland. Continuous monitoring of physical activity/step counts, sleep, and heart rate was performed with a smart wristband for 24 hours a day, 7 days a week over 7 months (6 months during pregnancy and 1 month postpartum). The smart wristband was connected to a cloud server. The total number of possible monitoring days during pregnancy weeks 13 to 42 was 203 days and 28 days in the postpartum period.

Results: Valid physical activity data were available for a median of 144 (range 13-188) days (75% of possible monitoring days), and valid sleep data were available for a median of 137 (range 0-184) days (72% of possible monitoring days) per participant during pregnancy. During the postpartum period, a median of 15 (range 0-25) days (54% of possible monitoring days) of valid physical activity data and 16 (range 0-27) days (57% of possible monitoring days) of valid sleep data were available. Physical activity decreased from the second trimester to the third trimester by a mean of 1793 (95% CI 1039-2548) steps per day (P<.001). The decrease continued by a mean of 1339 (95% CI 474-2205) steps to the postpartum period (P=.004). Sleep during pregnancy also decreased from the second trimester to the third trimester by a mean of 20 minutes (95% CI –0.7 to 42 minutes; P=.06) and sleep time shortened an additional 1 hour (95% CI 39 minutes to 1.5 hours) after delivery (P<.001). The mean resting heart rate increased toward the third trimester and returned to the early pregnancy level during the postpartum period.

Conclusions: The smart wristband with IoT technology was a feasible system for collecting representative data on continuous variables of health parameters during pregnancy. Continuous monitoring provides real-time information between scheduled appointments and thus may help target and tailor pregnancy follow-up.

(JMIR Form Res 2020;4(7):e12417) doi:10.2196/12417

KEYWORDS

prenatal care; postnatal care; wearable electronics; biosensing; cloud computing; mHealth; physical activity; sleep; heart rate

Introduction

Monitoring and follow-ups during pregnancy are vital to ensure the health and well-being of pregnant women and their unborn infants. Thus far, monitoring of pregnancy is performed at scheduled appointments by health care professionals in maternity care units [1]. Instead of intermittent measurements tied to time and place, remote and continuous monitoring could provide significant opportunities for health care professionals to observe the health-related parameters of their patients [2] and detect abnormal changes in maternal adaptation to pregnancy and liability to pregnancy complications early. Remote monitoring might support personalized care as maternity care could be tailored according to the received information. A personalized monitoring approach could also enhance a woman's self-management because it makes the woman more aware of her health and might commit her to pregnancy care [3].

The Internet of Things (IoT) provides methods for ubiquitous and continuous maternity monitoring. The IoT is a high-level network of objects (ie, things) that are wirelessly connected to servers to provide efficient and comprehensive services [4,5]. In practice, the pregnant woman wears sensors that monitor health-related parameters, enabling her and health care professionals to track the data through web-based user interfaces anywhere and at any time [5]. Recently, the use of various eHealth applications in pregnancy care has sharply increased—for example, for remote monitoring of blood glucose—and the number of maternity care visits for women with gestational diabetes mellitus has decreased with no differences in maternal or neonatal outcomes noted [6,7]. Further, pulse and blood pressure sensors in a smartphone application have been developed, and the feedback from pregnant women has been positive [8,9].

Physical activity and sleep are significant for a pregnant woman's general well-being and quality of life [10]. Physical activity decreases as pregnancy progresses [11], and sleep disorders are common among pregnant women, especially during the third trimester and during the postpartum period [12]. Both physical activity and sleep are usually measured by subjective self-reports [11-13], which is complicated because self-reports may over or underestimate the duration of sleep or activity [13,14]. On the other hand, device-based monitoring has the potential to increase physical activity because having visibility to one's daily activity is known to be an incentive to increase physical activity [15]. Although smart wristbands are widely utilized to measure activity, sleep, and heart rate, their use in pregnant women is still scarce [2]. Continuous device-based monitoring would provide unique and representative data on the levels and changes of activity and sleep during pregnancy and the postpartum period. Furthermore, heart rate measures could be utilized to evaluate the intensity of the physical activity of pregnant and postpartum women. Continuous monitoring could also be useful to target interventions to those needing them the most and to measure health outcomes systematically.

This study aimed to evaluate the feasibility of continuous monitoring of health parameters (physical activity, sleep, and heart rate) of nulliparous women throughout pregnancy and until one month postpartum with a smart wristband and an IoT-based monitoring system.

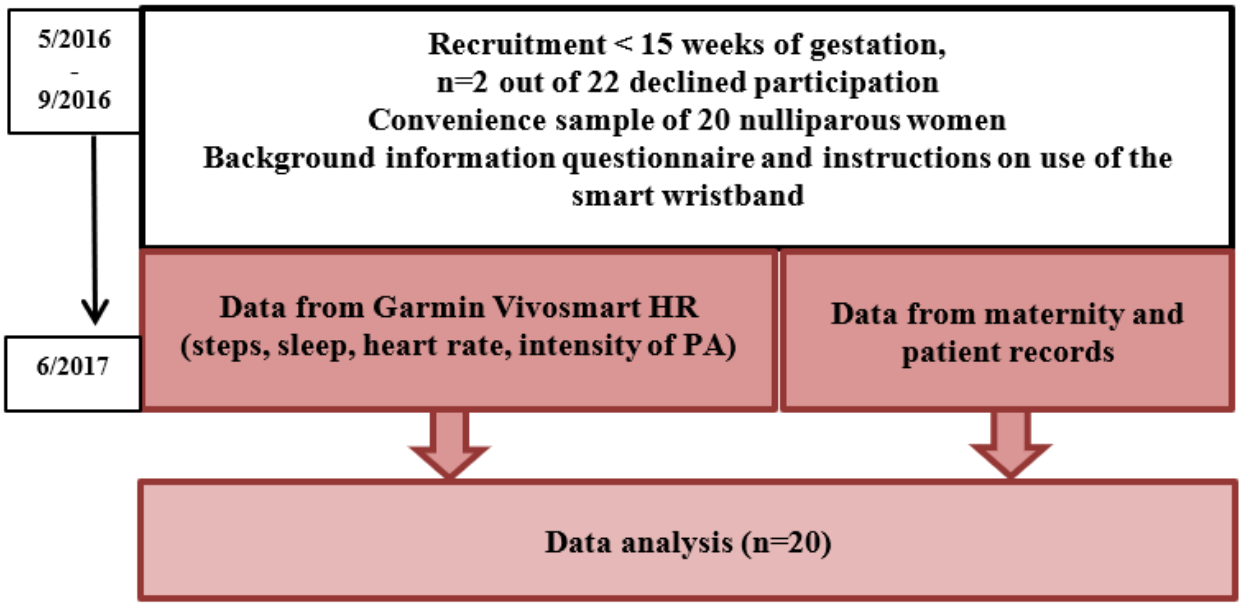
Methods

Study Design and Settings

This study was conducted as a prospective observational feasibility study on nulliparous women attending two maternity outpatient clinics in Southern Finland between May 2016 and June 2017. Physical activity, sleep, and heart rate data were collected with a smart wristband integrating a photoplethysmogram bio-sensor to measure heart rate [16] and an inertial measurement unit to track activity and sleep [17]. The study design is presented in Figure 1.

The study was carried out per the Code of Ethics of the World Medical Association (Declaration of Helsinki) and approved by the Joint Ethics Committee of the Hospital District of Southwest Finland (35/1801/2016) and the university hospital. Permission to use Garmin Vivosmart (HR, Garmin) smart wristbands in this study was obtained from the manufacturer. The study was not registered due to the feasibility design.

Figure 1. The study design.



Participants and Recruitment

In Finland, all pregnant women are offered a screening ultrasound free of charge at the end of their first trimester. A convenience sample of 20 pregnant women was recruited from this visit. Criteria for eligibility were women (1) expecting their first child, (2) ≥18 years of age, and (3) ≤15 of weeks of singleton gestation. Women were excluded if they did not understand Finnish or English or did not have a PC, tablet, or mobile device with which to synchronize data.

The midwives at the maternity outpatient clinics informed the eligible women about the study. After providing oral and written information, the midwives asked permission for the researchers to be in contact with potential participants (N=22). The researchers explained the study purpose and procedures to the women by telephone and scheduled a meeting if a woman was willing to participate. Two women declined participation because they felt that they would not wear the smart wristband. Written informed consent was obtained from all participants at the meeting with the researchers. In addition, a smart wristband and instructions on how to use it were given to each participant, and background information was obtained through a questionnaire.

Outcome Measures and Data Collection

Maternal background characteristics of age, BMI, marital status, education, employment status, smoking, and pre-pregnancy physical activity habits were collected with a questionnaire at the meeting with the researcher. Data on pregnancy and delivery were collected from the maternity card and hospital electronic patient records.

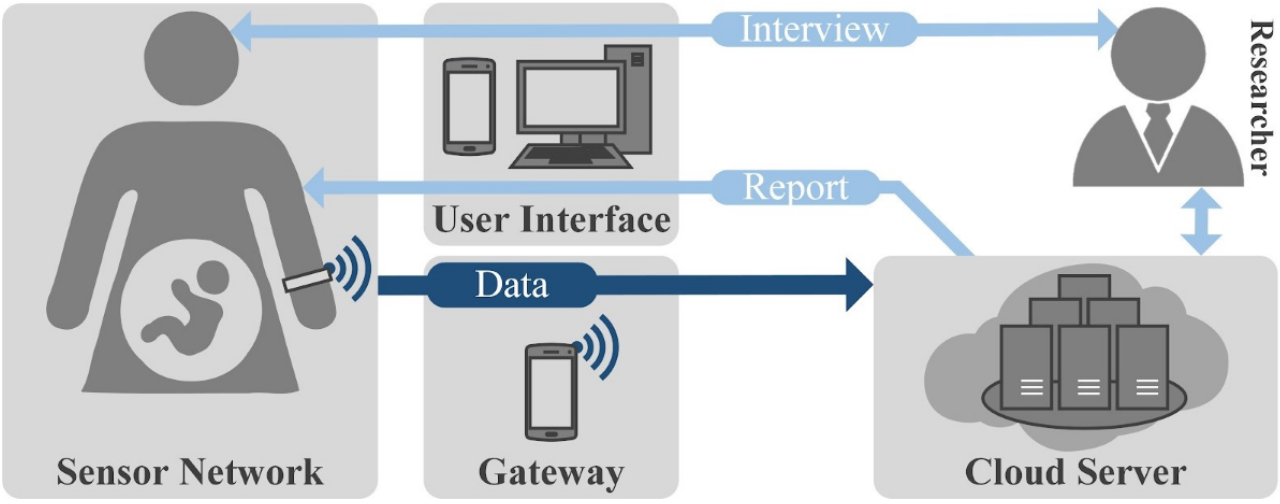
Physical activity, sleep, and heart rate data were collected objectively using a Garmin Vivosmart, small (21 mm × 12.3 mm) and light (29.6 g) smart wristband. The device has shown an acceptable level of validity for step counts tested under laboratory conditions against the Optogait system (OPTOGait, Microgate Srl) and a manual hand counter on the treadmill [18,19]. The device detected sleep automatically based on heart rate and hand movements during regular sleep hours, not during the day. Physical activity intensities were estimated by comparing heart rate data with the data collected by an accelerometer sensor when acceleration was detected, in comparison to the patient's average resting heart rate. If the heart rate sensor was not on, the device calculated the number of steps per minute to evaluate moderate-to-vigorous physical activity (MVPA) minutes. At least 10 consecutive minutes must be recorded to earn intensity minutes so that either step count rate or heart rate was elevated above the predefined moderate-intensity threshold [20]. The women were instructed to wear the activity tracker for 24 hours a day during their pregnancy and for 1 month after delivery, totaling 7 months. They were advised to synchronize the devices once a day or at least when charging the devices every 5 days. Data were accessible for both the researchers and the women.

An IoT-based system was implemented to provide remote health monitoring throughout pregnancy and postpartum. The monitoring system leverages an amalgamation of different sensing, communication, and computing resources to collect, transmit, and analyze data [5,21]. The architecture of this system is illustrated in Figure 2. First, a smart wristband was used to collect health data from the mother remotely. Second, the collected data were transmitted to cloud servers via a gateway device, which was a smartphone or a personal computer. Third, the cloud was responsible for storing the data and for performing the following data preprocessing and analysis methods.

Data collection rates were not fixed throughout monitoring, as the Garmin provided the parameters at different rates. Therefore, we first homogenized the heart rate values by interpolating or averaging them. We then selected a 15-minute interval between successive values. Moreover, step counts were obtained from the Garmin. We leveraged step count values to specify physical activity levels. Sleep duration per night was extracted from the

sleep data provided by Garmin. To validate this data, we carried out a manual cross-check between the sleep data and other data, including hand movement and heart rate values. The sleep data were corrected or removed in cases of mismatch.

Figure 2. An Internet of Things–based system for remote maternity monitoring.



Statistical Data Analysis

Physical activity (step count) and heart rate data were accepted as valid for analyses if the participant wore the device for at least 10 hours per day while awake [22,23] at least 4 days per week [24]. The only exception was the week of delivery, from which all awake data were included in the analyses because the delivery week is shorter in most cases (if woman deliverers at 40+1 gestational week, this week includes only 2 days). Sleep measurements were considered successful if there was no off-wrist time during the sleep period [25] and at least 4 sleep periods measured per week. Sleep duration was defined as the number of minutes scored as sleep by the Garmin algorithm.

To describe the women's background information and monitored health parameters (step counts, sleep and awake minutes, heart rate, and MVPA minutes), means, SD, or CI and medians with ranges were used as continuous variables and counts with proportions for categorical variables. Daily step counts, sleep, and heart rate data from 13 weeks of gestation to 1 month postpartum were presented for the second and third trimesters and the postpartum period (averaging all accepted data during each period). The changes in weekly step counts and sleep minutes were assessed throughout pregnancy. These analyses were performed with a linear mixed model with repeated measures, including one within factor (time = gestational weeks). A compound symmetry covariance structure was used for time. This method allows women with missing values to be included in the analyses. The same method was used to study differences in average step count, sleep, and awake minutes between pregnancy trimesters. Changes over trimesters in MVPA minutes were analyzed using the Friedman test because the normality assumption was not met.

Spearman correlation testing was used to explore associations between mean daily step count by the second and third trimesters and the postpartum period as well as mean daily sleep minutes by the trimesters and postpartum period. MVPA minutes were compared to the recommendation for physical activity in healthy pregnant women: 150 minutes of moderate-intensity or 75 minutes of vigorous-intensity activity per week [26]. All tests were performed as two-sided tests with a significance level set at .05. The analyses were performed using SAS System, version 9.4 for Windows (SAS Institute).

Results

Participants

A total of 20 of 22 eligible pregnant women (refusal rate of 9%) were enrolled in the study. The mean gestational age at recruitment was 12 (SD 2) weeks of gestation. The background characteristics of the women are described in Table 1 and the perinatal outcomes are described in Table 2.

Table 1. Descriptive statistics for maternal background characteristics (N=20).

Characteristic	Participants
Age (years), mean (SD)	26 (5.0)
Gestational age at recruitment (week), mean (SD)	12 (2.1)
Highest educational qualification, n (%)	
Primary education	4 (20)
Secondary education	9 (45)
College or university of applied sciences	4 (20)
University	3 (15)
Marital status, n (%)	
Married or living with a partner	17 (85)
Single	3 (15)
Employment status, n (%)	
Working	13 (65)
Unemployed	2 (10)
Student	5 (25)
Pre-pregnancy BMI (kg/m^2), median (range)	24.4 (17.7–43.5)
Pre-pregnancy smoking, n (%)	7 (35)
Smoking during pregnancy, n (%)	5 (25)
Pre-pregnancy physical activity (almost daily), n (%)	12 (60)

Table 2. Descriptive statistics for perinatal outcomes (N=20).

Perinatal outcomes	Participants
Gestational age at delivery (week), mean (SD)	39.4 (2.6)
Gestational diabetes, n (%)	5 (25)
Means of delivery, n (%)	
Vaginal	14 (70)
Vacuum-assisted	4 (20)
Emergency cesarean section	2 (10)
Birth weight (g), median (range)	3415 (1100-4445)

Physical Activity and Sleep

The recordings yielded valid physical activity data for a median of 144 (range 13-188) days, which represented 75% of the 6-month data collection period during pregnancy. During the postpartum period, valid data were available for a median of 15 (range 0-25) days, representing 54% of the 1-month data collection period.

Physical activity decreased from the second trimester to the postpartum period (Figure 3). The mean daily steps during the second trimester were 6838 (95% CI 5866-7810) and decreased by a mean of 1793 (95% CI 1039-2548) steps per day in the third trimester (P<.001). The daily steps further decreased by a mean of 1339 (95% CI 474-2205) steps in the postpartum period (P=.004). In weekly comparisons, the average daily step count was between 6000 and 7000 from 13-31 gestational weeks. After gestational week 32, the mean daily step count decreased to 5000 by 36 gestational weeks and further to approximately 4000 steps/day after that. The most significant decrease in physical activity occurred at 32 gestational weeks (P<.05 in most weekly comparisons) (Table 3).

Valid sleep data were available from a median of 137 days (range 0-184), representing 72% of data during pregnancy. During the postpartum period, a median of 16 days (range 0-27) of valid data (57%) was used in the analyses (Table 3). Sleep minutes decreased, and nightly awake minutes increased from the second trimester to the postpartum period (Figure 3). The participants slept a mean of 8 hours (95% CI 7.6-8.3 hours) during the second trimester and a mean of 20 minutes (95% CI –0.7 to 42 minutes) less per night in the third trimester (P=.06). The total night sleep time shortened an additional 1 hour (95% CI 39 minutes to 1.5 hours) in the postpartum period (P<.001). In weekly comparisons, the average daily sleep minutes were between 450 and 500 from 13 to 38 gestational weeks. After

gestational week 38, the mean daily sleep minutes decreased to 435 (P<.05 in most weekly comparisons; Table 3).

Figure 3. Mean and standard deviation of the sleep data and step counts per day by week during pregnancy and postpartum (n=2-19).

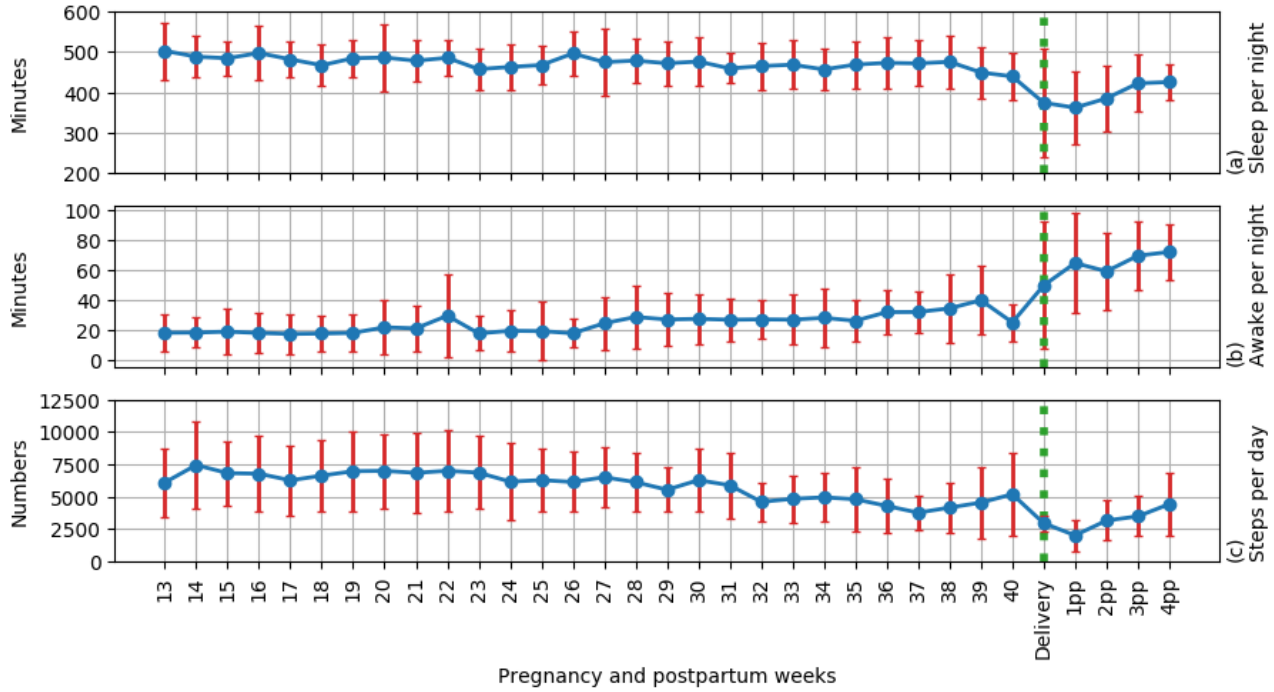


Table 3. Mean step counts and sleep minutes by trimesters and the postpartum period.

Variables	Second trimester (n=19)	Third trimester (n=17)	Postpartum period (n=12)	Overall differences between periods (P value)
Valid physical activity days, n (%[a])	76 (70)	51 (67)	12 (42)	
Steps per day[b]				
Mean (CI)	6838 (5866-7810)	5045 (4049-6041)	3705 (2635-4776)	<.001
Pairwise differences between periods (P value)	<.001[c]	.004[d]	<.001[e]	
Valid sleep days, n (%[a])	69 (63)	48 (63)	12 (41)	
Sleep (min/night)[b]				
Mean (CI)	477 (433-501)	457 (433-481)	393 (367-420)	<.001
Pairwise differences between periods (P value)	.06[c]	<.001[d]	<.001[e]	
Awake (min/night)[b]				
Mean (CI)	21 (15-27)	32 (25-40)	67 (52-81)	<.001
Pairwise differences between periods (P value)	.02[c]	<.001[d]	<.001[e]	

[a]Calculated days during the follow-up period.

[b]Measured by Garmin Vivosmart.

[c]Second trimester versus third trimester.

[d]Third trimester versus postpartum period.

[e]Second trimester versus postpartum period.

Heart Rate and the Intensity of Physical Activity

The resting heart rate increased progressively by 17% from 60 bpm (SD 5) at 13 gestational weeks to 70 bpm (SD 8) at 32 gestational weeks and remained at that level until delivery. The resting heart rate decreased to the early pregnancy level by 4 weeks postpartum (Figure 4).

The intensity of physical activity decreased from the second trimester to the postpartum period. A median of weekly MVPA minutes during the second trimester was 46 (range 0-288) and decreased by a mean of 14 minutes (95% CI –3 to 32 minutes) per week in the third trimester. After delivery, the MVPA minutes further decreased to a median of 15 (range 0-188)

minutes per week, but the change was not significant (*P*=.08; Figure 4). Comparing the MVPA minutes with the general recommendation of physical activity for pregnant women, only 47%, 24%, and 25% of the participants achieved the recommended activity level in at least 1 week in the second and third trimesters and the postpartum period, respectively (Table 4).

Figure 4. Resting heart rate values and moderate-to-vigorous physical activity minutes per week during pregnancy and postpartum (n=3-19). HR: heart rate; MVPA: moderate-to-vigorous physical activity.

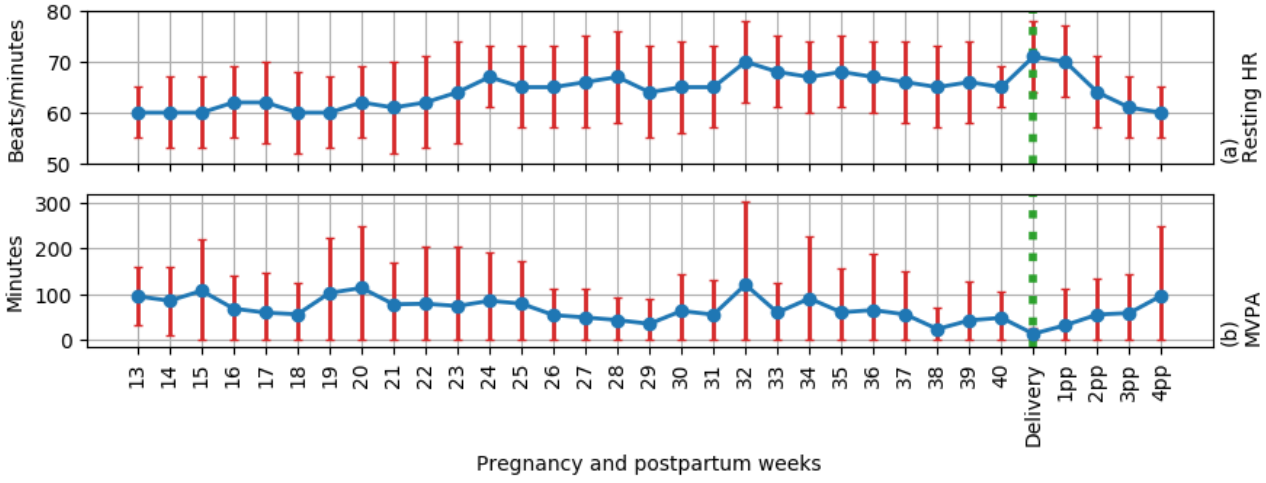


Table 4. Median MVPA minutes by trimester and the postpartum period.

Variables	Second trimester (n=19)	Third trimester (n=17)	Postpartum period (n=12)	Overall difference between periods (*P* value)
MVPA[a] (min/week)[b], median (range)	46 (0-288)	27 (0-279)	15 (0-188)	.08
Participants meeting the recommended MVPA[c], n/N (%)	9/19 (47)	4/17 (24)	3/12 (25)	N/A[d]
Weeks of 150 MVPA minutes, n/N[e] (%[c])	43/239 (18)	19/154 (12)	7/40 (18)	N/A

[a]MVPA: Moderate-to-vigorous physical activity.

[b]Measured by Garmin Vivosmart.

[c]Moderate-to-vigorous physical activity at least in 1 week during follow-up.

[d]N/A: not applicable

[e]Calculated weeks during the follow-up period.

Discussion

Principal Findings

To our knowledge, this is the first continuous, long-term follow-up study measuring physical activity, sleep, and HR during pregnancy and the postpartum period. The results of this study confirmed that this IoT-based system was feasible to monitor health parameters during pregnancy when we collected a large amount of valid physical activity (75% of the data collection period) and sleep (72%) data. After delivery, the amount of data obtained was not representative. Overall physical activity, measured objectively with step counts, was low during pregnancy and postpartum. Sleep time decreased during pregnancy and after delivery, and heart rate increased due to hemodynamic changes. These expected results support the preliminary reliability of a smart wristband with IoT-technology. However, more research is needed to validate the continuous monitoring of pregnant and postpartum women.

Evaluation of the Health Parameters: Physical Activity, Sleep, and Heart Rate

Physical activity (daily steps) decreased during pregnancy and were quite low in the third trimester and especially during the first month after the delivery. Previous studies using intermittent and subjective measurements also showed that physical activity decreases as the pregnancy progresses [27,28], and one previous study using continuous monitoring also found that step counts decrease in inactive pregnant women as pregnancy proceeded [29]. According to a previous review [15], the smart wristband itself might increase participant physical activity. However, the effects of physical activity interventions are often short term. Thus we estimated that this was not relevant in our study due to the long follow-up period. After the delivery, it was not expected that the women's physical activity would have returned to the early pregnancy level during the first month. In addition to the recovery from delivery, the transition to first-time parenthood takes time. A smart wristband could be used to illustrate the level of physical activity to a pregnant woman herself and, by implication, support her in changing her lifestyle to a more active one. Furthermore, physical activity improves

sleep quality during the postpartum period [12,30], and well-rested women are most likely to be physically active; thus, both elements of health should be equally supported.

Our study confirmed previous findings on declining sleep quality and more frequent nocturnal awakenings. In line with the previous study [31], we found sleep quality decline starting from the first trimester. In addition, both the number and duration of nocturnal awakenings seemed to increase as pregnancy proceeded [30] and in the postpartum period [12]. However, the sleep time remained rather high throughout pregnancy and after delivery. Pregnant women's abnormal sleep duration is associated with several maternal problems, such as hypertensive disorders [32], increased body mass index after delivery [33], and depression and anxiety [34]. Moreover, extreme sleep duration during pregnancy has been associated with gestational diabetes [35]. Although the sleep duration of the study participants was mostly sufficient according to the present recommendations, the awake minutes during sleep periods increased during pregnancy. The nocturnal awake minutes reduced sleep quality and probably compromised sleep efficiency. The data on sleep quality in pregnancy are limited; thus, the present study adds significantly to the present knowledge of this subject. Smart wristbands provide many opportunities to measure and support the sleep hygiene of pregnant women in maternity care.

The resting heart rate increased until 32 gestational weeks, a normal hemodynamic change during singleton pregnancy [36]; thus, it appears that a smart wristband is an appropriate tool for measuring heart rate during pregnancy. In the future, heart rate and heart-rate variability could be utilized to study the level of stress or the recovery from physical or mental exertions. Maternal stress is associated with some pregnancy complications [37], and identifying the increase in stress could help with finding suitable interventions to support the pregnant woman. The intensity of the physical activity levels in this study was low, as expected, and decreased considerably during the last month of pregnancy. A comparison with the global physical activity recommendation for pregnant women showed that over half of the participants did not reach the recommended 150 minutes of moderate or vigorous physical activity per week during the second or third trimester.

Limitations

The study includes distinct limitations due to the feasibility design. The participants were recruited as a convenience sample; however, the sample consisted of a diverse group of primiparous women. Due to the small sample size, the results from this study may not be generalizable. Furthermore, the small sample size did not allow us to perform statistical analyses between outcome predictors and maternal background characteristics or perinatal outcome variables. Although Garmin Vivosmart has been shown to provide a valid measure of step count [18,19] and total sleep time compared with a sleep diary in a healthy adult population [38], it is notable that the device has not been validated in pregnant women. A sleep diary would have strengthened our results. However, due to the extended data collection period, we made the decision not to use such a method. The intensity of physical activity was estimated based on heart rate elevation. This approach is not unambiguous in pregnant women, because even if the resting heart rate increases, the maximal heart rate decreases during pregnancy, resulting in a smaller heart rate reserve.

Furthermore, the specificity of intensity measures is only moderate as the device may underestimate the intensity at higher rates [19]. Therefore, moderate and vigorous activity minutes may be biased. Missing data due to the participants not wearing the smart wristband or not remembering to synchronize the device might have affected our results. The data collection period was, however, very long, and only valid weeks were included in the analysis. The data during pregnancy were representative and covered the long follow-up period well. After delivery, however, other methods for monitoring should be considered. In addition, the smart wristband used in this study was found to be a feasible tool for continuous monitoring during pregnancy [39].

Conclusions

The smart wristband with the IoT solution seems to be a feasible system to collect representative data on continuous variables such as physical activity, sleep, and heart rate throughout pregnancy. However, more research is needed, and other methods for monitoring after delivery should be considered. Continuous monitoring provides personalized information and thus may help target and tailor pregnancy follow-up. Remote monitoring provides vast opportunities to observe health-related parameters in pregnancy and thus detect pathological signs at an early stage and partly replace traditional appointments. Future studies on predicting complications or support of women with high-risk pregnancies are needed.

Acknowledgments

This research was supported by the University Foundation (Award 5-863) and Academy of Finland (Awards 313448, 313449, 316810, and 316811). The authors would like to thank the personnel of the maternity clinics for their assistance in data collection. Special thanks to the women who participated in the study.

Authors' Contributions

JS drafted the manuscript and contributed to the design of the study and the acquisition of data from the participants. JS also analyzed and interpreted the data from the questionnaires and the participant data related to background information and health records. HNV contributed to the design of the study and the acquisition of data from the participants. HNV was also a major contributor to drafting the manuscript and revising it critically. EE contributed to the design of the study and revised the manuscript

critically throughout the process. LH contributed to drafting the manuscript and analyzed and interpreted the data associated with physical activity, sleep, and heart rate. IA, AR, and PL contributed to the design of the study, facilitated the cloud service for the data collection, and analyzed and interpreted the data from the smart wristbands. EL was responsible for the overall biostatistical data analysis and interpretation. AA was a major contributor to the design of the study, data analysis, and interpretation. AA also contributed to drafting the manuscript and to revising it critically throughout the process. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Klemetti R; Hakulinen-Viitanen T. Äitiysneuvolaopas. 2013. URL: http://urn.fi/URN:ISBN:978-952-245-972-5 [accessed 2020-06-25]
2. Phillips SM, Cadmus-Bertram L, Rosenberg D, Buman MP, Lynch BM. Wearable Technology and Physical Activity in Chronic Disease: Opportunities and Challenges. Am J Prev Med 2018 Jan;54(1):144-150. [doi: 10.1016/j.amepre.2017.08.015] [Medline: 29122356]
3. Fantinelli S, Marchetti D, Verrocchio MC, Franzago M, Fulcheri M, Vitacolonna E. Assessment of Psychological Dimensions in Telemedicine Care for Gestational Diabetes Mellitus: A Systematic Review of Qualitative and Quantitative Studies. Front Psychol 2019;10:153 [FREE Full text] [doi: 10.3389/fpsyg.2019.00153] [Medline: 30804842]
4. Sugawara E, Nikaido H. Properties of AdeABC and AdeIJK efflux systems of Acinetobacter baumannii compared with those of the AcrAB-TolC system of Escherichia coli. Antimicrob Agents Chemother 2014 Dec;58(12):7250-7257 [FREE Full text] [doi: 10.1128/AAC.03728-14] [Medline: 25246403]
5. Mieronkoski R, Azimi I, Rahmani AM, Aantaa R, Terävä V, Liljeberg P, et al. The Internet of Things for basic nursing care-A scoping review. Int J Nurs Stud 2017 Apr;69:78-90. [doi: 10.1016/j.ijnurstu.2017.01.009] [Medline: 28189116]
6. 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]
7. Ming W, Mackillop LH, Farmer AJ, Loerup L, Bartlett K, Levy JC, et al. Telemedicine Technologies for Diabetes in Pregnancy: A Systematic Review and Meta-Analysis. J Med Internet Res 2016 Nov 09;18(11):e290 [FREE Full text] [doi: 10.2196/jmir.6556] [Medline: 27829574]
8. Li KHC, White FA, Tipoe T, Liu T, Wong MC, Jesuthasan A, et al. The Current State of Mobile Phone Apps for Monitoring Heart Rate, Heart Rate Variability, and Atrial Fibrillation: Narrative Review. JMIR Mhealth Uhealth 2019 Feb 15;7(2):e11606 [FREE Full text] [doi: 10.2196/11606] [Medline: 30767904]
9. Runkle J, Sugg M, Boase D, Galvin SL, C Coulson C. Use of wearable sensors for pregnancy health and environmental monitoring: Descriptive findings from the perspective of patients and providers. Digit Health 2019;5:2055207619828220 [FREE Full text] [doi: 10.1177/2055207619828220] [Medline: 30792878]
10. Kolu P, Raitanen J, Luoto R. Physical activity and health-related quality of life during pregnancy: a secondary analysis of a cluster-randomised trial. Matern Child Health J 2014 Nov;18(9):2098-2105. [doi: 10.1007/s10995-014-1457-4] [Medline: 24585400]
11. Coll CDVN, Domingues MR, Hallal PC, da Silva ICM, Bassani DG, Matijasevich A, et al. Changes in leisure-time physical activity among Brazilian pregnant women: comparison between two birth cohort studies (2004 - 2015). BMC Public Health 2017 Dec 25;17(1):119 [FREE Full text] [doi: 10.1186/s12889-017-4036-y] [Medline: 28122524]
12. Bei B, Coo S, Trinder J. Sleep and Mood During Pregnancy and the Postpartum Period. Sleep Med Clin 2015 Mar;10(1):25-33. [doi: 10.1016/j.jsmc.2014.11.011] [Medline: 26055670]
13. Petrov Fieril K, Glantz A, Fagevik Olsen M. The efficacy of moderate-to-vigorous resistance exercise during pregnancy: a randomized controlled trial. Acta Obstet Gynecol Scand 2015 Jan;94(1):35-42. [doi: 10.1111/aogs.12525] [Medline: 25287282]
14. Bonn SE, Bergman P, Trolle Lagerros Y, Sjölander A, Bälter K. A Validation Study of the Web-Based Physical Activity Questionnaire Active-Q Against the GENEA Accelerometer. JMIR Res Protoc 2015 Jul 16;4(3):e86 [FREE Full text] [doi: 10.2196/resprot.3896] [Medline: 26183896]
15. Brickwood K, Watson G, O'Brien J, Williams AD. Consumer-Based Wearable Activity Trackers Increase Physical Activity Participation: Systematic Review and Meta-Analysis. JMIR Mhealth Uhealth 2019 Apr 12;7(4):e11819 [FREE Full text] [doi: 10.2196/11819] [Medline: 30977740]
16. Allen J. Photoplethysmography and its application in clinical physiological measurement. Physiol Meas 2007 Mar;28(3):R1-39. [doi: 10.1088/0967-3334/28/3/R01] [Medline: 17322588]
17. Snyder JC, Rochelle LK, Marion S, Lyerly HK, Barak LS, Caron MG. Lgr4 and Lgr5 drive the formation of long actin-rich cytoneme-like membrane protrusions. J Cell Sci 2015 Mar 15;128(6):1230-1240 [FREE Full text] [doi: 10.1242/jcs.166322] [Medline: 25653388]

18. Wahl Y, Düking P, Droszez A, Wahl P, Mester J. Criterion-Validity of Commercially Available Physical Activity Tracker to Estimate Step Count, Covered Distance and Energy Expenditure during Sports Conditions. Front Physiol 2017;8:725 [FREE Full text] [doi: 10.3389/fphys.2017.00725] [Medline: 29018355]
19. Kooiman TJM, Dontje ML, Sprenger SR, Krijnen WP, van DSCP, de GM. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabil 2015;7:24 [FREE Full text] [doi: 10.1186/s13102-015-0018-5] [Medline: 26464801]
20. Garmin Vívosmart HR/HR+ Owner´s Manual. 2015. URL: https://static.garmin.com/pumac/vivosmart_HR_OM_EN.pdf [accessed 2020-06-25]
21. Al-Fuqaha A, Guizani M, Mohammadi M, Aledhari M, Ayyash M. Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. IEEE Commun. Surv. Tutorials 2015;17(4):2347-2376. [doi: 10.1109/comst.2015.2444095]
22. McCarthy M, Grey M. Motion Sensor Use for Physical Activity Data: Methodological Considerations. Nurs Res 2015;64(4):320-327 [FREE Full text] [doi: 10.1097/NNR.0000000000000098] [Medline: 26126065]
23. Matthews CE, Hagströmer M, Pober DM, Bowles HR. Best practices for using physical activity monitors in population-based research. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S68-S76 [FREE Full text] [doi: 10.1249/MSS.0b013e3182399e5b] [Medline: 22157777]
24. Khan T, Macaulay S, Norris SA, Micklesfield LK, Watson ED. Physical activity and the risk for gestational diabetes mellitus amongst pregnant women living in Soweto: a study protocol. BMC Womens Health 2016 Dec 18;16(1):66 [FREE Full text] [doi: 10.1186/s12905-016-0345-z] [Medline: 27756289]
25. Reid KJ, Facco FL, Grobman WA, Parker CB, Herbas M, Hunter S, et al. Sleep During Pregnancy: The nuMoM2b Pregnancy and Sleep Duration and Continuity Study. Sleep 2017 May 01;40(5) [FREE Full text] [doi: 10.1093/sleep/zsx045] [Medline: 28369543]
26. American College of Obstetricians Gynecologists' Committee on Obstetric Practice, Breastfeeding Expert Work Group. Committee Opinion No. 658: Optimizing Support for Breastfeeding as Part of Obstetric Practice. Obstet Gynecol 2016 Feb;127(2):e86-e92. [doi: 10.1097/AOG.0000000000001318] [Medline: 26942393]
27. Hayes L, McParlin C, Kinnunen TI, Poston L, Robson SC, Bell R, UPBEAT Consortium. Change in level of physical activity during pregnancy in obese women: findings from the UPBEAT pilot trial. BMC Pregnancy Childbirth 2015 Mar 01;15:52 [FREE Full text] [doi: 10.1186/s12884-015-0479-2] [Medline: 25886590]
28. Nascimento SL, Surita FG, Godoy AC, Kasawara KT, Morais SS. Physical Activity Patterns and Factors Related to Exercise during Pregnancy: A Cross Sectional Study. PLoS One 2015;10(6):e0128953 [FREE Full text] [doi: 10.1371/journal.pone.0128953] [Medline: 26083416]
29. Huberty JL, Buman MP, Leiferman JA, Bushar J, Adams MA. Trajectories of objectively-measured physical activity and sedentary time over the course of pregnancy in women self-identified as inactive. Prev Med Rep 2016 Jun;3:353-360 [FREE Full text] [doi: 10.1016/j.pmedr.2016.04.004] [Medline: 27419036]
30. Mindell JA, Cook RA, Nikolovski J. Sleep patterns and sleep disturbances across pregnancy. Sleep Med 2015 Apr;16(4):483-488. [doi: 10.1016/j.sleep.2014.12.006] [Medline: 25666847]
31. Wilson DL, Barnes M, Ellett L, Permezel M, Jackson M, Crowe SF. Decreased sleep efficiency, increased wake after sleep onset and increased cortical arousals in late pregnancy. Aust N Z J Obstet Gynaecol 2011 Feb;51(1):38-46. [doi: 10.1111/j.1479-828X.2010.01252.x] [Medline: 21299507]
32. Williams MA, Miller RS, Qiu C, Cripe SM, Gelaye B, Enquobahrie D. Associations of early pregnancy sleep duration with trimester-specific blood pressures and hypertensive disorders in pregnancy. Sleep 2010 Oct;33(10):1363-1371 [FREE Full text] [Medline: 21061859]
33. Rognmo K, Sivertsen B, Eberhard-Gran M. Self-reported short sleep duration and insomnia symptoms as predictors of post-pregnancy weight change: Results from a cohort study. Womens Health (Lond) 2016 Dec;12(5):465-474 [FREE Full text] [doi: 10.1177/1745505716668871] [Medline: 27634975]
34. Yu Y, Li M, Pu L, Wang S, Wu J, Ruan L, et al. Sleep was associated with depression and anxiety status during pregnancy: a prospective longitudinal study. Arch Womens Ment Health 2017 Dec;20(5):695-701. [doi: 10.1007/s00737-017-0754-5] [Medline: 28685391]
35. Xu Y, Shi L, Bao Y, Chen S, Shi J, Zhang R, et al. Association between sleep duration during pregnancy and gestational diabetes mellitus: a meta-analysis. Sleep Med 2018 Dec;52:67-74. [doi: 10.1016/j.sleep.2018.07.021] [Medline: 30286382]
36. Hunter S, Robson SC. Adaptation of the maternal heart in pregnancy. Br Heart J 1992 Dec;68(6):540-543 [FREE Full text] [doi: 10.1136/hrt.68.12.540] [Medline: 1467047]
37. Coussons-Read ME. Effects of prenatal stress on pregnancy and human development: mechanisms and pathways. Obstet Med 2013 Jun;6(2):52-57 [FREE Full text] [doi: 10.1177/1753495X12473751] [Medline: 27757157]
38. Lee J, Byun W, Keill A, Dinkel D, Seo Y. Comparison of Wearable Trackers' Ability to Estimate Sleep. Int J Environ Res Public Health 2018 Dec 15;15(6) [FREE Full text] [doi: 10.3390/ijerph15061265] [Medline: 29914050]
39. Grym K, Niela-Vilén H, Ekholm E, Hamari L, Azimi I, Rahmani A, et al. Feasibility of smart wristbands for continuous monitoring during pregnancy and one month after birth. BMC Pregnancy Childbirth 2019 Jan 17;19(1):34 [FREE Full text] [doi: 10.1186/s12884-019-2187-9] [Medline: 30654747]

Abbreviations

GDM: gestational diabetes mellitus
IoT: Internet of Things
MVPA: moderate-to-vigorous physical activity

Edited by G Eysenbach; submitted 24.05.19; peer-reviewed by M Aittasalo, L Mackillop, J van den Heuvel; comments to author 09.12.19; revised version received 03.02.20; accepted 14.05.20; published 24.07.20.

Please cite as:
Saarikko J, Niela-Vilen H, Ekholm E, Hamari L, Azimi I, Liljeberg P, Rahmani AM, Löyttyniemi E, Axelin A
Continuous 7-Month Internet of Things–Based Monitoring of Health Parameters of Pregnant and Postpartum Women: Prospective Observational Feasibility Study
JMIR Form Res 2020;4(7):e12417
URL: http://formative.jmir.org/2020/7/e12417/
doi:10.2196/12417
PMID:32706696

Original Paper

MyStrengths, a Strengths-Focused Mobile Health Tool: Participatory Design and Development

Stian Jessen[1,2], MPhilED; Jelena Mirkovic[1], PhD; Lise Solberg Nes[1,2], PhD, LP

[1]Department of Digital Health Research, Division of Medicine, Oslo University Hospital, Oslo, Norway

[2]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

Corresponding Author:
Stian Jessen, MPhilED
Department of Digital Health Research
Division of Medicine
Oslo University Hospital
Trondhjemsveien 235
Oslo, 0586
Norway
Phone: 47 98049759
Email: stian.jessen@rr-research.no

Abstract

Background: People living with chronic illnesses are an increasingly large group. Research indicates that care and self-management should not only focus on the illness and problem-oriented aspects of these individuals' lives but also support them in recognizing and leveraging their personal strengths in daily life.

Objective: This paper presents the design and developmental process of *MyStrengths*, a mobile health (mHealth) app designed to help its users (people with chronic conditions) both find and make use of their personal strengths in their daily lives. Through 4 consecutive phases, this paper presents participant- and researcher-driven activities, discussions regarding design, and development of both the *MyStrengths* app and its content.

Methods: During the 4 phases, we used a range of methods and activities, including (1) an idea-generating workshop aimed at creating ideas for strengths-supporting features with different stakeholders, including patients, caregivers, relatives, and designers (N=35); (2) research seminars with an international group of experts (N=6), in which the concept, theoretical background, and design ideas for the app were discussed; (3) a series of co-design workshops with people in the user group (N=22) aiming to create ideas for how to, in an engaging manner, design the app; and (4) in 4 developmental iterations, the app was evaluated by people in the user group (N=13). Content and strengths exercises were worked on and honed by the research team, the expert groups, and our internal editorial team during the entire developmental process.

Results: The first phase found a wide range of stakeholder requirements to, and ideas for, strengths-focused mHealth apps. From reviewing literature during the second phase, we found a dearth of research on personal strengths with respect to people living with chronic illnesses. Activities during the third phase creatively provided numerous ideas and suggestions for engaging and gameful ways to develop and design the *MyStrengths* app. The final phase saw the output from all the earlier phases come together. Through multiple increasingly complete iterations of user evaluations testing and developing, the final prototype of the *MyStrengths* app was created.

Conclusions: Although research supports the use of strengths-focused mHealth tools to support people living with chronic illnesses, there is little guidance as to how these tools and their content should be designed. Through all activities, we found great support among participating users for strengths-focused apps, and we can consider such apps to be both appropriate and valuable. This paper illustrates how combining a range of user-, researcher-, literature-, and designer-based methods can contribute to creating mHealth tools to support people with chronic illnesses to find and use more of their own personal strengths.

(JMIR Form Res 2020;4(7):e18049) doi:10.2196/18049

KEYWORDS

mHealth; participatory design; personal strengths; gameful design; gamification; iterative development; positive approach; co-design; user engagement; mobile phone; chronic care; self-management

Introduction

Background

The number of people living with chronic illnesses is continually increasing [1,2]. Between 2010 and 2020, the World Health Organization has predicted the number of persons with chronic conditions to rise by 15% [3]. For instance, in the United States, 6 in 10 people have one chronic disease, and as many as 4 in 10 have 2 or more chronic diseases [4]. Living with chronic illnesses is a highly demanding task, and many of the challenges faced are shared across conditions and demographics. For instance, chronically ill people can be required to cope with symptoms, disabilities, medication regiments, lifestyle changes, or emotional consequences such as depression and fear [1,5,6]. Managing these aspects of life with chronic illnesses is often described as self-management [7,8]. Traditionally, self-management interventions in health care have focused on patients' actual or potential health problems, risks, and deficits [7,9]. This is usually done by providing support to solve, alleviate, or prevent such problems through information, skill training, and teaching coping techniques. However, evidence from psychology and biobehavioral sciences points to the traditional health care approach, that is, focusing on people's problems and deficits, as not optimal to aid people in reaching their best health potentials [10,11]. Increasingly, self-management interventions include more balanced approaches that not only assess a person's deficiencies (eg, symptoms, problems, and needs) but also integrate positive personal resources and values. They are, in essence, providing a more holistic approach to health and well-being [8,12,13].

As the group of people living with chronic illnesses grows, a readily available mHealth tool that is not disease specific but instead aims at supporting its users overcome commonly shared challenges could be of great use. One particularly promising approach to provide support anchored in an individual's positive personal resources and values is to help them in recognizing and using more of their personal strengths [10,14]. In this paper, we describe the process of creating such a tool, the *MyStrengths* app.

Personal Strengths

The concept of personal strengths has its foundation in positive psychology [10,15] and has been defined as "traits/capabilities that are personally fulfilling, do not diminish others, ubiquitous, and valued across cultures, and aligned with numerous positive outcomes for oneself and others" [14]. Simply put, this means emphasizing what is possible, valuable, and doable, as opposed to focusing on deficits and problems [16,17]. A focus on people's own strengths has been shown to contribute positively to better moods and happiness [10] as well as increased general health and well-being [14,18,19]. Focusing on health care, Sturgeon and Zautra [20] reported how people with chronic pain use traits such as positive emotions, optimism, and social engagement to maintain a good life. Similarly, Rotegård et al [21] found cancer patients to employ strengths items such as will power and trust in health care providers to meet their daily challenges. In a study on adults with major depressive disorder, Cheavens et al [22] found better outcomes from personalizing treatment to focus on the patients' strengths rather than on their deficits and problems. In a study among people with one or more chronic conditions, positive emotions have been connected to increasing patient activation [23]. In addition to the strengths reported in the studies cited earlier, a multitude of different strengths used to overcome challenges and live a good life has been identified by participants in studies on strengths among people living with chronic illness in general: having a positive outlook on life; being persistent; being kind and caring; having courage; having support from family, friends, peers, and health care providers; and having constructive self-management strategies [16,24,25].

In sum, being aware of and mobilizing one's personal strengths can lead to a wide range of positive effects on health and well-being both for people in the general population as well as people living with chronic illnesses.

mHealth and Design for Engagement

In addition to the increasing ubiquity of smartphones and personal computers, both electronic health (eHealth) and mobile health (mHealth) interventions have been developed to support a wide range of health-related goals, for instance, medication adherence [26], symptom monitoring [27], support of smoking cessation [28], managing rheumatic and musculoskeletal diseases [29], and stress management [30]. However, to our knowledge, no tools have been designed focusing on identifying and mobilizing people's personal strengths in support of people living with chronic illnesses.

Although mHealth tools or services show great promise for supporting people with chronic illnesses [31], their success is often contingent on them being used as intended [32,33]. However, focusing on use alone is not necessarily enough, and as Kelders [33] argues, one should also take into consideration users' sense of involvement with and enjoyment of the mHealth tools. Together, these 3 aspects make up what can be referred to as engagement [34,35]. To increase users' engagement with tools or services, the use of design approaches known from the world of games, typically called gamification or gameful designs, is increasingly popular [36-39]. There are several definitions of gameful designs, and in this project, we define it similar to the one presented by Huotari and Hamari [40]: using design approaches and implementations from the world of games (in our otherwise nongame tools) to add a sense of playfulness and increase users' enjoyment and engagement. To date, gameful eHealth and mHealth tools have been designed for a variety of purposes, such as mental health [41], smoking cessation [42], and promoting physical activity among patients with rheumatoid arthritis [43].

Concerning the efficacy of gameful designs, Johnson et al [36] reviewed gameful designs for health and well-being. They found a positive influence of gameful designs for tools aimed at increasing physical activity and fitness, nutrition, health care utilization, medication misuse, blood glucose monitoring, and patient empowerment. Importantly, the authors also found gameful designed tools to have positive effects on personal growth, well-being, flourishing, stress, and anxiety. Thus, as they have the potential to positively affect not only behavior but also users' well-being through positive and engaging

experiences, gamefully designing tools or services should be particularly promising within the field of health [36,38].

Looking at specific gameful design techniques and implementations, the use of *points, badges, and leaderboards*, often referred to together as the PDB triad, is the most used approach among gameful interventions [36,37]. Such approaches have been critiqued for simplifying the nature of human motivation by assuming that humans primarily are motivated by achieving increasingly more or by merely leading or winning over others [44,45]. Gameful design approaches, such as chasing rewards, have also been reported as unfitting for mindfulness and well-being interventions [46]. It should be noted, however, that gamified well-being interventions have also been designed successfully using the aforementioned approaches [38].

Thus, when designing a tool that centers around self-awareness and reflection, it appears crucial to create engagement in ways that are aligned with the users' real interests and values. This can, for instance, be done by having the app represent *a better self* of its users with avatars and creating an involving and interesting narrative, rather than only external factors such as points and leaderboards [45,47]. Although research points to the positive effects of gameful designs, no evidence-based framework for designing eHealth or mHealth interventions gamefully currently exists [48,49]. Furthermore, what the *active ingredients* of successful gameful designs are is mostly unknown [36].

Participatory Design Processes

Users' participation in the design process is a productive way of ensuring that gameful designs are experienced as appropriate as well as meaningful to the end users [48,50-52]. Contributing to the literature on participatory eHealth development, this project takes a *participatory design* approach [53] to investigate ways of designing and creating the *MyStrengths* app.

Participatory design is guided by the fundamental ethical stance that the end users, whose future may be affected by the design, should have a say in the process. As such, participatory design processes do not merely include users as passive informers and evaluators but seek to include these as co-designers throughout the design process [54,55]. As such, participatory design is not merely concerned with collecting users' needs, ideas, or preferences but also with enabling meaningful participation. Examples include using techniques such as design games [56], role-playing [57], or future workshops [58]. In eHealth and mHealth, participatory approaches have, for instance, been involved in the design of apps to support teenagers with chronic illnesses in the transition from pediatric to regular care [59], to support young children living with cancer [60], and to facilitate stress management for cancer survivors [30].

Aims

As presented, the number of people living with one or more chronic illnesses is increasing. The rising ubiquity of smartphones affords mHealth research and designers to create support tools that can easily reach a large number of users without relying on their physical access to health care services and personnel. As personal strengths are common to us all, a tool supporting people living with chronic illnesses to find and use more of their own should thus be of great benefit. To our knowledge, neither the development nor the evaluation of any such tool has been published before.

Addressing this, our main goal was to present the activities, discussions, and decisions undertaken during designing the *MyStrengths* app. Our secondary goal addresses the lack of guidelines for creating both gameful designs and strengths features or activities for people living with chronic illnesses. Through this project, we explore and evaluate ways to integrate the focus of people's personal strengths into mHealth tools and how to make such tools more engaging through gameful and engaging designs in ways that are suitable for, respectful to, and appreciated by people living with chronic illnesses.

Methods

Overall Project Design

This paper presents activities from the project "The Power of Personal Strengths—using gamification to support patients in chronic illness management." The project was conducted at the Department of Digital Health Research at Oslo University Hospital, Norway, between 2016 and 2019, and funded by the Research Council of Norway (grant #248026).

To achieve its goal of creating the *MyStrengths* app, the project has a comprehensive, iterative, and participatory approach to combine patients' preferences and requirements with knowledge and expertise from the fields of self-management, positive psychology and well-being, and mHealth design. The overall design process follows the 4 stages presented in the Double Diamond design process [61]: discovering, defining, developing, and delivering. The 2 first and the 2 latter stages make up a diamond of diverging and converging activities. Divergence refers to activities where designers (and co-designers) go broadly out to discover new opportunities, solutions, and ideas for their design. Conversely, the converging activities are focused on narrowing down, concretizing, and creating based on the former phase [62].

Participants

The main project team consisted of 5 researchers (with backgrounds in nursing, public health, behavior change, informatics and design, and mHealth development), 2 editorial content producers, and a patient representative, to ensure that the user perspective was maintained. Although the project, outside of one of the activities, does not directly include health care workers, it should be noted that 4 of the 8 participants in the group had long clinical working experiences. The team met weekly and planned and coordinated the work throughout the entire project period.

In addition to the main project team, the project employed 3 main types of participants during the activities:

1. People in the user group, living with chronic illnesses. These people were recruited through collaborating institutions, such as learning and mastery centers at local hospitals or patient organizations, or from the team's professional network. Recruited in the same manner, patient

representatives and caregivers as well as designers and researchers took part during the first phase.
2. Designers and developers, external and from the in-house information technology (IT) system development group.
3. Researchers who are part of the project's international advisory group.

Ethical Considerations

The project was planned and conducted in adherence to the principles of the Helsinki Declaration [63] and approved by the Privacy Protection and Data Security Committee at Oslo University Hospital. All participants, or their legal guardians, signed informed consent before taking part.

Results

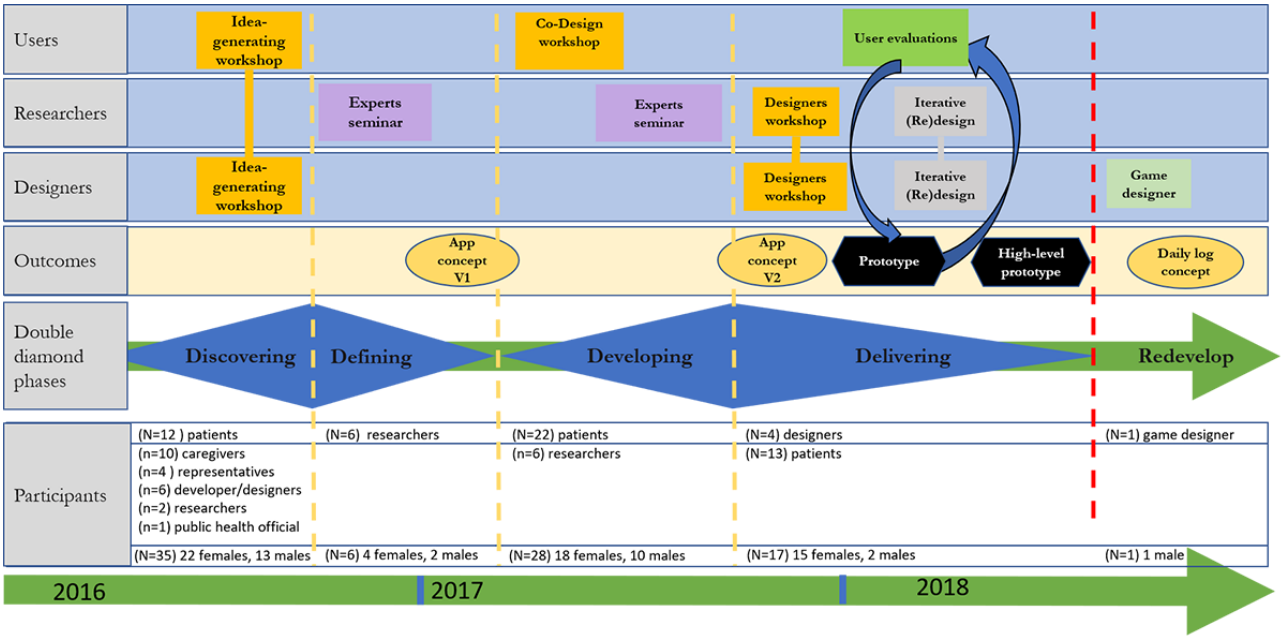
Design Activities and Results

The following section presents the iterative activities and the outcomes from the development process. As each phase builds on the earlier phase, the activities and results are presented grouped into the 4 phases of the design process:

1. In the discovering phase, users and other stakeholders took part in a full-day workshop, working toward identifying and ideating ideas for how a strengths-focused mHealth tool could work and be designed.
2. In the defining phase, inputs from the earlier phase as well as from literature and previous projects and experiences were discussed, and decisions on content and priorities were made.
3. The developing phase focused on the user experience of the tool.
4. In the delivering phase, all information and input gathered were put together, and a working high-fidelity prototype of the *MyStrengths* tool was developed.

A visual overview of the project's phases, activities, and participants is presented in Figure 1.

Figure 1. The project's phases, activities, and participants.



Phase 1: Discovering

The first phase of development explored, through an idea-generating workshop, how a mHealth tool could be created to support people living with chronic illness through discovering and using their own strengths.

Methods and Activities

Idea-Generating Workshop

To gather ideas and input from relevant stakeholders, a day-long idea-generating workshop was hosted, where patients, relatives, representatives from patient organizations, health care personnel and researchers, and designers and developers participated (n=35). The design of the workshop was inspired by Future Workshops, a common idea-generation activity within participatory design [57], and additionally adjusted with a positive and forward-looking approach based on principles from *appreciative inquiry* [64]. The participants worked in 6 groups hosted by members of the project team, and the first part of the workshop consisted of group activities focused on identifying and presenting each person's personal strengths, naming typical challenges experienced by people living with chronic illnesses, and what strengths people could use to overcome these. An illustration of this activity is shown in Figure 2. The latter part of the workshop had the same groups create concepts for an IT tool that could help them use their own strengths to overcome their daily challenges in a new and innovative manner.

In addition to the briefly described findings presented here, a detailed description of workshop methods and results are described in a separate publication [65]. Data from the workshops were qualitatively analyzed by the first and fourth authors of that publication using thematic analysis [66].

Figure 2. Activity from Idea Generation workshop.

Results

App Concepts—Strengths Focus

Each of the six participant groups created concepts for different eHealth tools that could support them in overcoming the selected challenges.

For example, Figure 3 shows the concept created by one of the participant groups. The suggestion is a personalized app that is designed to help the user identify her strengths and use them to complete challenges and tasks in everyday life. It could, for example, integrate sensors to interpret when a person is stressed and then prompt her with some of her strengths to help and motivate her to complete initiated tasks. The group participants also suggested more specific ideas for the strengths focus in mHealth tools, for instance, (1) having friends in the app suggest strengths the user has; (2) that the app provides the user reminders of her strengths; (3) based on previous input in the app, providing tips for what strengths the user can use to overcome new challenges; and (4) that the tool could offer exercises and tasks that would either build on or provide opportunities for the users to mobilize and use their own strengths.

Figure 3. Example of User-Concept.

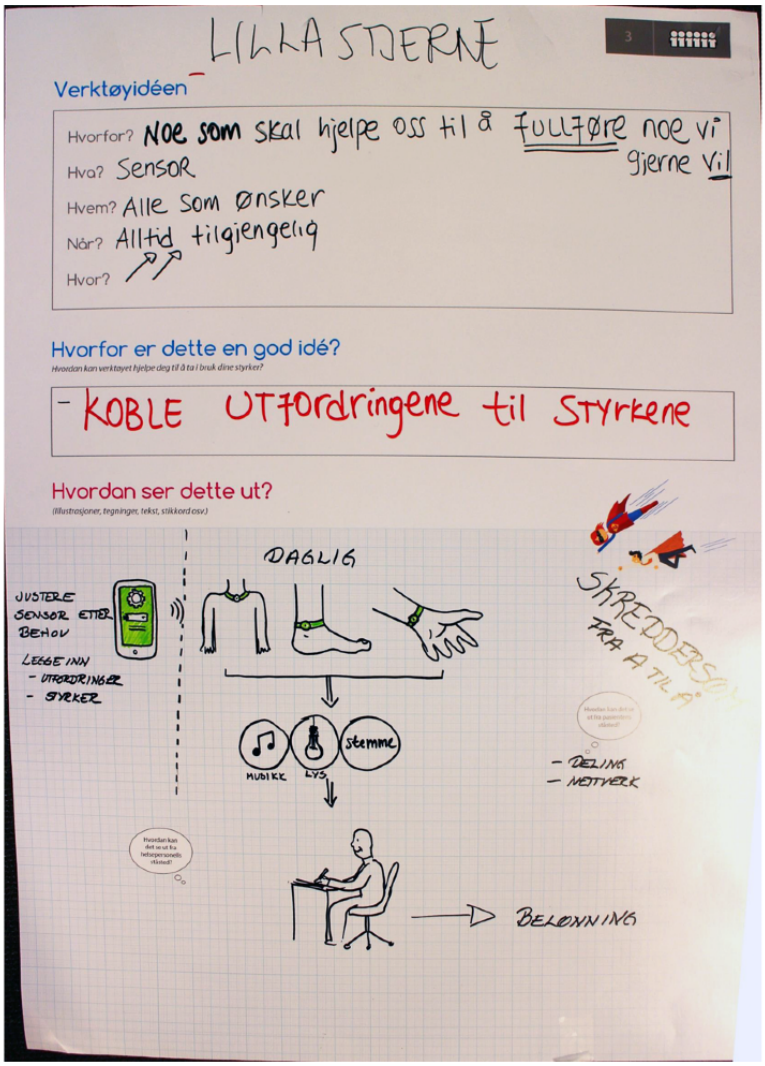


Table 1 presents the 6 groups and their mHealth tool suggestions. We planned to distribute the participants evenly between the groups based on gender and backgrounds, although some changes were outside our control. Although the groups vary in composition and only some of the ideas explicitly focused on mobilizing the personal strengths among people living with chronic illnesses, a positive and strengths-focused foundation exists in all the app ideas.

Table 1. Six proposed mobile health or electronic health tools from the idea-generating workshop (N=35, 13 males).

Group number (participants)	App idea	Challenges	Description
1 (n=5, 2 males)	Strengths treasure chest	Finding balance in life	A treasure chest app where a person can store his or her strengths, those written by himself or herself as well as strengths added by others (eg, friends and family)
Patient (n=1)			
Health care provider (n=1)			
Public health official (n=1)			
IT[a] developer (n=1)			
Researcher (n=1)			
2 (n=8, 3 males)	Cheering squad app	Mastering various aspects of life	An app where a person can invite people to join his or her own cheering squad
Patient (n=4)			
Health care provider (n=3)			
IT developer (n=1)			
3 (n=6, 3 males)	User-controlled personalized hospital	Challenging to be a young adult in a hospital	An app for a person transitioning from a pediatric ward to a unit for adults in a hospital
Patient (n=2)			
Health care provider (n=2)			
Researcher (n=1)			
IT developer (n=1)			
4 (n=7, 1 male)	Empathy simulator	Communication among relatives and the health care system	Virtual reality four-dimensional glasses that simulate experiences from different parties present in a consultation setting (eg, patient, caregiver, or a family member)
Patient (n=1)			
Patient organization Representative (n=2)			
Health care provider (n=3)			
IT developer (n=1)			
5 (n=6, 2 males)	Prioritizing app	Prioritize among important things	An app to help a person to make choices based on earlier knowledge and experiences
Patient (n=3)			
Health care provider (n=1)			
Patient organization Representative (n=1)			
Designer (n=1)			
6 (n=3, 1 male)	Task completer app	Finishing projects	An app that helps a person identify his or her strengths and use them to complete tasks in everyday life
Patient (n=1)			
Health care provider (n=1)			
Designer (n=1)			

[a]IT: information technology.

User Requirements and Functionality Ideas

The analysis of the workshops revealed 4 main themes of functionality requirements for the strengths-focused self-management tool:

1. Social support, that the tool should include support for social support and interaction, for example, by providing the possibility to chat with peers or role models, or one-directional messaging that friends can send to the user to cheer him or her up but that does not require the person to respond.
2. Supporting patient-health care providers' collaboration, for example, by supporting communication with providers, preparation for consultations, or easy sharing of information. One group suggested giving the user the possibility to use the app to give providers better insights into their feelings and values and using this information so that the treatment could be adjusted to the best fit these.
3. Awareness and reflection, allowing users to develop awareness and reflection about oneself and one's current situation. This could, for instance, be done by adding to and adapting the app to fit the users' values and situation or by helping the users to be more aware of, and use, their own strengths.
4. Supporting the users with coping strategies, by, for example, helping them to prioritize and make choices between different activities and goals that they wish to do or accomplish and by providing an overview of activities, goals, and choices done in the past.

Users' Preferences and Needs for Design and User Experience

The workshop also revealed a wide range of user requirements for the design of the strengths tool itself. The most prevalent requirement is that the design should have an overall positive focus and approach. For instance, it was suggested that all feedback be given in a supportive and positive manner or using *pleasant* and engaging metaphors such as treasure chests, islands in an ocean, or avatars. The participants also considered it important to be able to customize and adapt the tool to their own needs and preferences. This could include adjusting the tool content depending on the user's diagnosis or health situation, their level of experience, and simply being able to turn functionalities on or off.

Phase 2: Defining

To further inform the project, the second phase mainly consisted of us performing a literature search on strengths-focused interventions. These findings, as well as the outcomes from the first phase, were then discussed and evaluated in a research seminar.

Methods and Activities

Literature Review

To further explore the topic of strengths-focused management and support of individuals with chronic illnesses, we conducted a literature search using terms that employ positive approaches (eg, positive psychology interventions and mindfulness interventions). Although this review was intended to be published, the resource situation and practical changes on the project would regrettably thwart this. The findings are nonetheless outcomes of the project and are, therefore, presented in this paper.

Research Seminar

We hosted a 3-day research seminar with the main project team and 6 experienced national and international researchers from the field of behavior change, psychology, eHealth, service design, and participatory care. The primary goal of this research seminar was to combine the results from the previous phase with knowledge, evidence, and previous experiences to decide on a set of core features for the *MyStrengths* tool. Although the user representative was the only user taking part in the seminar, we took care to ensure that the opinions, needs, and ideas voiced by users were present and given equal weight during discussions and decisions. A secondary goal of the research seminar was to discuss the findings from the review of previous literature and how these findings might be used beneficially in the project.

The activities in the research seminar took many forms, including presentations, brainstorming, discussions, and versions of the activities from the idea-generating workshop held in the previous phase. The seminar was audio recorded, and a detailed summary and a key item take away form was created using these recordings as well as participants' notes, drawings, and photos.

Results

Review

From an initial search response of 6742 records, 7 publications on 6 different interventions were identified and included in the analysis. Three of these were delivered as in-person or face-to-face interventions [67-70], 2 interventions were delivered through web-based channels [71,72], and 1 intervention consisted of both offline and online intervention delivery modes, one per intervention group [73]. In all studies, personal strengths were implemented as one part of the intervention, and none of the studies focused solely on personal strengths. In all, a lack of consistency in the literature was found, and personal strengths were scarcely covered. This is also a common finding in more extensive reviews on behavior change techniques [74].

Of the more promising findings, the study by Nikrahan et al [67] identified a positive effect on hope and happiness at a 15-week follow-up of the participants. Participants were randomly assigned to 1 of the 3 positive intervention groups or a wait-list control group, and participants in the intervention groups received a 6-week in-person group training program. Two of the intervention groups included strengths activities. In the first group, participants were asked to identify a *signature strength* from a list of 24 personal strengths, use a signature strength in daily activities and identify strengths in their partners and children, and use a signature strength in a way that furthers a cause larger than oneself. In the second group, participants in parts of the intervention focused on positive personality traits to overcome fears about others' opinions; accept themselves; and initiate contact with people they would like to meet to foster authenticity, self-esteem, and extroversion. In another study by Cerezo et al [69], the experimental group received weekly face-to-face sessions aiming to provide positive

psychology-related coping strategies and enhance their psychological strengths. From this, the authors found positive between-group effects on positive emotions postintervention when comparing the experimental group with a control group.

In addition to the activities described earlier, strengths activities employed in all the studies identified ranged from questionnaires asking participants to select strength(s) items that most apply for them or naming strengths directly to exercises where they reflect on how they have used their strengths recently or aim to use a signature strength in a daily activity. Although such exercises are in line with common approaches for identifying personal strengths [14], none of the included studies reported on adaptations of strengths exercises and activities to the specific target groups.

A brief description of the 6 identified interventions' designs and findings is presented in Multimedia Appendix 1 [67-73]. As these interventions typically consist of multiple modules stemming from separate theories or approaches, it was not possible to conclude on either specific effects or mechanisms through which the interventions affected well-being. However, other studies have previously identified goal setting as a potential promising factor. More specifically, Linley et al [75] have shown strengths use to be associated with goal progress, which, in turn, was related to psychological need fulfillment and enhanced well-being in the general population, thus providing an example of how positive interventions could be combined with self-management interventions to support patients in better management of chronic illnesses.

Research Seminar Discussions and Project Development

When discussing review findings and further plans, we thoroughly discussed the overall approach of the *MyStrengths* tool and how it should be designed for the best possible effects. On the basis of the lack of evidence for specific mechanisms for increasing people's use of strengths, a general 3-step approach was suggested: (1) create awareness of people's strength; (2) help users to reflect on these; and (3) support users in using the strengths by, for instance, setting small and achievable goals. This was also along the lines of 2 of the identified publications [72,73], which also integrated personal strengths in the same module as personal goals. Concluding this phase, the outcomes from the idea workshop in the previous phase were presented, discussed, and sorted into main categories of possible features for the *MyStrengths* tool. These features are shown and briefly described in Textbox 1.

Textbox 1. Suggested MyStrengths tool features.

My strengths

Assessment and overview of the user's own personal strengths.

My goals

Larger goals the users want to achieve and can use his or her strengths to reach.

My small experiments

Small goals or activities that can serve as building blocks on the road to achieving the user's larger goals.

Exercises

Activities or tasks the user can do to use his or her strengths more. These can be created by the user or be preprogrammed within the app.

My experiences

A logbook that would allow the user to write down and reflect on activities or situations in relation to how they did or did not use their strengths. These reflections could then be accessed at a later stage and form part of the users planning or reflection on new activities, thus help them in using their strengths more productively.

Information

Content that explains as well as expands on the concept of strengths and its scientific background. This section could also provide specific information for the users, for instance, based on specific illnesses or a geographic area.

Social support

A variety of social features that could allow the user to communicate, support, or get support from other users, for instance, by sharing experiences or by supporting each other in reaching goals.

Timeline

A section of the app that would gather all input information in one place and present it in a visual and nice fashion.

About me

A profile page the users can set up to describe themselves, their values, and what is important to them.

Settings

General settings for the app.

Phase 3: Developing

Having created a set of features or functionalities of the *MyStrengths* tools, the goal for the third phase was to further develop and hone ideas and concepts for how to best design and implement features in an engaging and motivating way for users. This phase consisted of 2 main activities: a series of co-design workshops with users and another research seminar with experts in the field.

Methods and Activities

Co-Design Workshops

With the suggested features of the *MyStrengths* tool as a starting point, design challenges and low-fidelity mock-ups were created and used in a series of co-design workshops. In total, 2 workshops were conducted with each of the 3 different participant groups, from 2 educational centers and a youth council from hospitals in Norway.

Each of the workshops used participatory design methods, including design games, prototyping, and scenario making [57,76,77]. The first workshop focused on the design of the tool in a gameful and engaging way, and the second workshop built on the former and focused on the users' discovery and use of their own personal strengths into such tools. In addition to providing new ideas for the design and features of the *MyStrengths* tool, the workshops also allowed the participants to give feedback on the ideas and features that had been mocked up as part of the design activities. The detailed descriptions of the methods, procedures, and outcomes of these workshops have previously been presented in a separate publication [78].

Research Seminar

Next, a second 3-day research seminar was hosted with the project team and the same research experts from the fields of behavior change, positive psychology, eHealth, and participatory care taking part. As with the earlier seminar, the user representative was the sole participant from this group. Still, we also took great care to communicate and represent the perspectives and inputs stemming from user participants through all workshop activities. The seminar's goal was to discuss and gain feedback on the undertaken activities and their outcomes to confirm these and to prioritize features and functionalities for the tool. A third focus of the seminar was to inform and guide the team in the design and content development of the strengths exercises to be included. As with the seminar in phase 2, a range of activities such as presentations, brainstorming, and discussions were employed. To further the participants' understanding and appreciation of the user ideas from the earlier activities, we also used tasks from the co-design workshop, with participants splitting into smaller groups. The research seminar was audio recorded, and the recordings as well as participants' notes, photographs taken, and drawings were used to create a detailed summary and a listing of key decisions, discussions, and takeaway points.

Results

Co-Design Workshops

In total, 22 people with various chronic illnesses, in 3 locations, participated in 2 consecutive co-design workshops (see Table 2 for participants' age and gender). The workshops revealed further requirements from the user group in the form of numerous suggestions for how game-like features and elements could be designed to create a strength-based eHealth and mHealth tool.

Table 2. Participants in co-design workshops.

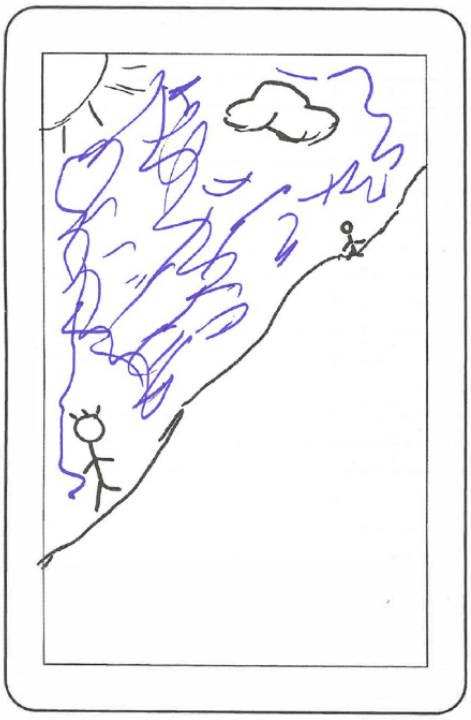
Group number	Participants, n	Age (years), range
Group 1	7 (2 males)	17-21
Group 2	7 (2 males)	21-58
Group 3	8 (3 males)	27-64
Total	22 (7 males)	17-64

In the workshops, the participants created posters and wireframes with their ideas for using gameful designs in strengths-focused mHealth apps. During the second workshop, they also rearranged and commented on simple mock-ups of the app created based on output from the first workshop. As with the idea-generating workshop in phase 1 (discovery), the importance of keeping the user experience positive and supporting emerged as one of the most important requirements. Typical and proverbial gameful design elements such as points, competitions, or trophies [37] were also often suggested. The participants also voiced several concerns, for instance, the need for the new tool to be experienced as motivating, yet not addicting, by overly focusing on scoring points and winning.

The design workshops also yielded feedback and input into using mHealth to support people in using their strengths more. For instance, one group suggested a strengths assessment system in which the tool narrows down one's strengths by asking a series of questions and then suggesting strengths one might have based on the answers.

Some of the groups created complete concepts for self-management apps. For instance, Figure 4 presents a sketch of a progress tracker in the app, visualized as the user and a friend competing to reach the top of a mountain. Other concepts further provided the user with opportunities to collaborate and share experiences or activities with friends as well as the option to set increasingly hard goals and targets. In terms of collaboration, it was also suggested that one could send anonymized small predefined texts or icons to other users as an uncomplicated way of creating social features while maintaining user privacy.

Figure 4. User sketch of idea.



Research Seminar

Features and Functionalities

In the research seminar, the proposed list of features and functionalities was thoroughly discussed in light of previous experiences and literature as well as with regard to the outcomes from the co-design workshops. One of the chief decisions made was to focus the remaining work on functionalities directly related to the user's strengths and the use of these, rather than features providing information or social interaction. Providing information was discussed and dropped as users, in general, were considered to already have considerable access to relevant information fitting their specific needs. As such, creating relevant content surpassing what is already available to the individual users was considered to be very time and resource intensive. This has also been reported by participants in the design workshops and is additionally supported in the literature [79].

The social features were cut primarily because of regulations from the privacy protection committee at our institution. To log in to patient-facing services that have social and/or sharing capabilities, even if this information is anonymized, these regulations mandate the use of a level 4 two-factor authentication system. These are, in Norway, available from a limited number of approved providers and are typically used to log in to banks and public services. Both the researcher's experiences and direct statements from participants in our various workshops pointed to this being overly cumbersome and would lead to the app not being used much. As such, designing ways to create social interaction between users was deemed challenging and resource intensive, and we were forced to drop these capabilities.

As a way of supporting users to easily use their strengths more, it was suggested to link specific strengths to goals and then have small, more easily achievable subgoals. It was proposed that this connection could be made by starting with a goal and then finding a strength or simply by starting with a strength the users want to use. The overall goal of this is to create a list of the personal strengths the user has. The app should also present the user with suggestions for how they could use their strengths in what the project team described as strengths exercises, in case the participants did not come up with activities themselves. The tool could then provide the user with feedback and positive reinforcement on the goals accomplished. An underlying goal should be to change the user's habits in small steps, as typically done in behavior change interventions [80], and help them gradually use their strengths more. To further contribute to users' well-being, it was also suggested that the app also should feature gratitude exercises, such as the *three good things* exercise [10,14].

Design

In terms of design, the second research seminar resulted in the overall idea of designing for gameful and enjoyable interactions that could help users learn for their own experiences and be more active in their self-management. During discussions, particular focus was placed on the users' first interaction with the tool and how this should be designed to simultaneously explain the rationale behind a strengths focus, introduce the app, and motivate the users to start exploring their strengths. For example, it was concluded that using avatars could be a positive source for motivation, as it could provide users with social motivations and support and, in some cases, allow the users to visualize a better self. Some suggestions included designing avatars to fulfill the role of either a guide or a companion or a narrator to the app (for instance, a friendly avatar to be one's climbing companion) or as a virtual representation of the users (designed by the users themselves during the initial use of the tool). Another way proposed to

provide relatedness with the app could be using videos (animated or real life) that could present the concept and rationale of strengths or compelling user stories.

The use of different metaphors in the app was proposed and discussed. Among other ideas, an expansion of the user-generated *climbing mountains* concept from the co-design workshop was suggested. Here, the user would still have climbing mountains as a goal, and the way to get to the top would contain several stops and base camps where the user could reflect and take stock of the tools (strengths) they are using for the expedition. Other ideas were to theme the app as a journey of discovery. Still, after discussion, a decision was made that a potential metaphorical theme or approach for the tool should be culturally neutral and not focus on metaphors that, for instance, are well known by some populations but can be unfamiliar to others. More traditional game features, such as points and unlockable content, were also suggested. However, a consensus opinion was that these should only be implemented if they could provide further value to the tool as a whole, as opposed to merely adding points for the sake of gamifying the tool.

At the end of this phase, the feature list for the *MyStrengths* tool was updated by removing, adding, and reorganizing features into a more detailed list. Specific ideas and features suggested by users and researchers or identified in the literature were added and annotated with its source. Although the overall categories were supported and suggested from users, researchers, and literature, the specific implementations of these often differed. For instance, at the end of an exercise, users suggested performing a reflection task on the difficulty of doing the task, whereas the researchers suggested performing a reflection task concerning how you had used your strengths to complete the task. When completed, the list (presented in Textbox 2) covered 4 main sets of features.

Textbox 2. Intervention feature overview list following phase 3.

My strengths

Assessment and overview of the user's strengths using a predefined list of strengths

Strengths exercises

Examples of exercises the user can try out on one or more of his or her specific strengths to try out to apply these in his or her daily life in a new way

Strengths experiments

Small activities that can serve as building blocks on the road to achieving the user's goals. Here, the user would first outline a goal and then plan out several small activities, or experiments, that could help toward reaching it

Daily log

Section of the app with 2 separate features. First, it allows the user to rate how the day has been and write down the *three good things* exercise. Second, it collects and visualizes the user's inputs and activities throughout the app and finally provides a summary for each day in an easily scrollable interface

Phase 4: Delivering

With a core set of features, the fourth phase of the project consisted of the final design and the technical development of the *MyStrengths* app through workshops with designers, development with our IT department, and iterative evaluations from users.

Methods and Activities

Designer Workshop

With the 4 primary features for the app (ie, strengths assessment, strengths exercises, strengths experiments, and daily log) as a basis, a 2-day design workshop was held to condense the gathered inputs and create a central concept for the tool. In addition to the core members of the project team, our in-house designer and one developer as well as 4 external designers experienced in designing for health and behavior change participated. For these workshops, inspiration was drawn from how *game jams* are organized [81]. Starting with a thorough description of the activities undertaken thus far, the participants worked in smaller groups. During the second day, a concept emerged, and the latter part of the workshop was spent on collectively elaborating and embellishing this concept.

Strengths and Strengths Exercises

With previous research on personal strengths at our research center as a basis [16,21], a list of 40 strengths to be included in the assessment part of the *MyStrengths* app was created. Connected to each of these strengths, we then created matching strengths exercises. These exercises are activities that the user can perform to employ a specific strength. For instance, if one person has kindness as a strength, the app could suggest for him or her to do something kind to a neighbor today. Most of these exercises were based directly on strengths exercises found in academic and popular literature [10,14,82-84] or related to items reported by participating patients in earlier related studies [16]. The wording and content of these exercises were iteratively refined during the user evaluations of the app.

Iterative Development—User Evaluation

Building on the outcomes from the workshop with the designers, the research team worked closely with the in-house designers and developers to sketch out and create mock-ups of the specific features of the app.

Over 4 iterations, with more and more of the features added, the app was evaluated by people living with chronic illnesses. On 3 occasions, colleagues at our center experienced in eHealth and mHealth development, and new to the tool, also evaluated the app using the same setup and methods as the evaluators from the user group. On the basis of feedback from the

evaluations, adjustments were discussed and made by the research and development teams.

All user evaluations were audio and video recorded (see Figure 5 for an example of the video frame). The first author (SJ) hosted the evaluations, whereas a member of the project team observed, supported the host when needed, and took notes. During evaluations, participants were not given instructions but were told to navigate through the app freely and describe their impressions and actions by thinking aloud [85]. Having gone through all the features available, the user and the project members openly discussed the experience and any thoughts or ideas the user might have. Both the facilitator and the observer wrote detailed notes. The first author rewatched the recorded evaluations and, drawing on these and the notes, made detailed reports from each evaluation. Individual reports were combined and presented to the research team and developers. This group then discussed the findings and ideas from the evaluations and then decided on what changes to make to both existing and planned functionalities and design.

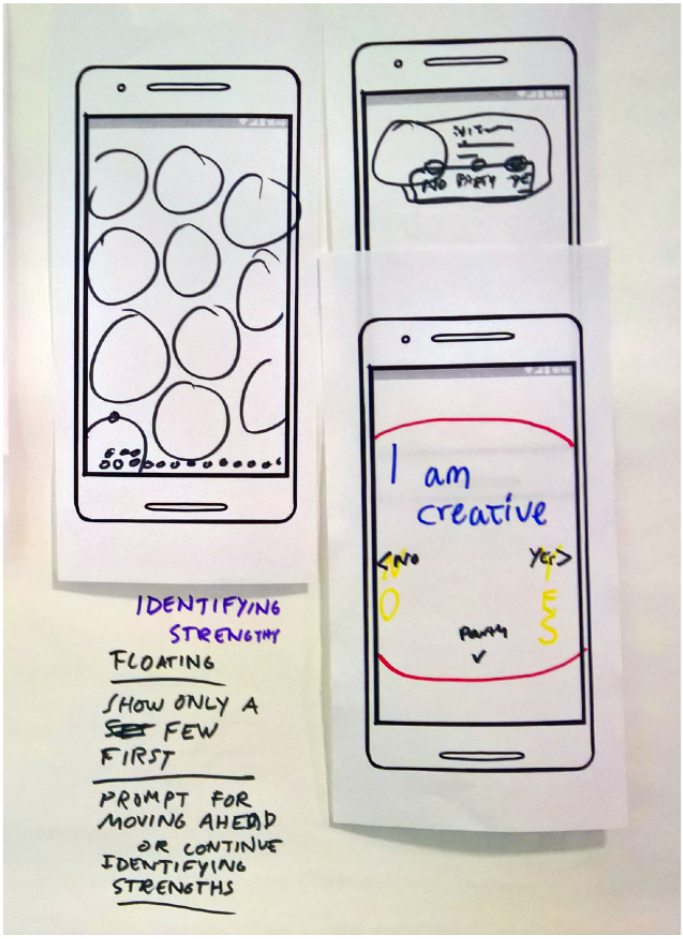
Figure 5. User evaluating the app (capture from video).

Results

Designer Workshop and Iterative Development

During the designers' workshop, on advice from researchers and experts, it was decided that concrete metaphors should be avoided. Following this, we decided on a concept with spheres visualizing the users' strengths floating on the home screen. See Figure 6 for an early sketch of the home screen. By putting the users' own strengths front and center, simply opening the app could serve as a positive reminder of all the strengths the user has. When starting the app for the first time, a single sphere would float up to the top of the screen. The users can then click on and rate the strength as having, partially having, and not having. The app includes 40 strengths, and based on the ratings given, the spheres have different colors. To emphasize the strengths the user has, the ones rated as having would float on top. Under these, the other spheres of the partially having would float, and at the bottom, the ones the users rated as not having. Unrated spheres would float up toward the top to vie for the users' attention and get him or her to rate them. Clicking on one of these spheres would then *open it up*, and the user would have the possibility to do exercises and register reflections and thoughts concerning the strength. Users are also easily able to add their own strengths and exercises should they want to.

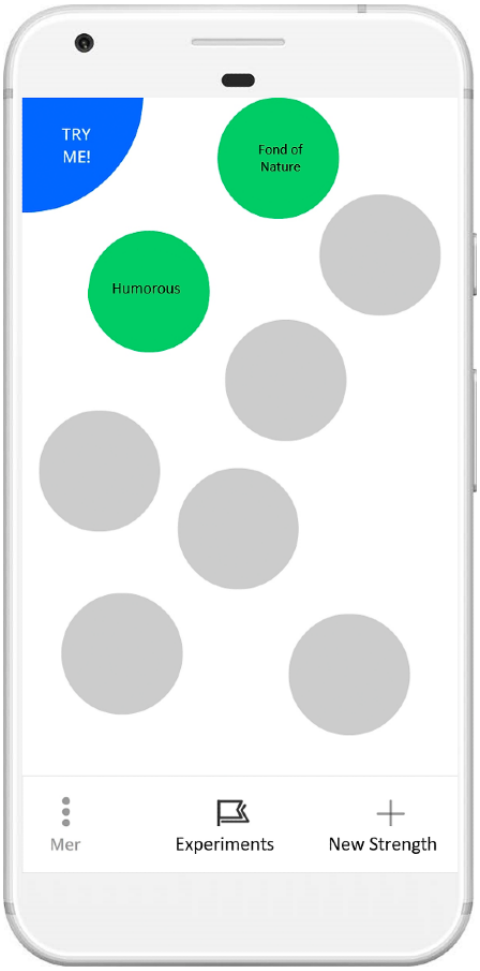
Figure 6. Home Screen Sketch (left).



Another key design decision from this workshop was to minimize the focus on the number of strengths a person has or the number of exercises done. This is because part of the rationale for a strengths focus posits that one uses strengths one already has [17]. As such, gaining new strengths is not a primary goal. To keep track of what users do in the app, we created the *daily log* section where each day's activity and input would be presented on cards listed onscreen. The log would additionally provide the users with the *three good things* exercise from positive psychology [10], daily asking the users for *three good things* they had experienced that day.

To engage users, time was spent discussing *novel* ways to interact with the app, and it was suggested that if the users shook the phone, this could, for example, result in the spheres on the screen moving around or prompting the users with a randomly suggested exercise. Although the *shaking idea* was eventually discarded because of both technical and design challenges, the concept of suggesting random exercises was refined through several iterations. Figure 7 presents a snapshot of an early interactive mock-up of the home screen with a *try me* button in the top left corner. Pushing the button, the app would then suggest an exercise for one of the users' strengths at random. In the final design, this feature took inspiration from the world of games and the universally known roll of a dice, in what became the *dice* feature. In addition to surprising the users with randomly selected exercises, this randomness could possibly also provide users with new exercises very easily.

Figure 7. Home screen interactive mockup.



After several iterations of both paper-based and interactive mock-ups, our internal IT department started the development of the *MyStrengths* app based on the Unity [86] game engine.

User Evaluations

In 4 iterations, 13 users and 3 of our colleagues tried and evaluated the app. Table 3 presents the features added and evaluated, the number of participants, and their age and gender.

Table 3. Development iterations and users (N=13, 1 male).

Iteration number	User evaluations, n	Age (years), range	Internal testers, n	Features introduced in the current iteration
1	4 (1 male)	22-62	2	• Introduction Pages • Home screen • Strengths assessment • Strengths exercises (Marvel-app mock-ups)
2	4	20-50	0	• Strengths exercises
3	3	51-59	1	• Daily log
4	2	46	0	—[a]

[a]No new features added.

Overall

In sum, the feedback from the user evaluations was quite positive, and everyone reported to like the concept. However, the users also reported that the *daily log* section of the app was not particularly fun or engaging to use, and in a few cases, they even stated that they expected something *more fun* when clicking on it. Following this feedback, several of the participants contributed creatively with suggestions for improvements, and during the evaluations, they suggested adding a more game-like way to create goals and write entries in the daily log. Some of the evaluators also wanted to add pictures from the phone in the daily log. Due to all content created or added in the app being encrypted and stored locally, we could not add this

functionality as it could possibly make the app's data use on the devices enormous. Several participants suggested adding more visual flair overall, for instance, by using more animations and interactivity throughout or by adding avatars to represent you in the app. Many evaluators reported they were expecting to receive some notifications from the app, for instance, to remind them to do exercises or simply reminding them of their strengths. It was also suggested that the app in the notifications could ask users how they feel each day and then provide appropriate feedback, for instance, by cheering the user on if the day was good or reminding them of better times or their strengths if the day was bad.

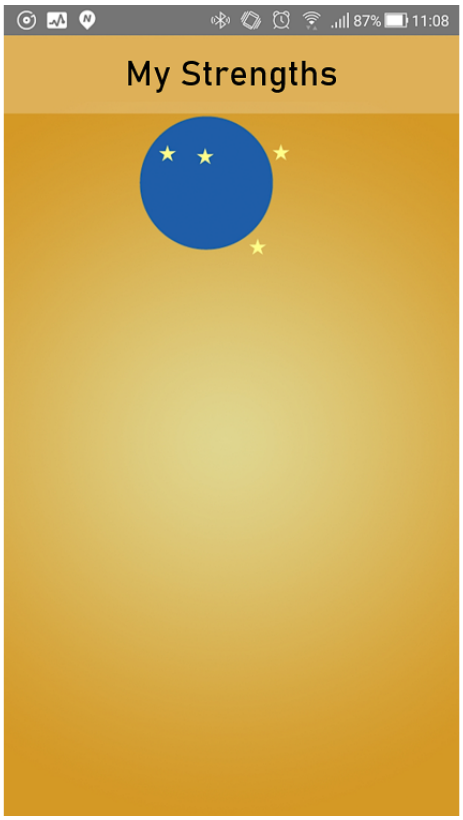
Strengths

When trying the strengths experiments, several participants reported that they would instead work on gaining the strengths they were lacking than focusing on using the ones they had more. When asked about this, one participant explained, "Because this is how we are always taught to think." Over the different iterations of the app, both the wording and style of the strengths exercises were subject to significant work and redesign. For instance, most users preferred exercises that were concise and simple to do (ie, not involving many steps of different actions needed). Along the same line, many users in the earlier iterations reported that some of the exercises were too complex and time consuming. Several users also commented on the impracticality of the strengths exercises asking them to write plans and thoughts down on paper, and they would instead prefer to do this on the phone itself. A large portion of the users also preferred exercises that were physical rather than cognitive or mental, such as "surprise a friend with something nice today" as opposed to "Sit down for 15 minutes and think about the good things you have in your life."

Usability

In terms of user friendliness, we experienced during the evaluations that the intuitiveness of some sections the tool was not satisfactory. For instance, when the first sphere was shown on the screen, as shown in Figure 8, most participants did not click on or try to interact with it. This was redesigned, so that *unopened* spheres would pulsate to attract the users' attention and indicate a possibility for interaction. Although we follow guidelines for universal design, a user with reduced dexterity still found some buttons difficult to hit. These were redesigned to be larger. We also encountered multiple instances of users not fully understanding icons and buttons. For example, on the page for performing strength exercises, one can choose between suggested exercises or add one's own by pushing a button with a plus sign (+). During the evaluations, few of the evaluators used this button, even if prompted to add new exercises, and it was both redesigned and placed more prominently on the page.

Figure 8. Screenshot of first strengths sphere.

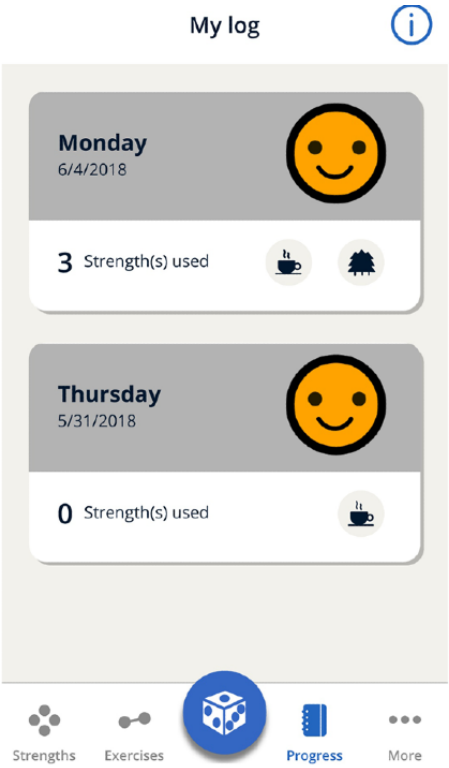


There were also multiple technical issues discovered during the evaluations. In the earliest iterations, the app would freeze if users tried adding emojis as part of textual input. We also found an issue where the strengths spheres that should be *floating* to the top of the screen instead sank to the bottom of the available screen space. Overall, more than 120 design and usability issues as well as numerous bugs and technical problems in the *MyStrengths* tool were identified over the different iterations. The issues were redesigned and fixed between iterations.

Work With External Game Designers

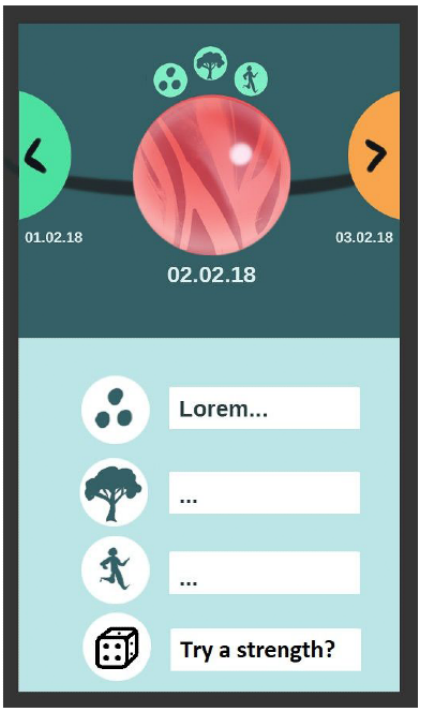
As feedback from user evaluations as well as the project group indicated a need for a more playful experience when using the *MyStrengths* app, we decided to consult with an external game designer to help expand on the gameful aspects of the daily log functionality, which at present visualized summaries of each day's activities and registrations as *cards* (Figure 9).

Figure 9. Daily log screenshot.



Working with the designer yielded a detailed concept description and rationale as well as mock-ups of the further developed daily log section of the app, an example of which is presented in Figure 10. The concept of spheres was also central to this, and entries into the daily log would look like pearls on a string one can swipe between. Like in the home screen, clicking on a sphere would *open it up* and show the details of that day's entries. Hovering above each sphere would be small icons visualizing different activities registered during the day.

Figure 10. Daily log new concept sketch.



Although the project team found this concept to hold great potential, we were not able to implement these designs into the app because of both administrative and resource-related challenges as well as the need for the app to be ready in time for its feasibility trial. On the basis of the outcomes of this trial, we have planned to make adjustments to the app before making it generally available, and we aim to implement the new design concept during this period.

The Final MyStrengths App High-Fidelity Prototype

The finished high-fidelity prototype of the *MyStrengths* app centers on a list of strengths spheres floating around on the screen (Figure 11). These are colored red, yellow, blue, or green and signify which strengths one thinks one possesses, strengths one partly possess, do not possess, or have not yet assessed (or do not find relevant). Although the apps ask users to rate the 40 strengths, the user can also add as many new strengths as they please. From this screen, the user can access the 2 other key features of the app: strengths exercises and reflections and the daily log. Clicking on a sphere opens it up and provides the user with a list of suggestions for small exercises and the opportunity to create new exercises themselves or a note-taking area for writing small reflections on how this strength can help them in their lives (Figure 12). Active or completed exercises are listed in the exercises menu at the bottom of the screen. The daily log asks for an entry each day and when doing this first asks for a rating of how the day was, using 5 smiley faces (from sad to happy), then asks the person to write 3 good things that happened and pick icons for these (Figure 13). For each day, the log puts all this information into a summary and represents each daily entry as a card on a scrollable list (Figure 14). The dice at the bottom center of the screen randomly selects a strength exercise from the strengths that the user has and suggests exercises for the user to use this strength (Figure 15). A secondary goal of both the dice and the opportunity to add own strengths and exercises is that the app can provide users with new and interesting experiences, despite its basic set of features and content.

Figure 11. *MyStrengths* Home Screen.

Figure 12. Strengths Exercise.

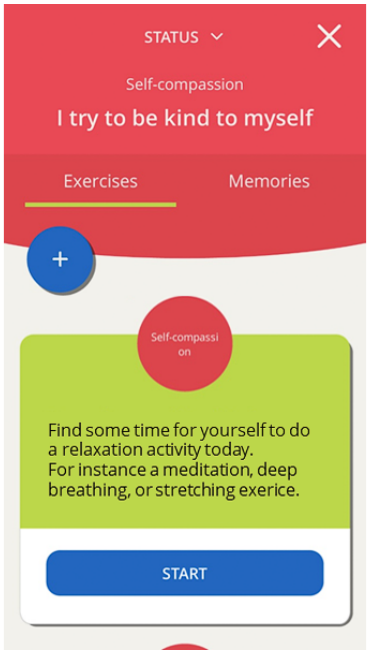


Figure 13. Rating the day in the daily log.

Figure 14. Daily Log.

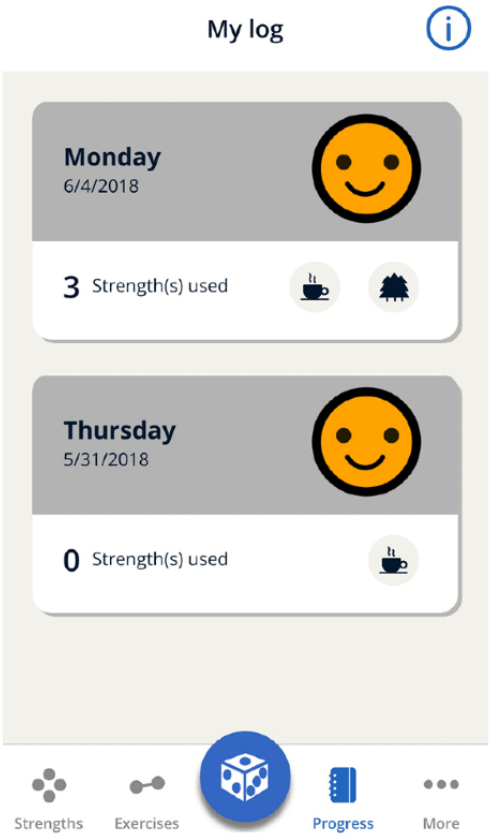
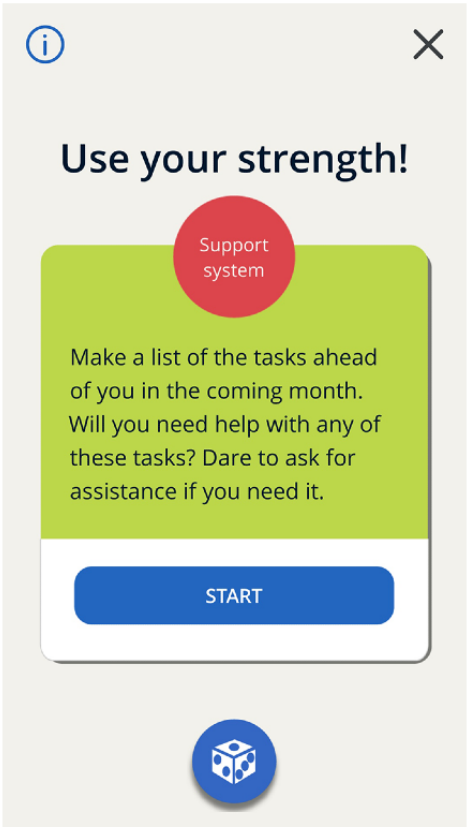


Figure 15. Dice Suggestion.



Privacy and Data Security

Maintaining user privacy is of high importance. As presented in phase 3, the privacy protection and data security committee at our institution has strict guidelines that led us to discard the interaction between users of the *MyStrengths* app. As such, users will not be able to identify one another. The app is planned to be available through both Google Play and the Apple App Store, and besides having an account in these stores, there is no need for the users to identify themselves or register before using the app. Although there are slight differences between data available to developers on Google Play and Apple App stores, the most detailed information on users available would be an aggregated number of users with different operative system versions. All data generated by users are encrypted and stored locally on the device. When launching the app for the first time, users create a four-digit pin code needed to access the app. A key for data encryption is generated, and this key is used to encrypt data before being stored. The encryption key itself is stored after being encrypted using a temporary key derived from the user's pin code. The app also conforms to the standards set by the European General Data Protection Regulation.

During a 4-week feasibility trial, the app will send usage data over a secure connection to the *Service for Sensitive Data* at the University of Oslo. Once the trial is completed, the registration and transmission of usage data will be disabled, and the app will only run locally on the users' device. The feasibility trial has also been approved by the privacy protection and data security committee.

Discussion

Principal Findings

The availability of strengths-focused mHealth apps or guidelines for their design is scarce. This paper contributes to this field by presenting, in detail, the various activities, decisions, and outcomes from the design of the *MyStrengths* app. We have presented 3 types of findings stemming from the process of development:

1. From individuals in the target group who took part throughout, we have gotten both ideas and design requirements for strengths-focused mHealth tools. Chief among this input is the need to have such apps thoroughly focused on the user themselves and if using gameful designs to keep these mostly noncompetitive and positively oriented.
2. Having reviewed the existing literature on strengths-focused self-management interventions, we found a general lack of existing guidelines or design descriptions for strengths-focused mHealth tools. However, goal setting might be a productive way through which one might help people find and use more of their strengths.
3. Researchers and designers contributed, among other things, with knowledge and input on how strengths activities can be designed and connected to the users achieving set goals. This group also took part in prioritizing between, and merging, the various features and parts into the final *MyStrengths* app prototype.

From our experiences on this project, we would like to further discuss a few points and present some recommendations for others, creating strengths-focused and positive mHealth tools.

Gameful Designs in a Positive and Strengths-Focused Environment

Recent literature reviews of gamefully designed eHealth and mHealth tools report the most popularly used game elements to all to be externally oriented: points, rewards, and leaderboards [36,37]. Still, through activities involving both users and experts, it was repeatedly suggested to be cautious with such elements, and during the co-design workshops, users voiced a specific dislike for designs with reward schemes that facilitate *addictive* use. This is similar to findings from the study by Ahtinen et al [46], where participants reported chasing rewards unfit for mindfulness exercises. However, in a review of apps promoting well-being and mindfulness [38], the authors found the archetypical game elements of points' badges and competitions to continue to be the approaches mainly used.

Not overly using externally oriented motivational features, such as points and rewards, has also been discussed in more theoretical works on engaging and gameful designs [45,87,88]. These also highlight the importance of not only playing the game mechanics to win but to do the activities or tasks for the *right reasons*. This seems especially relevant for the *MyStrengths* tool, with its focus on the user's reflections and awareness concerning themselves and their situation. As such, we decided to *tone down* the focus on gameful features that foster competition and aggregation of points or rewards and instead focus on engaging users with pleasant and positive user experiences.

Combining both social and goal-directed game elements into features such as competitions and collaborations is also very popular in gameful designs [37]. In addition, their potential for providing social comparisons and role modeling can make such features valuable in the field of mHealth [36]. However, users taking part reported concerns regarding visibly losing to, or receiving negative communications from, other users. To address this issue, the users during the co-design activities proposed ideas such as one-way communication using predefined texts or elements as a way of ensuring a positive focus. Anonymized social interaction or competition might be one way of leveraging the positive and motivational aspects of social interaction. This would also allow the user to avoid any perceived obligation of reciprocity or the possible stigmas of losing to one's friends. An example of such an approach is presented by Mylonopoulou [89], who created leaderboards for progress in a mHealth tool, in which the user was shown and compared with 3 other anonymous users who were slightly better than them.

A gameful feature that, although not highly used, still holds potential is randomness [38]. During this project, we developed the *dice* feature that randomly selects strengths exercises for the users with the press of a button on the app's home screen. The rationale for using randomness as a design element is that it can induce a sense of variety and anticipation of *what the future holds* for the user [38,90,91]. Randomness can be suitable for tools such as *MyStrengths*, in which there is no set route or user journey through the app. However, it is likely more

challenging to include randomness in tools that are based on strict treatment regimens such as cognitive behavioral theory, in which the structure and order of activities is important [38]. The *dice* feature was well received by users in the evaluations, and we consider such approaches to hold great potential for providing gameful or playful experiences in an overall positive and gentle fashion. Thus, in technical terms, the design approach for this project can be described as focusing more on designing for *playfulness*, free play, and exploration, rather than *gamefulness*, which, in addition to free play and exploration, is more rule structured and often concerns a pursuit of points or scores [48].

Strengths Approach to mHealth and eHealth

At the core of the strengths concept is changing the focus from deficits to the positive to achieve better well-being or happiness [14,17]. In keeping with this approach, emphasis should be placed on using existing strengths rather than focusing on turning non-strengths into actual strengths. Even so, during our user evaluations, several participants reported that they would instead work on gaining new strengths, as opposed to focusing on the ones they already have. This way of thinking is not that surprising and falls in line with the typical deficit focus of health care in general [92,93]. Therefore, future self-management tools using a positive or strengths approach should consider how to support users in shifting this way of thinking. In *MyStrengths*, the *dice* feature, which selects exercises at random, only draws exercises related to strengths that the user has assessed as having and thus guides them to build on their already identified strengths. This approach can be described as a way of nudging [94], a form of altering people's behavior without forcing their options or activities. Furthermore, as an additional way of emphasizing the users' existing strengths, the *MyStrengths* app sorts the strengths spheres so that the ones already marked as a possessed strength by the user are at the top and visible first when the app is opened.

Although assessment of people's personal strengths has been done before (for instance, through tools or services such as *VIA Character strengths* [14] and *StrengthsFinder 2.0* [95]), these assessments have been made for the general population, not specifically for people living with chronic illnesses. This project, therefore, used a set of personal strengths that have been reported by, and found important for, people living with chronic illnesses [16]. Still, people's strengths vary from person to person, and to improve the personal relevance of the tools for the users, we would also recommend to do as in the *MyStrengths* app and allow users to create their own strengths.

Despite previous research showing that the strength concept can be challenging for people to grasp, particularly when referring to one's own health [96], users participating in co-design activities as well as evaluations of the tool mostly understood the strengths concept quite well and were able to relate to it. One the other hand, users who took part in evaluating the first iterations of the app reported finding the strengths exercises difficult to grasp, and most of these participants also preferred the more physical and tangible exercises (such as "do something nice for a friend today") over exercises that were either more cognitively challenging or consisted of many steps (such as "go to a library, borrow and read this book, then think how your journey as a patients likens the hero in the story"). On the basis of this type of feedback, the exercises were redesigned to be simpler to both understand and perform.

Privacy and Ethical Considerations for Designing mHealth Tools

During all design phases in this project, users participating discussed privacy and the different ways in which the features of *MyStrengths* could affect this. For instance, during the co-design workshop, multiple participants talked about how one's strengths were a very personal thing and something one might not want to share with anyone. Similarly, in the idea-generating workshop, several participants also voiced concerns about privacy and suggested using nicknames in communication with others. Clearly not wanting to share this information about themselves, one participant simply said, "I would never put something like this on Facebook" [78]. Participants in the user evaluations also highlighted concerns regarding privacy, and many would ask the facilitator whether anyone could view the information they entered into the *MyStrengths* app. As such, we can surmise that although people living with chronic illnesses, for instance, often are active in support and interest groups on social media [79], they should also likely be the ones to control what, if anything, to share.

As voiced by participants throughout this project, losing is no fun when it is owing to your own medical situation. Thus, it became important to make sure the *MyStrengths* app is respectful to users changing shapes, and we avoided using gameful design techniques such as awarding users for streaks, that is, using the app according to set schedules or intervals. Furthermore, when creating engaging, gameful, or persuasive designs, designers are steering users toward a desired action. In doing so, one should be careful not to steer users toward making choices that are against their interest or will or create a very limited set of perceived choices for interaction [97,98]. In *MyStrengths*, the *dice* only draws exercises for strengths that each person assesses themselves as having and will thus not suggest exercises for strengths that the user considers herself or himself not to have. In the *MyStrengths* app, this is explained to the users when going through the first-use tutorial of the app. However, users still have the possibility of selecting activities for all strengths in the app through the home screen. As such, it is important to consider the ethical aspects of mHealth tools. This can, as discussed above, mean not metaphorically *forcing the users hand* too much or using game design techniques that, for instance, are experienced as inappropriate, unfair, or even trivialize the user's situation [45,47,90].

Participatory Approach

In creating mHealth tools, it is highly beneficial for the design group to include researchers and/or designers with health, design, and psychology experience or background as well as actual representatives of the user group [51,52,97]. In this project, both the diverging phases started with involving users in a broad way to create ideas and suggestions before the converging phases relied more on evidence and input from researchers and designers to narrow the initial input down to concrete and implementable solutions. From their participation,

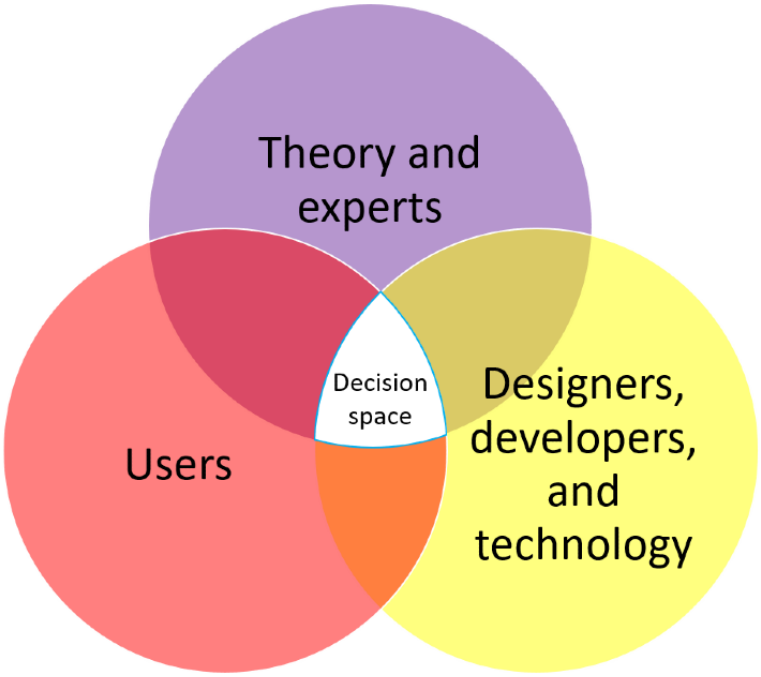
we see that the users contributed valuable contributions such as (1) emphasizing a liking for a gameful design but not necessarily when it consists of stiff competition and possibly losing; (2) providing important input into the design and wording of the strengths exercises; and (3) how the *MyStrengths* app should fully be about the users, putting them at its center.

Although users did not actively take part in most of the converging phases, we were always focused on correctly representing their contributions and needs. In general, user inputs and concerns were maintained through 3 different strategies: (1) having a user representative, who is equal to everyone else in the project group (this person, it is important to highlight, works as a nurse and has a medical or nursing background, something that may also reduce the hierarchical distance between them and the researchers); (2) using the outcome from the co-design workshops and user evaluations in the discussions; and (3) having all project members and researchers do the same design tasks as the participants in the co-design workshop. This third strategy also supported the important task of communicating and creating a deeper understanding of the users' needs and requirements within the project group [99]. This allowed the group to see not only the outcomes of the participants' work but also to go through the process of creating similar types of work themselves.

In terms of the nonuser participants, a range of professions have been involved, including researchers with health, psychology, and informatics background and expertise; developers; designers; and game designers. Each of these participants contributed to broadening the group's repertoire of possibilities with respect to the design of the *MyStrengths* tool. To name a few, (1) the psychology and health care researchers provided important information and feedback regarding the strengths approach, (2) the developers and designers contributed creatively in putting everything together into a working and user-friendly tool, and (3) involving the external game designer provided suggestions for more interactive and immersive solutions.

As such, the outcomes of this project can be considered to stem from a space between 3 different groups (see illustration in Figure 16): (1) theory, evidence, and researchers' knowledge; (2) designers, developers, as well as the limitations and opportunities of the technology itself; and (3) the users, with their needs and ideas. Although the power of decision in this project was mostly placed with the researchers and experts, most decisions were the result of a negotiation between the 3 groups and their inputs into the project. Even though no users besides the user representative took part in the 2 converging phases, following the 3 strategies described above still allowed us to maintain and represent their needs, ideas, and requirements in the decision-making processes.

Figure 16. Venn diagram of decision space.



As shown by the range of outputs from the idea and design workshops, users are indeed capable of creating ideas and concepts with great potential. Although we thoroughly maintain the user's inputs and ideas throughout the project, they are not able to contribute new input without taking part. This could, in particular, have changed the way in which we implemented and combined ideas and concepts stemming from the users in the first place. Still, users have been included and given the power and opportunity to create ideas or suggest ways to use the tool and to evaluate and suggest improvements to the design, the 2 aspects of a project that have the *strongest possibility for participation* [100]. As presented in this paper as well as in previous publications from this project [65,78], we have shown how the participants in our co-design activities contribute creatively and productively to both the content and design of the *MyStrengths* app.

Lessons Learned From Creating *MyStrengths*

During the process of developing *MyStrengths*, we collaborated with a multitude of different stakeholders in a wide range of

activities and processes. From these, we wish to highlight some points and issues that may be of interest to others developing positive or strengths-focused eHealth or mHealth tools for people living with chronic illnesses.

(Try to) Involve Users in All Activities

When projects clear the proverbial *fuzzy front end* [54] and finally start building the actual product, the pace of the work tends to increase. In earlier phases of the project, recruiting participants to activities was often done in collaboration with the institutions participating in the project's activities. By recruiting only for specific activities, we had to reach out again during later phases and re-recruit the same participants. This made organizing the recruitment and activities for participation time consuming and thus more challenging than envisioned. For instance, we would have benefited from additional design activities with users, both in connection with the designer workshop and in response to the user evaluations in phase 4. To achieve this, we could have recruited participants early in the project into a pool of available participants that we could easily contact to work more closely with as the project moved on. Although all design projects are highly dynamic, and schedules and priorities are often subject to unforeseen changes in resources, time, or technology, we would recommend creating a systematic plan for recruitment and participation for the entirety of the project.

Use User Representatives

Having a user representative as a member of the project group helped us in keeping touch with the users' perspectives and challenges throughout the work, even when the users were not participating. The user representative in our project was a full-fledged member of the group and took part in all activities and decision-making processes. For others employing representatives, we would also recommend including these as much as possible in activities and discussions to ensure they stay active and up to date and not end up passive sources, which the rest of the team only taps for information when needed. In addition, by the representative being a full member of the team, and not merely invited when needed, he or she gets to know the other project members better, and hierarchical differences in power are hopefully diminished.

Ideate Freely

We had to abandon the implementation of social features, primarily because of restrictions from the privacy officers at our institution. Although we could have foreseen this issue and removed the possibility of social features in our activities, we might have restricted the participants' creativity and range of possibilities. Furthermore, by working without restrictions during the workshops, participants also give important information and feedback that, for instance, are relevant not only to social features but also to mHealth tools as a whole.

It may seem more efficient concerning time and money to keep participatory activities focused on what is possible or advisable to create. However, we still recommend allowing for free creativity and ideation in such activities, as this can yield not only interesting ideas and concepts but also valuable insights into the user group and their needs and wishes.

Participatory Approaches in the Face of Incomplete Guidelines

Participatory approaches should be well suited in design situations where the aim is to create something new in the face of a shortage of guidelines for designing. This project aimed at both exploring ways of creating an engaging and strengths-focused mHealth app as well as actually developing one, *MyStrengths*. It is our position that those goals together make participatory approaches well suited. In addition, as our secondary goal is to explore opportunities for designing strengths-focused tools, it seems right indeed to include users, primarily as they should be considered to be experts of their own situation, and additionally as they can *widen the design space* [101] by, for instance, contributing ideas that the researchers and designers would not think of. Thus, users can also contribute productively in situations with solid evidence forming the basis of mHealth tools. As presented earlier, even how text is written and phrased benefited from user input. Thus, we would recommend considering all forms of creating mHealth tools, even small projects that, for instance, translate an existing tool, processes in which user participation can be of great value.

Strengths and Limitations

The activities as well as outcomes presented in this paper can serve as a foundation for future research on the development of strengths-focused mHealth or eHealth interventions for people living with chronic illnesses. It is important to note that because of the explorative nature of this project, any generalizations as to what designs or functionality people living with chronic illnesses enjoy or want is neither possible nor intended. However, based on both the quality and range of output presented in this paper, it is likely that our chosen activities and methods worked well and may be applicable to others as well.

However, there are also some limitations to this study. First, all users and stakeholders participating in the study volunteered to participate on their own or were contacted by the project team or collaborating institutions. As such, the participating groups are likely biased toward being more motivated, resourceful, and managing their life with chronic illness well. Thus, they may not perfectly represent the entire user group. However, this is too common for this kind of research. Using other means for recruitment, such as social media [102], might have eased access to harder-to-reach users, but this was not within the project's mandate.

Second, by aiming to design a mHealth tool for such a broad group (ie, people aged >16 years living with chronic illness), this project sought to design for a target group that is practically the entire population. However, creating features for specific or smaller groups was not the goal of the project, among others, as having strengths is something shared by everyone irrespective of illnesses, age, or background.

Third, the gender balance during most of the project activities was skewed toward female participants, with around only one-third of the participants being male. During the user evaluations, only one of the participants was male. During all recruiting, we tried to recruit more male participants, and unfortunately, we were never able to reach equal number of

males and females. Similarly, both age and gender distribution among the participants changed throughout the project. However, having participants from different age groups also means that they represent the age range of the intended user group.

Fourth, regulations concerning privacy and data security made us decide to drop the inclusion of social functionalities in the *MyStrengths* app. Although necessary for us to do, this also meant that we discarded one of the more popular features suggested by participants throughout all our activities. Similarly, because of unforeseen challenges at our end as well as for our external collaborators, we were never able to implement the redesign of the daily log part of the app.

As it stands, some of the features and concepts suggested by users, designers, or researchers throughout this project for various reasons ended up unused. Technical challenges and restrictions or overruns on time and resources are not uncommon in these kinds of projects. Many of the technical challenges faced are a product of us trying out new designs or forms of interactions, such as the strengths spheres floating on the home screen or the initial idea of shaking the device to move spheres around. Although time consuming, this experimentation is also in line with the projects' goal of exploring new ways of creating strengths-focused and engaging mHealth tools. The design of strengths-focused mHealth or eHealth interventions has rarely been reported, and both the activities and outcomes reported in this paper add important new information to this growing field.

We also wish to highlight the fact that the *MyStrengths* app was developed in a very collaborative way with members of the users and stakeholder groups. In fact, 2 of the participants took part in all activities with user involvement, idea workshops, co-design workshops, and user evaluations. Although caregivers only participated in the first phase, their perspective was not lost in the later phases, as several members of the project team have backgrounds in nursing and health promotion work. The repeated involvement of many users is an excellent strength to the project. It allows researchers, developers, and participants to create a thorough and deep understanding of each other and the tools being designed.

Everyone has their own strengths [14,17], and as such, the *MyStrengths* app has great potential in helping the increasingly large group of people living with chronic illnesses. Although the *MyStrengths* app holds great promise, we cannot, at present, speak to the effects or benefits of its use. However, a real-world feasibility trial with the *MyStrengths* app is currently ongoing, the results of which will be presented in a future publication.

Conclusions

Supporting people living with chronic illnesses in focusing on their strengths and positive resources can lead to higher well-being and quality of life. This project developed the *MyStrengths* app, a tool supporting its users to use their strengths more actively in their everyday life. As there currently exists little guidelines on developing mHealth tools with a strengths focus, this project took a participatory approach to create an engaging and playful mobile app to address this gap in research. During this process, we explored a range of ways of implementing strengths into mHealth tools and how to make such tools more engaging for its users. Although participatory design projects take time and are resource intensive, designing a tool such as the *MyStrengths* app based solely on the knowledge and ideas from literature and researchers is likely not a sound strategy as the user perspective is often lacking. In this project, we have shown how users can contribute productively to ensure that mHealth and eHealth tools being developed are both accepted and understood well.

Adding to the growing field of designing strengths-focused mHealth tools, this paper presents our approach for creating the *MyStrengths* app, which is based on evidence and established theories as well as created with contemporary design methods and with a high degree of user participation. Outcomes from, and methods used during, this project can be used as a starting point for future studies exploring strengths or strengths based on mHealth and eHealth tools. We kindly invite others to further develop, adapt, and build on findings, ideas, and activities presented in this paper for their own contexts.

Acknowledgments

The funding for this project was provided by the Norwegian Research Council (grant #248026). The authors would like to thank their larger project team, Cornelia M Ruland, Kurt Stange, Shirley M Moore, Luis Fernandez Luque, Hein De Vries, and Christina Bode. The authors also thank the staff at the learning and mastery centers at the Hospital of Southern Norway and the Nordland Hospital as well as the Youth Council at Akershus University Hospital for help and coordination. Finally, the authors would like to thank all participants who took part throughout this project for their time, creativity, and effort.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Brief overview of review findings.
[DOCX File , 30 KB - formative_v4i7e18049_app1.docx]

References

1. Liddy C, Blazkho V, Mill K. Challenges of self-management when living with multiple chronic conditions: systematic review of the qualitative literature. Can Fam Physician 2014 Dec;60(12):1123-1133 [FREE Full text] [Medline: 25642490]
2. Helsetilstanden I Norge 2018. Folkehelserapporten – Kortversjon. FHI: Folkehelseinstituttet. 2018. URL: https://www.fhi.no/publ/2018/fhr-2018/ [accessed 2020-03-10]
3. Global Action Plan for the Prevention and Control of NCDs 2013-2020. World Health Organization. 2013. URL: https://apps.who.int/iris/bitstream/handle/10665/94384/9789244506233_rus.pdf [accessed 2020-03-10]
4. Buttorff C, Ruder T, Bauman M. Multiple Chronic Conditions in the United States. The RAND Corporation. 2017. URL: https://www.rand.org/pubs/tools/TL221.html [accessed 2020-03-02]
5. Ambrosio L, García JM, Fernández M, Bravo S, de Cerio Ayesa SD, Sesma ME, et al. Living with chronic illness in adults: a concept analysis. J Clin Nurs 2015 Sep;24(17-18):2357-2367. [doi: 10.1111/jocn.12827] [Medline: 25951949]
6. Carrier J. Managing Long-term Conditions and Chronic Illness in Primary Care: A Guide to Good Practice. Abingdon, UK: Routledge; 2009.
7. 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]
8. Grady PA, Gough LL. Self-management: a comprehensive approach to management of chronic conditions. Am J Public Health 2014 Aug;104(8):e25-e31. [doi: 10.2105/AJPH.2014.302041] [Medline: 24922170]
9. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. J Am Med Assoc 2002 Nov 20;288(19):2469-2475. [doi: 10.1001/jama.288.19.2469] [Medline: 12435261]
10. Seligman ME, Steen TA, Park N, Peterson C. Positive psychology progress: empirical validation of interventions. Am Psychol 2005;60(5):410-421. [doi: 10.1037/0003-066X.60.5.410] [Medline: 16045394]
11. Niemiec R, Shogren K, Wehmeyer M. Character strengths and intellectual and developmental disability: a strengths-based approach from positive psychology. Educ Train Autism Dev Disabil 2017;52(1):25 [FREE Full text] [doi: 10.2307/26420372]
12. Grey M, Knafl K, McCorkle R. A framework for the study of self- and family management of chronic conditions. Nurs Outlook 2006;54(5):278-286. [doi: 10.1016/j.outlook.2006.06.004] [Medline: 17027605]
13. Schulman-Green D, Jaser S, Martin F, Alonzo A, Grey M, McCorkle R, et al. Processes of self-management in chronic illness. J Nurs Scholarsh 2012 Jun;44(2):136-144 [FREE Full text] [doi: 10.1111/j.1547-5069.2012.01444.x] [Medline: 22551013]
14. Niemiec R. Character Strengths Interventions: A Field Guide for Practitioners. First Edition. Boston, UK: Hogrefe Publishing; 2017.
15. Seligman ME, Csikszentmihalyi M. Positive psychology: an introduction. Am Psychol 2000;55(1):5-14 [FREE Full text] [doi: 10.1037/0003-066x.55.1.5]
16. Kristjansdottir OB, Stenberg U, Mirkovic J, Krogseth T, Ljoså TM, Stange KC, et al. Personal strengths reported by people with chronic illness: a qualitative study. Health Expect 2018 Aug;21(4):787-795 [FREE Full text] [doi: 10.1111/hex.12674] [Medline: 29478260]
17. Rapp C, Goscha R. The Strengths Model: A Recovery-Oriented Approach to Mental Health Services. Oxford, UK: Oxford University Press; 2012.
18. Harzer C. The eudaimonics of human strengths: the relations between character strengths and well-being. In: Vittersø J, editor. Handbook of Eudaimonic Well-Being. Switzerland: Springer; 2016:307-322.
19. Gander F, Proyer RT, Ruch W, Wyss T. Strength-based positive interventions: further evidence for their potential in enhancing well-being and alleviating depression. J Happiness Stud 2012 Sep 4;14(4):1241-1259 [FREE Full text] [doi: 10.1007/s10902-012-9380-0]
20. Sturgeon JA, Zautra AJ. Resilience: a new paradigm for adaptation to chronic pain. Curr Pain Headache Rep 2010 Apr;14(2):105-112 [FREE Full text] [doi: 10.1007/s11916-010-0095-9] [Medline: 20425199]
21. Rotegård AK, Fagermoen MS, Ruland CM. Cancer patients' experiences of their personal strengths through illness and recovery. Cancer Nurs 2012;35(1):E8-17. [doi: 10.1097/NCC.0b013e3182116497] [Medline: 21558850]
22. Cheavens J, Strunk D, Lazarus S, Goldstein L. The compensation and capitalization models: a test of two approaches to individualizing the treatment of depression. Behav Res Ther 2012 Nov;50(11):699-706. [doi: 10.1016/j.brat.2012.08.002] [Medline: 22982085]
23. Hibbard JH, Mahoney E. Toward a theory of patient and consumer activation. Patient Educ Couns 2010 Mar;78(3):377-381. [doi: 10.1016/j.pec.2009.12.015] [Medline: 20188505]
24. Mirkovic J, Kristjansdottir OB, Stenberg U, Krogseth T, Stange KC, Ruland CM. Patient insights into the design of technology to support a strengths-based approach to health care. JMIR Res Protoc 2016 Aug 24;5(3):e175 [FREE Full text] [doi: 10.2196/resprot.5906] [Medline: 27558951]
25. Kristjánsdóttir OB, Zangi H, Ruland C, Mirkovic J, Stange KC, Stenberg U. Helping patients mobilize their personal strengths within a rheumatology setting: a qualitative study with healthcare providers. Eur J Person Center Healthcare 2019;7(2):376 [FREE Full text] [doi: 10.5750/ejpch.v7i2.1713]
26. Pérez-Jover V, Sala-González M, Guilabert M, Mira JJ. Mobile apps for increasing treatment adherence: systematic review. J Med Internet Res 2019 Jun 18;21(6):e12505 [FREE Full text] [doi: 10.2196/12505] [Medline: 31215517]

27. Wolf A, Fors A, Ulin K, Thorn J, Swedberg K, Ekman I. An ehealth diary and symptom-tracking tool combined with person-centered care for improving self-efficacy after a diagnosis of acute coronary syndrome: a substudy of a randomized controlled trial. J Med Internet Res 2016 Feb 23;18(2):e40 [FREE Full text] [doi: 10.2196/jmir.4890] [Medline: 26907584]
28. Vilardaga R, Rizo J, Zeng E, Kientz JA, Ries R, Otis C, et al. User-centered design of learn to quit, a smoking cessation smartphone app for people with serious mental illness. JMIR Serious Games 2018 Jan 16;6(1):e2 [FREE Full text] [doi: 10.2196/games.8881] [Medline: 29339346]
29. Najm A, Gossec L, Weill C, Benoist D, Berenbaum F, Nikiphorou E. Mobile health apps for self-management of rheumatic and musculoskeletal diseases: systematic literature review. JMIR Mhealth Uhealth 2019 Nov 26;7(11):e14730 [FREE Full text] [doi: 10.2196/14730] [Medline: 31769758]
30. Børøsund E, Mirkovic J, Clark MM, Ehlers SL, Andrykowski MA, Bergland A, et al. A stress management app intervention for cancer survivors: design, development, and usability testing. JMIR Form Res 2018 Sep 6;2(2):e19 [FREE Full text] [doi: 10.2196/formative.9954] [Medline: 30684438]
31. Marcolino MS, Oliveira JA, D'Agostino M, Ribeiro AL, Alkmim MB, Novillo-Ortiz D. The impact of mhealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
32. Kelders SM, Kok RN, Ossebaard HC, van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
33. Kelders S. Involvement as a Working Mechanism for Persuasive Technology. In: Proceedings of the International Conference on Persuasive Technology. 2015 Presented at: PERSUASIVE'15; June 3-5, 2015; Chicago, IL, USA. [doi: 10.1007/978-3-319-20306-5_1]
34. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
35. Kelders SM, van Zyl LE, Ludden GD. The concept and components of engagement in different domains applied to ehealth: a systematic scoping review. Front Psychol 2020;11:926 [FREE Full text] [doi: 10.3389/fpsyg.2020.00926] [Medline: 32536888]
36. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
37. Hamari J, Koivisto J, Sarsa H. Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification. In: Proceedings of the 47th Hawaii International Conference on System Sciences. 2014 Presented at: HICSS'14; January 6-9, 2014; Waikoloa, HI, USA. [doi: 10.1109/hicss.2014.377]
38. Cheng VW, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in apps and technologies for improving mental health and well-being: systematic review. JMIR Ment Health 2019 Jun 26;6(6):e13717 [FREE Full text] [doi: 10.2196/13717] [Medline: 31244479]
39. Muangsrinoon S, Boonbrahm P. Game elements from literature review of gamification in healthcare context. J Technol Sci Educ 2019 Feb 1;9(1):20 [FREE Full text] [doi: 10.3926/jotse.556]
40. Huotari K, Hamari J. A definition for gamification: anchoring gamification in the service marketing literature. Electron Mark 2016 Jan 15;27(1):21-31. [doi: 10.1007/s12525-015-0212-z]
41. Shaw A, Paul D, Billingsley W, Kwan P, Wilson R. Gamification in E-mental Health: Development of a Digital Intervention Addressing Severe Mental Illness and Metabolic Syndrome. In: 8th International Conference on Internet Technologies & Society. 2017 Presented at: ITS'17; December 11-13, 2017; Sydney, Australia p. 83-90 URL: http://www.iadisportal.org/digital-library/gamification-in-e-mental-health-development-of-a-digital-intervention-addressing-severe-mental-illness-and-metabolic-syndrome
42. Rath JM, Williams V, Rubenstein R, Smith L, Vallone D. Assessing the impact of an interactive mobile game on tobacco-related attitudes and beliefs: the truth campaign's 'flavor monsters'. Games Health J 2015 Dec;4(6):480-487 [FREE Full text] [doi: 10.1089/g4h.2015.0005] [Medline: 26230971]
43. Allam A, Kostova Z, Nakamoto K, Schulz P. The effect of social support features and gamification on a web-based intervention for rheumatoid arthritis patients: randomized controlled trial. J Med Internet Res 2015 Jan 9;17(1):e14 [FREE Full text] [doi: 10.2196/jmir.3510] [Medline: 25574939]
44. Deterding S. A Quick Buck by Copy and Paste. Gamification Research Network. 2011. URL: http://gamification-research.org/2011/09/a-quick-buck-by-copy-and-paste/ [accessed 2017-06-17]
45. Sicart M. Playing the good life: gamification and ethics. In: Walz SP, Deterding S, editors. The Gameful World: Approaches, Issues, Applications. London, UK: MIT Press; 2015.
46. Ahtinen A, Mattila E, Välkkynen P, Kaipainen K, Vanhala T, Ermes M, et al. Mobile mental wellness training for stress management: feasibility and design implications based on a one-month field study. JMIR Mhealth Uhealth 2013 Jul 10;1(2):e11 [FREE Full text] [doi: 10.2196/mhealth.2596] [Medline: 25100683]

47. Nicholson S. A RECIPE for meaningful gamification. In: Gamification in Education and Business. London, UK: Springer; 2015.
48. Deterding S. The lens of intrinsic skill atoms: a method for gameful design. Hum-Comput Interac 2015 May 15;30(3-4):294-335. [doi: 10.1080/07370024.2014.993471]
49. Mora A, Riera D, González C, Arnedo-Moreno J. Gamification: a systematic review of design frameworks. J Comput High Educ 2017 May 23;29(3):516-548. [doi: 10.1007/s12528-017-9150-4]
50. Nicholson S. A User-Centered Theoretical Framework for Meaningful Gamification. In: Games+ Learning+ Society 8.0. 2012 Presented at: GLS'12; June 13-15, 2012; Madison, WI p. 223-230 URL: http://scottnicholson.com/pubs/meaningfulframework.pdf [doi: 10.1184/R1/6686786]
51. Edwards EA, Lumsden J, Rivas C, Steed L, Edwards LA, Thiyagarajan A, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. BMJ Open 2016 Oct 4;6(10):e012447 [FREE Full text] [doi: 10.1136/bmjopen-2016-012447] [Medline: 27707829]
52. Ludden GD, van Rompay TJ, Kelders SM, van Gemert-Pijnen JE. How to increase reach and adherence of web-based interventions: a design research viewpoint. J Med Internet Res 2015 Jul 10;17(7):e172 [FREE Full text] [doi: 10.2196/jmir.4201] [Medline: 26163456]
53. Simonsen J, Robertson T. Routledge International Handbook of Participatory Design. New York, USA: Routledge; 2013.
54. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008 Mar 26;4(1):5-18. [doi: 10.1080/15710880701875068]
55. Kensing F, Greenbaum J. Heritage: having a say. In: Simonsen J, Robertson T, editors. Routledge International Handbook of Participatory Design. New York, USA: Routledge; 2013:21-36.
56. Brandt E. Designing Exploratory Design Games: A Framework for Participation in Participatory Design? In: Proceedings of the Ninth Conference on Participatory Design: Expanding Boundaries in Design - Volume 1. 2006 Presented at: PDC'06; August 1-5, 2006; Trento, Italy. [doi: 10.1145/1147261.1147271]
57. Brandt E, Binder T, Sanders EB. Tools and techniques. In: Routledge International Handbook of Participatory Design. New York, USA: Routledge; 2006:145-181.
58. Vidal RV. The Future Workshop: Democratic Problem Solving. IMM, DTU, Image Analysis and Computer Graphics. 2005. URL: http://www2.imm.dtu.dk/pubdb/pubs/4095-full.html [accessed 2020-07-07]
59. Velden MV, Sommervold M, Culén A, Nakstad B. Designing interactive technologies with teenagers in a hospital setting. In: Little L, Fitton D, Bell BT, Toth N, editors. Perspectives on HCI Research with Teenagers. New York, USA: Springer; 2016:103-131.
60. Ruland CM, Starren J, Vatne TM. Participatory design with children in the development of a support system for patient-centered care in pediatric oncology. J Biomed Inform 2008 Aug;41(4):624-635 [FREE Full text] [doi: 10.1016/j.jbi.2007.10.004] [Medline: 18082468]
61. -. Design Council. Design for Public Good. 2013. URL: https://www.designcouncil.org.uk/resources/report/design-public-good [accessed 2020-06-30]
62. Löwgren J, Stolterman E. Thoughtful Interaction Design: A Design Perspective on Information Technology. Cambridge, Massachusetts: MIT Press; 2005.
63. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. J Am Med Assoc 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
64. Whitney D, Trosten-Bloom A. The Power of Appreciative Inquiry: A Practical Guide to Positive Change. Oakland, CA: Berrett-Koehler Publishers; 2010.
65. Mirkovic J, Jessen S, Kristjansdottir OB, Krogseth T, Koricho AT, Ruland CM. Developing technology to mobilize personal strengths in people with chronic illness: positive codesign approach. JMIR Form Res 2018 Jun 5;2(1):e10774 [FREE Full text] [doi: 10.2196/10774] [Medline: 30684404]
66. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
67. Nikrahan GR, Suarez L, Asgari K, Beach SR, Celano CM, Kalantari M, et al. Positive psychology interventions for patients with heart disease: a preliminary randomized trial. Psychosomatics 2016;57(4):348-358 [FREE Full text] [doi: 10.1016/j.psym.2016.03.003] [Medline: 27137709]
68. Nikrahan GR, Laferton JA, Asgari K, Kalantari M, Abedi MR, Etesampour A, et al. Effects of positive psychology interventions on risk biomarkers in coronary patients: a randomized, wait-list controlled pilot trial. Psychosomatics 2016;57(4):359-368 [FREE Full text] [doi: 10.1016/j.psym.2016.02.007] [Medline: 27129358]
69. Cerezo MV, Ortiz-Tallo M, Cardenal V, de la Torre-Luque A. Positive psychology group intervention for breast cancer patients: a randomised trial. Psychol Rep 2014 Aug;115(1):44-64. [doi: 10.2466/15.20.PR0.115c17z7] [Medline: 25153949]
70. Zangi HA, Mowinckel P, Finset A, Eriksson LR, Høystad TO, Lunde AK, et al. A mindfulness-based group intervention to reduce psychological distress and fatigue in patients with inflammatory rheumatic joint diseases: a randomised controlled trial. Ann Rheum Dis 2012 Jun;71(6):911-917. [doi: 10.1136/annrheumdis-2011-200351] [Medline: 22186709]

71. Müller R, Gertz KJ, Molton IR, Terrill AL, Bombardier CH, Ehde DM, et al. Effects of a tailored positive psychology intervention on well-being and pain in individuals with chronic pain and a physical disability: a feasibility trial. Clin J Pain 2016 Jan;32(1):32-44. [doi: 10.1097/AJP.0000000000000225] [Medline: 25724020]
72. Cohn MA, Pietrucha ME, Saslow LR, Hult JR, Moskowitz JT. An online positive affect skills intervention reduces depression in adults with type 2 diabetes. J Posit Psychol 2014 Jan 1;9(6):523-534 [FREE Full text] [doi: 10.1080/17439760.2014.920410] [Medline: 25214877]
73. Cheung EO, Cohn MA, Dunn LB, Melisko ME, Morgan S, Penedo FJ, et al. A randomized pilot trial of a positive affect skill intervention (lessons in linking affect and coping) for women with metastatic breast cancer. Psychooncology 2017 Dec;26(12):2101-2108 [FREE Full text] [doi: 10.1002/pon.4312] [Medline: 27862646]
74. 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]
75. Linley PA, Nielsen KM, Gillett R, Biswas-Diener R. Using signature strengths in pursuit of goals: effects on goal progress, need satisfaction, and well-being, and implications for coaching psychologists. Int Coach Psychol Rev 2010 Dec;5(1):6-15 [FREE Full text]
76. Brandt E, Messeter J. Facilitating Collaboration Through Design Games. In: Proceedings of the Eighth Conference on Participatory Design: Artful Integration: Interweaving Media, Materials and Practices - Volume 1. 2004 Presented at: PDC'04; July 27-31, 2004; Toronto, Ontari. [doi: 10.1145/1011870.1011885]
77. Halskov K, Dalsgård P. Inspiration Card Workshops. In: Proceedings of the 6th conference on Designing Interactive Systems. 2006 Presented at: DIS'06; June 26-28, 2006; University Park, PA, USA. [doi: 10.1145/1142405.1142409]
78. Jessen S, Mirkovic J, Ruland CM. Creating gameful design in mhealth: a participatory co-design approach. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e11579 [FREE Full text] [doi: 10.2196/11579] [Medline: 30552080]
79. Gabarron E, Bradway M, Årsand E. Social media for adults. In: Diabetes Digital Health. Amsterdam, Netherlands: Elsevier; 2020:119-129.
80. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
81. Fullerton T. Game Design Workshop - A Playcentric Approach To Creating Innovative Games. New York, USA: A K Peters/CRC Press; 2014.
82. Niemiec R, Wedding D. Positive Psychology at the Movies: Using Films to Build Virtues and Character Strengths. Boston, MA: Hogrefe Publishing; 2013.
83. Schutte N, Malouff J. The impact of signature character strengths interventions: a meta-analysis. J Happiness Stud 2018 May 21;20(4):1179-1196. [doi: 10.1007/s10902-018-9990-2]
84. Berdal G, Sand-Svartrud A, Bø I, Dager TN, Dingsør A, Eppeland SG, et al. Aiming for a healthier life: a qualitative content analysis of rehabilitation goals in patients with rheumatic diseases. Disabil Rehabil 2018 Apr;40(7):765-778. [doi: 10.1080/09638288.2016.1275043] [Medline: 28084842]
85. Ericsson K, Simon H. Protocol Analysis: Verbal Reports as Data. Cambridge, USA: MIT press; 1993.
86. Unity Technologies. URL: https://unity.com/ [accessed 2020-07-01]
87. Kim TW, Werbach K. More than just a game: ethical issues in gamification. Ethics Inf Technol 2016 May 12;18(2):157-173. [doi: 10.1007/s10676-016-9401-5]
88. Triberti S, Kelders S, Gaggioli A. User engagement. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. EHealth Research, Theory and Development: A Multi-Disciplinary Approach. London & New York: Routledge; 2018:271-289.
89. Mylonopoulou V. Design for Health Behavior Change Supportive Technology: Healthcare Professionals' Perspective. In: Proceedings of the 10th Nordic Conference on Human-Computer Interaction. 2018 Presented at: NordiCHI'18; October 1-3, 2018; Oslo, Norway. [doi: 10.1145/3240167.3240196]
90. Malone T. Heuristics for Designing Enjoyable User Interfaces: Lessons From Computer Games. In: Proceedings of the 1982 Conference on Human Factors in Computing Systems. 1982 Presented at: CHI'82; March 15-17, 1982; Maryland, United States URL: http://portal.acm.org/citation.cfm?id=800049.801756 [doi: 10.1145/800049.801756]
91. Huber S, Röpke K. How gamification can help companies to become more sustainable: a case study on ride sharing. In: Reiners T, Wood LC, editors. Gamification in Education and Business. New York, USA: Springer; 2015:615-636.
92. Swartz MK. A strength-based approach to care. J Pediatr Health Care 2017;31(1):1. [doi: 10.1016/j.pedhc.2016.10.008] [Medline: 27939585]
93. Gottlieb LN. CE: strengths-based nursing. Am J Nurs 2014;114(8):24-32. [doi: 10.1097/01.naj.0000453039.70629.e2]
94. Vugts A, van den Hoven M, de Vet E, Verweij M. How autonomy is understood in discussions on the ethics of nudging. Behavioural Public Policy 2020;4(1):108-123. [doi: 10.1017/bpp.2018.5]
95. Rath T. StrengthsFinder 2.0. Omaha: Gallup Press; 2007.
96. Gottlieb LN, Gottlieb B. The spiraling process for uncovering and discovering strengths. In: Strengths-Based Nursing Care: Health And Healing For Person And Family. New York, USA: Springer Publishing Company; 2018.

97. Calvo R, Peters D. Positive Computing: Technology for Wellbeing and Human Potential. Cambridge, Massachusetts: MIT Press; 2014.
98. Gemert-pijnen LV, Kelders S, Jong N, Oinas-kukkonen H. Persuasive health tehnologies. In: Gemert-pijnen L, Kelders S, Kip H, Sanderman R, editors. EHealth Research, Theory and Development: A Multi-Disciplinary Approach. London & New York: Routledge; 2018:228-246.
99. Kelders SM, Pots WT, Oskam MJ, Bohlmeijer ET, van Gemert-Pijnen JE. Development of a web-based intervention for the indicated prevention of depression. BMC Med Inform Decis Mak 2013 Feb 20;13:26 [FREE Full text] [doi: 10.1186/1472-6947-13-26] [Medline: 23425322]
100. Bratteteig T, Wagner I. Disentangling Participation: Power and Decision-making in Participatory Design. New York, USA: Springer International Publishing; 2014.
101. Bratteteig T, Wagner I. Design Decisions and the Sharing of Power in PD. In: Proceedings of the 13th Participatory Design Conference: Short Papers, Industry Cases, Workshop Descriptions, Doctoral Consortium papers, and Keynote abstracts - Volume 2. 2014 Presented at: PDC'14; October 6-10, 2014; Windhoek, Namibia. [doi: 10.1145/2662155.2662192]
102. Russomanno J, Patterson JG, Jabson Tree JM. Social media recruitment of marginalized, hard-to-reach populations: development of recruitment and monitoring guidelines. JMIR Public Health Surveill 2019 Dec 2;5(4):e14886 [FREE Full text] [doi: 10.2196/14886] [Medline: 31789598]

Abbreviations

eHealth: electronic health
mHealth: mobile health
IT: information technology

Edited by G Eysenbach; submitted 30.01.20; peer-reviewed by V Mylonopoulou, S Hewage, J Ropero, J Patterson, M Marcolino; comments to author 21.02.20; revised version received 06.06.20; accepted 13.06.20; published 24.07.20.

Please cite as:
Jessen S, Mirkovic J, Nes LS
MyStrengths, a Strengths-Focused Mobile Health Tool: Participatory Design and Development
JMIR Form Res 2020;4(7):e18049
URL: http://formative.jmir.org/2020/7/e18049/
doi:10.2196/18049
PMID:32706651

Original Paper

Utility and Perceived Value of a Provincial Digital Diagnostic Imaging Repository: Multimethod Study

Lisa Wickerson[1,2], BSc, PT, PhD; Jamie K Fujioka[2], MPH; Vanessa Kishimoto[2], BSc; Trevor Jamieson[2,3,4], MBI, MD; Ben Fine[2,5], SM, MD; R Sacha Bhatia[2,3,4,6], MD, MBA; Laura Desveaux[2,6], PT, PhD

[1]University Health Network, Toronto, ON, Canada

[2]Institute for Health Systems Solutions and Virtual Care, Women's College Hospital, Toronto, ON, Canada

[3]Unity Health Toronto, Toronto, ON, Canada

[4]Department of Medicine, University of Toronto, Toronto, ON, Canada

[5]Department of Medical Imaging, University of Toronto, Toronto, ON, Canada

[6]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Laura Desveaux, PT, PhD
Institute for Health Systems Solutions and Virtual Care
Women's College Hospital
76 Grenville Street
Toronto, ON, M5S 1B2
Canada
Phone: 1 4163236400 ext 4302
Email: laura.desveaux@wchospital.ca

Abstract

Background: Timely and comprehensive diagnostic image sharing across institutional and regional boundaries can produce multiple benefits while supporting integrated models of care. In Ontario, Canada, the Diagnostic Imaging Common Service (DICS) was created as a centralized imaging repository to enable the sharing and viewing of diagnostic images and associated reports across hospital-based and community-based clinicians throughout the province.

Objective: The aims of this study were as follows: (1) to explore real-world utilization and perceived clinical value of the DICS following the provision of system-wide access and (2) to identify strategies to optimize the technology platform functionality and encourage adoption.

Methods: This multimethod study included semistructured interviews with physicians and administrative stakeholders and descriptive analysis of the current DICS usage data.

Results: In this study, 41 participants were interviewed, that is, 34 physicians and 7 administrative stakeholders. The following 4 key themes emerged: (1) utilization of the DICS depended on the awareness of the technology and the preferred channels for accessing images, which varied widely, (2) clinical responsibilities and available institutional resources were the drivers of utilization (or lack thereof), (3) centralized image repositories were perceived to offer value at the patient, clinician, and health care system levels, and (4) the enabling factors to realize value included aspects of technology infrastructure (ie, available functionality) alongside policy supports. High-volume DICS usage was not evenly distributed throughout the province.

Conclusions: Suboptimal adoption of the DICS was driven by poor awareness and variations in the clinical workflow. Alignment with physician workflow, policy supports, and investment in key technological features and infrastructure would improve functionality and data comprehensiveness, thereby optimizing health system performance, patient and provider experience, population health, and health care costs.

(JMIR Form Res 2020;4(7):e17220) doi:10.2196/17220

KEYWORDS

diagnostic imaging; eHealth; health care delivery

Introduction

The ability to electronically share patient-level information across institutional and geographic boundaries can facilitate an integrated and coordinated model of health care delivery among hospital-based and community-based health care professionals by providing timely transfer of relevant and complete information to inform clinical decision making [1]. Within the diagnostic imaging landscape, rapid and comprehensive sharing of imaging data has demonstrated multiple benefits at the patient, clinician, and health care system levels. Patient experience and quality of care is improved by reducing the number of unnecessary tests and radiation exposure [2-5] and by facilitating timely access to specialist consultation and treatment [4]. At the clinician level, timely diagnosis and treatment improves clinical and administrative workflows [2,4,6,7]. Reducing duplicate imaging improves the overall health care system efficiency and associated costs [2-5,8,9]. Technology-enabled models of electronic image sharing include institutional and multi-institutional regional picture archiving and communication systems (PACSs), onsite and offsite vendor neutral archives, cloud-based image transfer, and cross-enterprise document image sharing [10-13]. The technology environment of a health care organization or a clinical practice (ie, technology infrastructure, storage, and resources) influences the platform availability and channels of access for internal and external imaging studies [7]. The import and display of external priors (ie, imaging performed at external institutions) can be time-intensive and lead to clinical inefficiencies, treatment delays, and duplicate testing [2-4]. Physicians may temporarily retrieve and import external imaging data from regional diagnostic imaging repositories (DIRs) directly into their local PACS by using import and display of external priors or foreign exam management [14], whereas community-based health care professionals may rely on accessing a third-party provider portal, image upload from a compact disc or, most commonly, a fax of an image report. Multiple channels of access to diagnostic imaging can lead to variability in clinical and administrative workflow.

In order to alleviate the logistical and administrative burden of transferring imaging data across multiple systems, several jurisdictions have implemented centralized imaging repositories through health information exchanges [13,15]. For example, Scotland has adopted a unified nationwide approach with a single supplier PACS and a central data archive for long-term data storage and sharing [16]. In jurisdictions with fragmented PACSs and electronic medical record systems such as the United Kingdom and Estonia, blockchain technology has been introduced to increase interoperability and decentralize data for easy access and exchange [17]. Blockchain systems have been shown to improve interoperability and reduce administrative costs involved in transporting data, without compromising the security [17]. Within Ontario, Canada, a centralized imaging repository was created by eHealth Ontario, an agency affiliated with the Ontario Ministry of Health, to enable and support real-time sharing and viewing of diagnostic images and reports, which was the focus of this study.

The objective of this clinically focused evaluation was to explore physician engagement and the perceived value following system-wide access to a centralized DIR in Ontario, Canada (the Diagnostic Imaging Common Service [DICS]). This evaluation was completed in partnership with the Ontario Ministry of Health to directly inform future strategies to optimize the technology platform and increase adoption and meaningful use.

Methods

Study Design

This multimethod study included one-on-one semistructured interviews that explored the utilization and perceived value of the DICS from the perspective of the physicians across a wide range of specialties and geographic areas. Usage data relating to access of the repository was also examined to understand the utilization of the centralized repository and the practice characteristics of the users who accessed large volumes of data. Ethics approval was obtained from the research ethics board of Women's College Hospital (REB# 2018-0177-E).

Study Setting

Medically necessary hospital and physician services (including diagnostic imaging) is publicly funded in Ontario, Canada according to the Canada Health Act [18]. Ontario has a population of over 14 million, which represents 38.6% of the Canadian population in 2018 [19]. There are over 36,000 physicians in active practice in Ontario, including 1000 radiologists and 13,500 primary care physicians [20-22]. In 2012-2013, more than 21 million diagnostic imaging procedures were performed in Ontario, 60% of which were performed in hospitals and the remainder in stand-alone independent health facilities [23]. eHealth Ontario, an agency affiliated with the Ontario Ministry of Health, created the DICS to enable and support real-time sharing and viewing of diagnostic images and reports.

The DICS provides a single front-end web-based viewer that, on the back end, consolidates access to imaging procedures stored across multiple DIRs, thereby providing long-term storage for hospitals and contributing independent health facilities within a specified geography-based catchment area [24]. The DICS is a federated repository that uses a cross-enterprise document sharing to facilitate registration, distribution, and access to images across health care organizations. When a health care practitioner initiates an incoming query to access an image or a report (via a web-based viewer), the cross-enterprise document sharing integration profile sources the relevant data from the DIRs. This enables the health care practitioners to view images and reports across the entire province (ie, outside of their traditional institutional and regional boundaries). To ease the access for these health care practitioners, the DICS was embedded into 3 pre-existing clinical viewing portals as of August 2018. These portals provide the heath care practitioners access to a range of patient-level information, including diagnostic imaging reports, dispensed medications, laboratory results, hospital visits, and home and community care information (ie, referral details, risk assessments, and care plans). The DICS initiative provides the additional ability to

view images in a web-based viewer on top of the baseline ability to view image reports. As of January 2019, there were over 35 million images and over 47 million reports available on the DICS. Of these, <0.001% of the images and <0.002% of the reports were accessed during the month of January. While there is some variation based on the clinical viewing portal used, it takes the DICS system an average of 0.49 seconds to respond to a query (range, 0.01-40.85 seconds). This response time is distinct from the time it takes the image to load, which is known to be variable.

Recruitment and Data Collection

Physicians were recruited using a combination of convenience and purposive sampling to achieve a diverse sample of participants that reflect the breadth of the current use and the future potential of the DICS. The research team, representatives of the Ontario College of Family Physicians, and working group members of the clinical viewing portals were asked to refer contacts who could provide relevant insight. A snowball recruitment strategy was employed, wherein interview participants were asked to refer colleagues who may have relevant insights related to the DICS platform or access to imaging. Purposive data-driven recruitment was also used by asking eHealth Ontario to send out recruitment emails to high-volume users of the DICS based on the viewing statistics. The recruitment expanded to include administrative stakeholders (ie, those with experience in hospital PACS administration, regional DIRs, and independent health facilities) in response to emerging themes around multiple channels of accessing diagnostic imaging and role of the clinical viewing portals in order to fully understand access and engagement. The practice characteristics (ie, geographic location, professional role, and specialty) of high-volume users accessing the DICS during the study timeframe were obtained from eHealth Ontario's usage statistics.

Data Analysis

All interviews were audio recorded and professionally transcribed verbatim. Two researchers (LW and JF) independently and inductively coded 3 transcripts to develop a coding framework using NVivo (QSR International), which was applied to the remaining transcripts. Emerging insights were coded inductively and added to the codebook, as necessary. An inductive thematic analysis was applied to identify prominent and recurring themes at regular intervals. Codes were reviewed by 3 members of the research team (ie, LW, JF, VK) who then began the process of thematic mapping to understand the relationships and to generate preliminary themes and subthemes. Refinements and specifications of the thematic categories, subcategories, and relationships between the themes were discerned based on in-depth discussion and negotiated consensus with the fourth member of the research team (LD). Descriptive statistics was performed on usage statistics.

Results

Between February 22, 2019 and June 30, 2019, 41 participants were interviewed, that is, 34 physicians and 7 administrative stakeholders. The physicians represented a cross-section of the practice areas and a broad range of specialties (Table 1). The administrative stakeholders were managers or directors working in independent health facilities, hospitals, or the regional DIRs who were involved in the implementation or oversight of diagnostic images.

1. Variable utilization is driven by awareness and access preferences: A lack of awareness of prior imaging studies alongside multiple channels available to access internal and external diagnostic imaging and reports (ie, local viewers, regional DIRs, and the DICS through clinical viewing portals) leads to inconsistent and suboptimal engagement with the DICS.
2. Clinical roles and institutional resources inform utilization practices: The functionality of the clinical viewing portals connecting to a centralized DIR needs to fully support the intricacies of the clinical workflow requirements of multiple users, including radiologists, specialists, and primary health care professionals. Radiologists require high-resolution capabilities for images with a link to transcription software to interpret and produce a diagnostic image report, whereas specialists (eg, oncologists, surgeons) use images to plan a surgical approach and monitor change over time in order to plan medical treatments and evaluate responses. When image fidelity, speed of upload, measurement tools, and viewing features for comparative studies (ie, side-by-side viewing) in clinical viewing portals were inferior to local PACS, engagement with the DICS was low. In primary health care, clinicians need efficient and comprehensive access to full reports (rather than the images themselves) to inform their clinical decision making and navigate the patient thorough the health care system. Incomplete or delayed access to reports resulted in decreased utilization. The institutional and technological ecosystem influences how external diagnostic imaging is utilized. While integrating regional DIRs with a local PACS through foreign exam management is a common modality for image viewing, small hospitals and community services may have less technological resources and updated software to support this.
3. Centralized diagnostic imaging was perceived to offer value at the patient, clinician, and health system levels: A centralized DIR with efficient image access was perceived to increase patient satisfaction and safety (ie, reduced radiation exposure, timely diagnosis, and timely treatment), improve the clinical and administrative workflows and communication of the health care professionals, and optimize the health care organizational efficiency by reducing unnecessary repeat imaging and subsequently reducing the health care costs and wait times.
4. Enabling factors to realize value include technology infrastructure and policy supports: High-value technology infrastructure for radiologists and specialists include automatic integration and downloading of images into local systems to enable a comprehensive view of the imaging history, alongside specific functionality of the DICS to support this. Policy supports and infrastructure to promote interoperability between systems (ie, standardization and regulation) and to reduce the image contribution gaps from community-based diagnostic imaging services (ie, independent health facilities) and specific clinical specialties

outside of traditional radiology and the scope of DIRs (ie, cardiac imaging) would significantly increase the value and utilization of the centralized repository.

Over 570,000 reports and 135,000 images were viewed between September 2018 and April 2019 across the 3 clinical viewing portals. In February 2019, there were 11,070 users who accessed the DICS at least once during the month. The majority of the high-volume users (566/658, 86%), defined by the research team as viewing >15 images or reports in that month, included health care professionals who were located primarily in the Greater Toronto Area and Southeast Ontario region. This distribution mirrored the demographic profile of the interviewed participants (Table 1).

Four key themes emerged (Table 2), which described physician experiences with accessing images and engaging with the DICS and their overall perceptions of value.

Table 1. Participant characteristics (N=41).

Participant characteristics	Physicians (n=34), n (%)	Administrative stakeholders (n=7), n (%)
Professional specialty		
Primary care	7 (20)	N/A[a]
Radiology	6 (17)	N/A
Oncology	7 (20)	N/A
Thoracic surgery	3 (9)	N/A
Emergency medicine	2 (6)	N/A
Orthopedic surgery	2 (6)	N/A
General medicine	2 (6)	N/A
Respirology	1 (3)	N/A
Cardiology	1 (3)	N/A
Urology	1 (3)	N/A
Gastroenterology	1 (3)	N/A
Neurosurgery	1 (3)	N/A
Geographic region		
Greater Toronto Area and Southeast Ontario	29 (85)	N/A
Southwest Ontario	3 (9)	N/A
Northern and Eastern Ontario	2 (6)	N/A
Health care setting		
Hospital	30 (88)	N/A
Community	4 (12)	N/A
Years of clinical practice		
0-5	7 (20.5)	N/A
6-10	7 (20.5)	N/A
>10	18 (53)	N/A
Unavailable	2 (6)	N/A
Role		
Administrator for diagnostic imaging repository	N/A	5 (72)
PACS[b] administrator	N/A	1 (14)
Imaging director of an independent health facility	N/A	1 (14)
Geographic region		
Greater Toronto Area and Southeast Ontario	N/A	3 (43)
Southwest Ontario	N/A	1 (14)
Northern and Eastern Ontario	N/A	3 (43)

[a]N/A: not applicable.

[b]PACS: picture archiving and communication system.

Table 2. Themes describing physician experiences and their illustrative quotes.

Theme	Illustrative quote
Variable utilization is driven by awareness and access preferences	• *When I read conceptual imaging, it is like 20, 30, 40 cases a day. I do not know whether there was outside imaging done—if there would be a prompt, that would be very helpful, that is one. It is just not practical. I cannot log in to the clinical viewing portal and check whether any outside imaging was done for every single patient.*[Radiologist 1] • *It is my strong preference to use our own viewer in the hospital…because if I can access their images through our (regional DIR[a]) interface, and then I can transfer those images to our own PACS[b]system.* [Urologist 1]
Clinical roles and institutional resources inform utilization practices	• *Ultimately all these web-based viewers, no matter how good they are, they cannot be a PACS quality or caliber machine, because my PACS, yeah, sure, I can do other things on it, but really it is a computer dedicated to 1 task and 1 task only, whereas the website has lots of different tasks.* [Radiologist 4] • *I think the local PACS has more functionality in terms of the ability to display multiple images from different dates on the same screen and to synchronize the images so that I can make a direct comparison on 1 monitor between a current scan and a remote scan, like an older scan.* [Thoracic Surgeon 2] • *I think just, again, the integration into our workflow is really important. Primary care has a very busy, and a very, very chaotic workflow. If things do not fit into that, they often get left behind, even if they are potentially a helpful resource. I am just really thinking about how to access this in a very easy way, where I do not have to retype in the patient's date of birth, medical record number, name, etc. Ideally, if it is from the electronic health record, it is connected in. I think that is really important.* [Primary Health Care Professional 5] • *There are a few vendors that we do not have foreign exam management set up on yet. I know that they have reached out. They would really like to be able to do that…they do not have foreign exam management and I know that they would be a huge, huge user of it. Once they do decide on the vendor that they are going to use, hopefully we will be able to get foreign exam management set up with them, which would be a really big asset.* [DIR Administrator]
Centralized diagnostic imaging was perceived to offer value at the patient, clinician, and health system levels	• *I would say that more often than not, it will lead to one or both of the following outcomes. Number one, the patient care is delayed that day and clinics start running late. That is probably minor. It is annoying and it obviously costs patients more parking money, etc, but the more important one was that I think clinicians just go and say, you know what? Mrs. Robinson, I cannot read this fax. We cannot see your computed tomography image. I am going to call the radiologist. You clearly have an urgent problem and we are going to try to urgently book you for a new computed tomography session here within the next week. So then, clinicians start calling radiologists and they are like hey, I have got a patient in clinic, she is here today, it is kind of urgent, she looks like she has got jaundice, I do not know what is going on, I cannot see the image from the outside, can you just try and find a slot…..…I think you guys now can see the downstream impact that this does for health care delivery efficiency…it is huge. And that is just 1 patient's story but imagine 80 clinicians all doing this all at the same time in 1 center, each 1 individual. You can see how imaging departments across the province are going ugh.* [Oncologist 2].
Enabling factors to realize value include technology infrastructure and policy supports	• *I guess what I would need is that my PACS facilitator, the informatics team at my hospital, ideally, automatically, without having to do any work, would automatically be able to import those images in Digital Imaging and Communications in Medicine format from this large system into our system and be matched with the patient's name so that when I open the study, I can see all the prior imaging data.* [Radiologist 2] • *The final downside of the clinical viewing portal is that I cannot transfer the images to my own PACS system. That is always handy because then I can have access to them when I am on the network and if there is network downtime, then also it would allow me to use that internal viewer to do all that stuff more expeditiously.* [Urologist 1] • *The biggest downside to all these things is the fact that most outside community-based diagnostic imaging services are simply not available on any electronic format. In particular, in the ultrasound imaging performed outside (ie, independent health facilities), we are missing a vast majority of things that would really be sort of important and would prevent further testing. We order a lot of repeat ultrasounds in patients who have just had an ultrasound, which is a significant drain on resources because unfortunately, the report, I mean I really do not care what the image is but the actual report is simply not available in any repository. So that is kind of the biggest disadvantage is that community access is not there.* [Emergency Physician 2] • *It is logical, it is obvious, to be able to look at an echocardiogram and an electrocardiogram with a cardiac magnetic resonance image makes complete sense, and it would be used widely. But currently, the systems are firewalled, you cannot view echocardiograms on our system, and you cannot view nuclear medicine heart testing on our system….* [Radiologist 3]

[a]DIR: diagnostic imaging repository.

[b]PACS: picture archiving and communication system.

Discussion

Principal Findings

Although a centralized DIR was perceived to offer clinical value by physicians across a wide range of specialties, there was inconsistent and suboptimal engagement. This was driven by several factors, including a lack of awareness, nuances of clinical workflow and professional roles, functionality of the viewing portals, and policies around interoperability with the local viewing systems. A key driver that was identified to increase clinical utility was the addition of data sources from community-based diagnostic imaging services and other medical specialties to expand the comprehensiveness of the repository. The provision of broad, system-wide access to the DICS through clinical viewing portals was not reflective of, or sensitive to, the heterogeneity of the clinical roles, workflows, and diagnostic image requirements of physician specialists. The inability to upload images from the clinical viewing portals into the local PACS leads to limited utility for the physicians who rely primarily on the PACS for their workflow (ie, radiologists and in-hospital specialists) [25]. Rather than providing similar access and functionality to all the physicians, greater value may be realized if optimization of the digital platform is targeted toward high-priority clinical areas, wherein diagnostic imaging is integral to specific professional roles and workflows. The potential priority areas identified from this evaluation included oncology, surgery, and orthopedics, wherein diagnostic images are routinely used to inform a surgical and medical approach, monitor response to treatment, and track progression/resolution over time. Previous research performed in a surgical oncology center in Ontario found that a shared regional DIR decreased repeat imaging and reduced the waiting times for surgical consultation and surgery [4]. An additional specialized image sharing technology is the Emergency Neuro Imaging Transfer System. This system is a centralized web-based image archive distinct from the DICS that is available in select acute care centers in Ontario that provides temporary access to “on demand” neurological, vascular, and cardiac computed tomography images, magnetic resonance images, and ultrasound images for urgent or critical care [26]. To better understand how to optimize and implement the DICS to increase adoption and utilization, targeted clinician consultation is needed to elucidate value propositions in high-value clinical areas alongside robust clinical and administrative workflow mapping, co-design with intended clinician users [27,28], and education to increase the awareness of the centralized repository.

Beyond the functionality of the DICS and the alignment with the clinical and administrative workflows, a major reported barrier to clinical utility and perceived value of the DICS was the lack of data comprehensiveness. This arose from the reality that community-based independent health facilities do not regularly contribute images to regional DIRs, despite performing up to 40% of the radiology procedures in Ontario [23]. In addition, images that were outside of traditional radiology, such as cardiac images, were not within the scope of the original DIRs and thus were not consolidated into the DICS. Further, multiple channels for accessing imaging data often exist (ie, DIRs, local or regional PACS, and the DICS in Ontario), which can lead to image duplication and fragmented data sets. Consolidated access to comprehensive images and reports through a single portal would produce the most value for clinicians and the health system by facilitating timely access and reducing duplicate imaging [29]. However, administrative stakeholders described challenges surrounding system variation in terminology mapping, data formats, and protocols, storage, and sustainment models for ongoing image contribution from community-based facilities.

Successful implementation of diagnostic image sharing platforms is supported by digital formats and standards that ensure interoperability across multiple vendors and establishing image quality and reporting standards to meet the diverse needs of multiple clinical areas [30-32]. Policy supports around incentives and mandatory reporting of any imaging studies receiving public payment would also encourage data contributions from heterogeneous organizations. Such approaches are best supported by a coordinated government strategy such as the Nationwide Interoperability Roadmap in the United States [33], wherein vendors, health care systems, and medical institutions have committed to information technology exchange standards. Consequently, increasing comprehensive and consistent availability of imaging reports across the system is best facilitated through multistakeholder collaborative efforts, which would, in turn, optimize the clinical value of the centralized repository.

Limitations

This formative study provides insight into the perspectives of diverse health care professionals, including radiologists, physician specialists, and primary care practitioners. As this study was meant to generate hypotheses about the current use and future potential of the DICS, the next step is to conduct a thorough analysis of the quantitative usage data and identify high users of imaging services in Ontario. In collaboration with the Ontario Ministry of Health, we will identify target users and engage in specialty-specific assessments of imaging requirements and workflow patterns to increase the generalizability of the results. The majority of the participants (34/41, 82.9%) were physicians practicing at academically affiliated hospitals and localized to 1 region of the province. Although this was reflective of the demographic data of the DICS users, future work is needed to understand a broad range of the clinician perspectives as current usage is suboptimal, and current user demographic data are not reflective of the optimal target population. In addition, other types of health care professionals who may access the DICS, such as nurses and allied health care professionals, were not interviewed. The administrative stakeholder group was not the primary recruitment target; therefore, further engagement is needed to build upon their preliminary insights and gain a comprehensive understanding of the administrative and policy drivers that influence engagement with the DICS.

Conclusions

The clinical utility and perceived value of a system-wide, one-size-fits-all approach to a centralized DIR has not been fully realized owing to suboptimal awareness and lack of alignment with end-user workflows. Further engagement with potentially high-value clinical information users (ie, those who access large volumes of diagnostic images and reports) will help in aligning the technology platform with the nuances of different medical specialist end-user workflows and diagnostic imaging needs. In parallel, investment in data comprehensiveness and inclusion of all imaging reports in Ontario in 1 system would enhance the value by strengthening the utility of the available data.

Acknowledgments

We would like to acknowledge our project partners at the Digital Health Division, Ontario Ministry of Health and eHealth Ontario for their collaboration and participation in this evaluation.

Conflicts of Interest

None declared.

References

1. Graetz I, Reed M, Shortell SM, Rundall TG, Bellows J, Hsu J. The next step towards making use meaningful: electronic information exchange and care coordination across clinicians and delivery sites. Med Care 2014 Dec;52(12):1037-1041 [FREE Full text] [doi: 10.1097/MLR.0000000000000245] [Medline: 25304020]
2. Sheppard CW, Groll AL, Austin CL, Thompson SJ. Impact of duplicate CT scan rate after implementation of transfer image repository system at a level 1 trauma center. Emerg Radiol 2018 Jun;25(3):275-280 [FREE Full text] [doi: 10.1007/s10140-017-1575-6] [Medline: 29330668]
3. Jung H, Vest JR, Unruh MA, Kern LM, Kaushal R, HITEC Investigators. Use of Health Information Exchange and Repeat Imaging Costs. J Am Coll Radiol 2015 Dec;12(12 Pt B):1364-1370 [FREE Full text] [doi: 10.1016/j.jacr.2015.09.010] [Medline: 26614881]
4. Hallet J, Coburn NG, Alberga A, Fu L, Tharmalingam S, Beyfuss K, et al. Reducing repeat imaging in hepato-pancreatico-biliary surgical cancer care through shared diagnostic imaging repositories. HPB (Oxford) 2019 Jan;21(1):96-106. [doi: 10.1016/j.hpb.2018.06.1807] [Medline: 30297304]
5. Watson JJJ, Moren A, Diggs B, Houser B, Eastes L, Brand D, et al. A statewide teleradiology system reduces radiation exposure and charges in transferred trauma patients. Am J Surg 2016 May;211(5):908-912. [doi: 10.1016/j.amjsurg.2016.01.010] [Medline: 27012476]
6. Morgan MB, Young E, Harada S, Winkler N, Riegert J, Jones T, et al. Ditching the Disc: The Effects of Cloud-Based Image Sharing on Department Efficiency and Report Turnaround Times in Mammography. J Am Coll Radiol 2017 Dec;14(12):1560-1565. [doi: 10.1016/j.jacr.2017.08.024] [Medline: 28969976]
7. Aas IHM, Geitung JT. Choosing networks for picture archiving and communication systems and teleradiology. J Telemed Telecare 2003;9 Suppl 1:S27-S29. [doi: 10.1258/135763303322196231] [Medline: 12952712]
8. Welk B, Liu K, Al-Jaishi A, McArthur E, Jain AK, Ordon M. Repeated Diagnostic Imaging Studies in Ontario and the Impact of Health Information Exchange Systems. Healthc Q 2016;19(1):24-28. [doi: 10.12927/hcq.2016.24613] [Medline: 27133604]
9. You JJ, Yun L, Tu JV. Impact of picture archiving communication systems on rates of duplicate imaging: a before-after study. BMC Health Serv Res 2008 Nov 12;8:234 [FREE Full text] [doi: 10.1186/1472-6963-8-234] [Medline: 19014501]
10. Bolan C. A view of the future image exchange.: Appl Radiol; 2013. URL: https://appliedradiology.com/articles/technology-trends-a-view-of-the-future-image-exchange [accessed 2020-04-05]
11. Nagels J, MacDonald D, Parker D. Foreign exam management in practice: seamless access to foreign images and results in a regional environment. J Digit Imaging 2015 Apr;28(2):188-193 [FREE Full text] [doi: 10.1007/s10278-014-9735-7] [Medline: 25273504]
12. Caffery L, Manthey K. Implementation of a Web-based teleradiology management system. J Telemed Telecare 2004;10 Suppl 1:22-25. [doi: 10.1258/1357633042614230] [Medline: 15603599]
13. Schwind F, Münch H, Schröter A, Brandner R, Kutscha U, Brandner A, et al. Long-term experience with setup and implementation of an IHE-based image management and distribution system in intersectoral clinical routine. Int J Comput Assist Radiol Surg 2018 Nov;13(11):1727-1739. [doi: 10.1007/s11548-018-1819-2] [Medline: 29998400]
14. eHealth Ontario. Diagnostic Images. Ontario Health. URL: https://www.ehealthontario.on.ca/en/for-healthcare-professionals/diagnostic-imaging [accessed 2020-06-16]
15. Lebre R, Bastião L, Costa C. Shared Medical Imaging Repositories. Stud Health Technol Inform 2018;247:411-415. [Medline: 29677993]
16. Sutton LN. PACS and diagnostic imaging service delivery--a UK perspective. Eur J Radiol 2011 May;78(2):243-249. [doi: 10.1016/j.ejrad.2010.05.012] [Medline: 21600402]

17. Vazirani AA, O'Donoghue O, Brindley D, Meinert E. Implementing Blockchains for Efficient Health Care: Systematic Review. J Med Internet Res 2019 Feb 12;21(2):e12439 [FREE Full text] [doi: 10.2196/12439] [Medline: 30747714]
18. Canada Health Act. Government of Canada. 2018. URL: https://www.canada.ca/en/health-canada/services/health-care-system/canada-health-care-system-medicare/canada-health-act.html [accessed 2020-01-22]
19. Ontario Demographics Quarterly. Office of Economic Policy. 2019. URL: https://www.fin.gov.on.ca/en/economy/demographics/quarterly/dhiq1.html [accessed 2020-01-22]
20. 2018 Annual Report. College of Physicians and Surgeons of Ontario. 2018. URL: https://view.joomag.com/annual-report-2018/0076106001559747788?short [accessed 2020-01-22]
21. About the Ontario College of Family Physicians (OCFP). Ontario College of Family Physicians. URL: https://www.ontariofamilyphysicians.ca/advocacy/the-value-of-family-physicians/ocfp-infographic-september2019.pdf
22. About Us. Ontario Association of Radiologists. URL: https://oarinfo.ca [accessed 2020-01-22]
23. Peer review: a diagnostic imaging quality initiative for Ontario. Health Quality Ontario. URL: https://www.hqontario.ca/Portals/0/documents/health-quality/di-expert-panel-report-en.pdf [accessed 2020-01-22]
24. Hospital Diagnostic Imaging Repository Services (HDIRS). Hospital diagnostic imaging repository services (HDIRS): Report for Ed Clark. Ontario Government. URL: https://files.ontario.ca/22._hospital_diagnostic_imaging_repository_services.pdf [accessed 2020-01-22]
25. Canadian Agency for Drugs and Technologies in Health. Appropriate utilization of advanced diagnostic imaging procedures: CT, MRI and PET/CT environmental scan. Canadian Agency for Drugs and Technologies in Health. 2013. URL: https://www.cadth.ca/sites/default/files/pdf/DiagnosticImagingLitScan_e.pdf [accessed 2020-01-22]
26. Southwestern Ontario Diagnostic Imaging Network. ENITS. Southwestern Ontario Diagnostic Imaging Network. URL: https://swodin.ca/enits [accessed 2020-01-22]
27. Trischler J, Pervan SJ, Kelly SJ, Scott DR. The Value of Codesign. Journal of Service Research 2017 Jul 10;21(1):75-100. [doi: 10.1177/1094670517714060]
28. Shaw J, Agarwal P, Desveaux L, Palma DC, Stamenova V, Jamieson T, et al. Beyond "implementation": digital health innovation and service design. NPJ Digit Med 2018;1:48 [FREE Full text] [doi: 10.1038/s41746-018-0059-8] [Medline: 31304327]
29. Nagels J, Macdonald D, Coz C. Measuring the Benefits of a Regional Imaging Environment. J Digit Imaging 2017 Oct;30(5):609-614 [FREE Full text] [doi: 10.1007/s10278-017-9963-8] [Medline: 28299488]
30. Mendelson DS, Erickson BJ, Choy G. Image sharing: evolving solutions in the age of interoperability. J Am Coll Radiol 2014 Dec;11(12 Pt B):1260-1269 [FREE Full text] [doi: 10.1016/j.jacr.2014.09.013] [Medline: 25467903]
31. Roth CJ, Lannum LM, Dennison DK, Towbin AJ. The Current State and Path Forward For Enterprise Image Viewing: HIMSS-SIIM Collaborative White Paper. J Digit Imaging 2016 Oct;29(5):567-573 [FREE Full text] [doi: 10.1007/s10278-016-9887-8] [Medline: 27473474]
32. Chakera T, Nagree Y, Song S, Jones P. Bridging the communication gap between public and private radiology services. Med J Aust 2009 Nov 16;191(10):558-560. [Medline: 19912089]
33. Connecting health and care for the nation: a shared nationwide interoperability roadmap. The Office of National Coordinator for Health Information Technology Internet. 2015. URL: https://www.healthit.gov/sites/default/files/comments_upload/tascet_comments_to_onc_interoperability_roadmap_2015_04_01.pdf [accessed 2020-01-22]

Abbreviations

DICS: diagnostic imaging common service
DIR: diagnostic imaging repository
PACS: picture archiving and communication system

Edited by C Lovis; submitted 26.11.19; this is a non–peer-reviewed article;accepted 08.04.20; published 27.07.20.

Please cite as:
Wickerson L, Fujioka JK, Kishimoto V, Jamieson T, Fine B, Bhatia RS, Desveaux L
Utility and Perceived Value of a Provincial Digital Diagnostic Imaging Repository: Multimethod Study
JMIR Form Res 2020;4(7):e17220
URL: https://formative.jmir.org/2020/7/e17220
doi:10.2196/17220
PMID:32459644

Original Paper

A Companion App to Support Rheumatology Patients Treated with Certolizumab Pegol: Results From a Usability Study

Barbara Domańska[1*], MSc; Stijn Vansant[2*], MSc; Irina Mountian[2*], MD, PhD

[1]UCB Pharma, Slough, United Kingdom

[2]UCB Pharma, Brussels, Belgium

[*]all authors contributed equally

Corresponding Author:
Barbara Domańska, MSc
UCB Pharma
208 Bath Road
Slough
United Kingdom
Phone: 44 1615 197676
Email: barbara.domanska@UCB.com

Abstract

Background: Certolizumab pegol (CZP) is an anti-tumor necrosis factor drug approved for the treatment of multiple moderate to severe chronic inflammatory diseases. In the European Union, CZP is approved for administration by subcutaneous self-injection using a prefilled syringe, prefilled pen, or reusable electromechanical auto-injector (electronic device). CimplyMe is a companion app for use alongside CZP self-injection devices, designed to support CZP-treated patients self-managing their treatment and disease.

Objective: This study aimed to validate the usability of the companion app by demonstrating that tasks required for use can be performed successfully by intended end users.

Methods: We recruited 15 patients with moderate to severe rheumatoid arthritis, currently prescribed biologic treatment, and using apps on a smart phone. Patients were assessed on their ability to use the companion app in a setting designed to simulate a location where patients regularly administer biologic treatment. To assess the usability of the key features of the app, 8 critical and 3 noncritical tasks were designed. Patients' success on each task was recorded through observations or knowledge-based questions. Successes with difficulty and use errors were also recorded. If a patient made a use error at the first attempt, a second attempt was allowed. Second-attempt use errors were recorded as a task failure.

Results: A total of 207 first attempts at the 14 components of the 8 critical tasks were evaluated (3 patients failed to complete one component); 178 (86.0%) critical tasks were successfully completed at the first attempt. The remaining first attempts comprised 16 (7.7%) successes with difficulty and 13 (6.3%) use errors, which had to be repeated. One critical task was not re-attempted by one patient due to time constraints; however, there were no use errors in the 12 completed second attempts. A total of 107 first attempts at the 3 noncritical tasks were made, all of which (107/107, 100.0%) were completed without use errors.

Conclusions: In simulated testing, patients were able to successfully use the companion app without formal training. This study suggests the companion app is easy to use and could help patients prescribed CZP better manage their treatment and disease.

(JMIR Form Res 2020;4(7):e17373) doi:10.2196/17373

KEYWORDS

rheumatology; internet; digital health; mobile health; mHealth; smartphone; mobile phone; validation human factors study

Introduction

Anti-tumor necrosis factors (TNFs) are established and effective treatments for moderate to severe chronic inflammatory diseases [1]. The use of anti-TNFs alongside conventional disease-modifying antirheumatic drugs has been proven to result in better long-term disease control and reduced functional impairment [1].

Certolizumab pegol (CZP) is an Fc-free, PEGylated anti-TNF approved for use in 66 countries [2]. CZP is approved to treat adults with moderate to severe rheumatoid arthritis (RA), axial spondyloarthritis (including both radiographic and

nonradiographic axial spondyloarthritis), psoriatic arthritis, and plaque psoriasis in European Union (EU) countries and is also indicated for Crohn's disease in the United States [3,4]. In the EU, CZP treatment is administered via prefilled syringe or prefilled pen [5]. Recently, an electromechanical autoinjection device (electronic device [e-Device]), ava, has also been approved for use [3,5]. All CZP-injection devices were designed with patient input. The e-Device aims to provide patients with customizable features to improve patient satisfaction and overall self-injection experience [5].

Subcutaneous self-injection is associated with several specific challenges that can lead to reduced treatment adherence. Increasing use of smartphones and tablets provides a unique opportunity to address some of these challenges. For example, mobile phone apps can be designed to support patients and provide a means to overcome challenges such as forgetfulness [6]. Previous studies looking into the effectiveness of mobile phone health apps suggest that if well designed, they may provide an effective tool, empowering patients with long-term chronic conditions to self-manage their own health [7-9]. CimplyMe is a new mobile technology companion app designed to be used alongside any CZP self-injection device to further support and engage CZP-treated patients throughout their disease journey. By providing additional support, the companion app aims to improve patient satisfaction, treatment adherence, and resulting clinical outcomes.

The aim of this study was to validate the usability of the companion app by demonstrating that critical tasks can be performed successfully by a group of patients diagnosed with RA, representative of the intended end users.

Methods

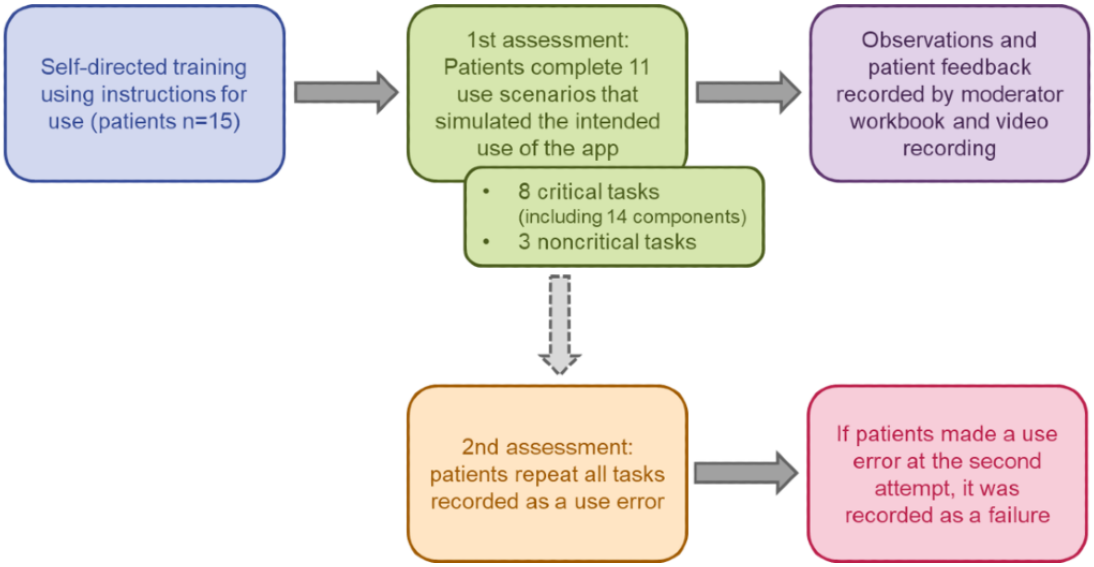
Study Design

This study assessed the usability of CimplyMe, a new mobile technology companion app designed to be used alongside any CZP self-injection device to further support and engage CZP-treated patients throughout their disease journey. The app provides step-by-step guidance through its timeline feature, providing information on the disease and CZP treatment. Guidance is focused on the initial 12 weeks of treatment but continues to support patients throughout their disease journey. Patients can track their pain, energy levels, mood, and activities using the Health Metrics Tracking feature. This is incorporated into the Health Summary that provides patients with an overview of their progress, including treatment adherence tracking, which they can choose to share with health care practitioners and their family. Milestones are included in the app and were designed to help patients acknowledge important moments in their disease journey and motivate patients to continue treatment. Additionally, injection and clinic appointment reminders aim to improve patient outcomes by supporting treatment adherence and disease monitoring.

This study was designed in line with ANSI/AAMI/IEC62366-1:2015 on applying human factors and usability engineering to medical devices [10]. Each participant gave informed consent before attending a single session; all sessions took place in London, United Kingdom. Figure 1 shows the study design. Patients were not given any formal training but had some time to familiarize themselves with the app. They were also able to use an instruction manual if requested, although the app was designed to be used without this document. Patients were asked to complete 11 use scenarios that were designed to simulate the intended use of the app. A moderator asked all patients to behave in the way they thought they would if the scenario was real. During the session, a moderator recorded all observations and patient feedback in a workbook. The sessions were also video recorded.

The primary study objective was to validate that the user interface of CimplyMe is intuitive, easy, and safe to use by intended users so as not to cause injury or harm. The secondary objective was to confirm that the risk of errors was reduced as far as reasonably possible through risk-mitigation methods implemented in the user interface (UI) of the beta version of the app used in this study.

Figure 1. Study design flowchart.



Study Procedures and Evaluations

Patients were assessed on their ability to complete 8 critical tasks composed of 14 component tasks and 3 noncritical tasks composed of 8 component tasks (Table 1). All participants were observed and timed while completing each task by study personnel. Each test session was scheduled to last for 105 minutes (1.75 hours); the moderator divided the time allocated between each task (between approximately 3 and 10 minutes depending on the task) to ensure that the participants could at least attempt each task once. Each component task had to be completed correctly; otherwise, a use error was recorded. If a patient struggled to complete the component tasks but avoided a use error, the task was recorded as a "success with difficulty." If a patient made a use error on their first attempt, they could attempt the task a second time. If a use error was made at the second attempt, it was recorded as a task failure. All patient actions that led to, or resulted from, successes with difficulty or use errors were recorded. Detailed root cause analysis was carried out on all tasks completed with difficulty or recorded as a use error to determine whether the cause was a result of the user interface.

Table 1. Critical and noncritical tasks.

Tasks	Use scenarios
Critical tasks	
1. Download companion app	Patients were asked to explain how they would go about downloading an app.
2. Jailbroken phone[a]	Patients were shown the app terms and conditions and asked to explain what the message means. They were then asked if the app can be used on a jailbroken phone.
3. Set up user profile (change language)	Patients were asked to change the language of the app.
4. Set and adjust medication reminder	Patients were asked to explain how they would set up medication reminders. They were asked to schedule an injection reminder and edit it, following instructions. Finally, they were asked to delete the injection reminder.
5. Manually enter injection dates	Patients were told to imagine they wanted to perform their injection ahead of schedule. They were given the date and locations of the injection they had just made and were asked to register the injection in the app.
6. Log CZP[b] injection date	Patients were asked to log an injection date that was outside of the logging time frame and asked if they understood the confirmation screen informing them that CZP was not logged within the correct time frame.
7. Add or edit a CZP medication schedule	Patients were asked to imagine their treatment plan has changed and asked to change their dose regimen.
8. Download app updates	Patients were asked what they would do if they opened the app and an "Updates Available" screen appeared.
Noncritical tasks	
1. Edit or delete clinical appointments	Patients were asked to imagine they wanted to schedule a physician appointment and given information to input. They were then asked to edit the information they just inputted and finally to delete the reminder.
2. Error messages	Several error messages from the app were presented. The patient was asked to explain what they thought the different messages mean.
3. Remove profile from the app	Patients were asked to remove their user profile from the app.

[a]A jailbroken phone is one that has bypassed restrictions to allow the user to download apps from websites other than the official app store.

[b]CZP: certolizumab pegol.

Participants

Adult patients (≥18 years) diagnosed with moderate to severe RA were invited to participate in the study. To ensure at least 97% of potential use errors were identified [11], a minimum of 15 intended users were recruited. To verify their RA diagnosis, patients were asked to provide a confirmation letter from their physician. In addition, patients had to have been prescribed biologic treatment at the time of enrollment. Patients must have been using a smart phone and had experience using apps, of any type, on their phone. Patients must also have stated that they would use a medical mobile app to assist with their disease treatment. The study aimed to recruit at least 7 patients with impaired vision, either correctable or not correctable with glasses or contact lenses.

Patients were excluded if they had participated in medical device research for a medical app in the previous 6 months, were app developers, or had stated that they would not use a mobile app to assist with their disease treatment. Patients were also excluded if they experienced <10 minutes of joint stiffness in the morning without treatment.

Results

Patient Disposition and Baseline Characteristics

Of the 30 patients who were screened to take part in the study, 15 patients completed the study. Of the 15 patients that did not take part, 8 were unable to provide a physician's letter, 4 were not able to take part due to scheduling, and 2 patients self-withdrew. One patient was excluded as they were known to the sponsor. Familiarity with the use of mobile apps was a stated requirement; therefore, there were no screen failures relating to this.

Patient baseline characteristics are shown in Table 2. The mean patient age was 53.5 years (SD 12.0 years), and 66.7% (10/15) patients were female. All patients were using either a prefilled syringe or prefilled pen to administer their current biologic medication.

Table 2. Patient baseline characteristics.

Parameter	RA[a] patients (n=15)
Age (years), mean (SD)	53.5 (12.0)
Gender (female), n (%)	10 (66.7)
Ethnicity, n (%)	
White British	11 (73.3)
Black British	2 (13.3)
British Asian	2 (13.3)
Dominant hand (right), n (%)	14 (93.3)
Highest education level achieved, n (%)	
Early years	0
Primary	0
Secondary	4 (26.7)
A-Level	1 (6.7)
College/university	6 (40.0)
Postgraduate/vocational	4 (26.7)
Vision (self-reported), n (%)	
Nearly perfect	4 (26.7)
Need glasses to read	3 (20.0)
Need glasses to see into the distance	2 (13.3)
Need glasses to read and see into the distance	6 (40.0)
Impaired vision and not correctable with glasses	0
Disease severity (moderate to severe), n (%)[b]	15 (100.0)
Morning joint stiffness (minutes in the morning), n (%)	
0[c]	1 (6.7)
≤30	4 (26.7)
30-60	6 (40.0)
>60	4 (26.7)
Joints affected by RA, n (%)	
Fingers	12 (80.0)
Hands	11 (73.3)
Wrists	8 (53.3)
Elbow	5 (33.3)
Current biologic medication, n (%)	
Etanercept	6 (40.0)
Certolizumab pegol	2 (13.3)
Adalimumab	6 (40.0)
Tocilizumab	1 (6.7)
Type of self-injection device used, n (%)	
PFS[d]	7 (46.7)
PFP[e]	8 (53.3)

[a]RA: rheumatoid arthritis.

[b]Physician-determined.

[c]When medicated; when unmedicated, stiffness ≥10 mins.

[d]PFS: prefilled syringe.

[e]PFP: prefilled pen.

Critical Task Successes and Use Errors

Throughout the study, 207 first attempts at critical component tasks were completed. Of these, 86.0% (178/207) tasks were successful; 7.7% (16/207) were successes with difficulty; and 6.3% (13/207) were use errors (Table 3). Twelve second attempts at critical tasks were completed in the study; one task was not re-attempted due to time constraints. No use errors were encountered when performing tasks for a second time, although 4 tasks were recorded as successes with difficulty (Table 3).

Reasons for patient difficulties and use errors during first and second attempts at critical tasks are shown in Multimedia Appendix 1. Several (6/13, 46.2%) first attempts classed as use errors were cases of patients not completing the task in the allotted session time. Patients were not given a set amount of time to complete each task; however, the total time of each session was limited to 110 minutes so not all patients were able to complete all tasks within this time frame.

Difficulties at the first attempt were due to various reasons; however, most (10/16, 62.5%) resulted from patients getting confused by the app terms, organization, or process (Multimedia Appendix 1). Additionally, one difficulty was caused by an app fault. There were 4 successes with difficulty during the 12 second attempts at the critical tasks. As with those observed at first attempt, most (3/4, 75.0%) of these difficulties were due to patients getting confused by the app terms or organization (Multimedia Appendix 1).

Table 3. Critical and noncritical task outcomes.

Task and attempt	Success	Success with difficulty	Use error
Critical tasks			
First (n=207)	178	16	13
Second (n=12)[a]	8	4	0
Noncritical tasks			
First (n=107)	104	3	0

[a]Only 12 patients re-attempted critical task 6, “log an injection date,” due to time constraints.

Root Cause Analysis of Critical Task Errors

The root cause analysis of all use errors is shown in Table 4. The UI contributed to 61.5% (8/13) of use errors. Of these, 42.9% (3/8) were attributable to the small font size in the instructions for use (IFU). In 37.5% (3/8) of the use errors, patients did not see or were not prompted to scroll to the necessary button(s). The remaining errors associated with the UI (2/8, 25.0%) were attributed to app complexity or lack of clarity.

Table 4. Root cause analysis of all use errors.

Task and patient actions	Root cause	Did the UI[a] contribute to the use error?
4. Set and adjust medication reminder		
P09: did not save the reminder; opened IFU[b] but did not read it after realizing it was 120 pages	Perception/cognition: expected the app to have a more assistive role in setting up a reminder	Yes: expected the app to do more and did not read the IFU
P02: tried to edit the reminder but deleted it; confused tracking reminder and injection reminder	Cognition: did not understand what "Tracking Reminder" meant	Yes: unsure what tracking reminder meant
5. Manually enter injection dates		
P03: did not register second injection	Perception/cognition: did not see the "Add Another" button and thought they would add both injections to the same image	No
P04: took a long time to understand the task	Perception/cognition: could not see or understand where to register an injection	No
P05: only registered one injection	Perception: did not see the "Add Another" button	Yes: did not see the "Add Another" button, only "Done"
P06: kept trying to register injections through health tracking	Cognition: did not know where to go to register injections	No
P10: did not scroll down to see "register injection"	Perception/cognition: did not see the button and did not think to continue scrolling	Yes: "Register Injection" was not visible on the screen, and there was nothing on the screen to suggest scrolling down
P12: registered an injection but could not register a second injection	Cognition: expected to add both injections on the same image	Yes: text was very small in the IFU, which led to the patient skimming the text
7. Add or edit a CZP[c] medication schedule		
P02: expected "Change Treatment Plan" to be on the timeline	Cognition: thought they were looking in the obvious place to look	No
P04: took a long time to complete the task	Cognition: struggled with the terminology of the app	No: patient struggled with the terminology throughout
P09: expected to edit their dose regimen through the calendar	Perception: did not see the section in profile	No
P11: not able to find where to edit the medication schedule	Perception/cognition: not able to find where to complete the task	Yes: patient said the complexity of the app made it more difficult
P04: thought the app was faulty when the confirmation screen appeared	Cognition: patient found it difficult to understand the language and reasoning behind the confirmation screen	Yes: text could have been bigger and use of colors clearer

[a]UI: user interface.

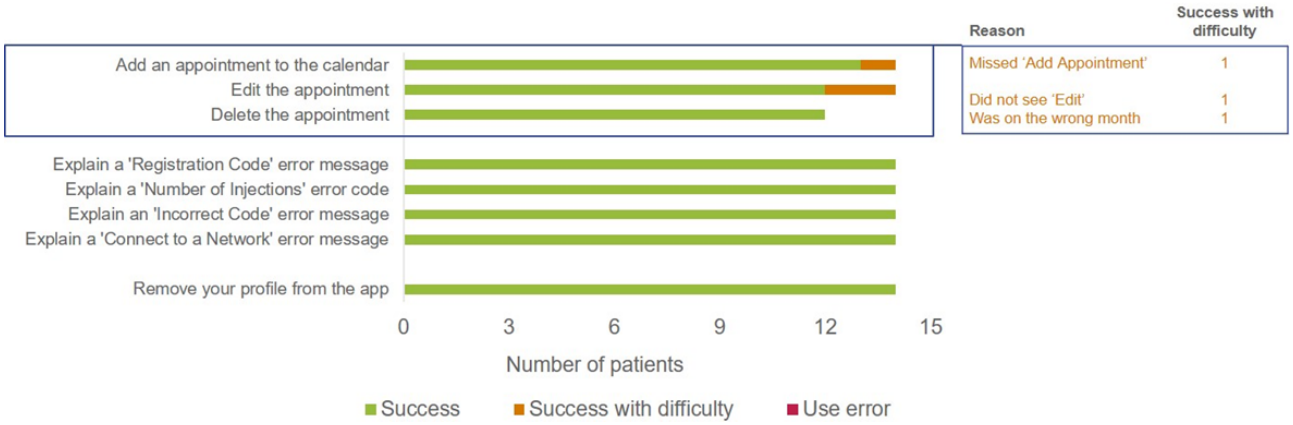
[b]IFU: instructions for use.

[c]CZP: certolizumab pegol.

Noncritical Task Successes and Use Errors

A total of 107 first attempts at the 3 noncritical tasks were performed in the study. Of these, 97.2% (104/107) were successful, 2.8% (3/107) were successes with difficulty, and none were use errors (Table 3). Reasons for patient difficulties when completing noncritical tasks are shown in Figure 2.

Figure 2. Reasons for noncritical task difficulties and use errors at first attempt. The number of patients who completed noncritical tasks was <15 as some patients did not have time to complete all tasks during the session.



Discussion

Principal Findings

No failures were observed for any of the attempted tasks; all patients were able to successfully complete both critical and noncritical tasks using the companion app on their first or second attempt without any formal training.

There were several instances where use errors could be attributed to the UI; these can all be resolved by small changes to the app. For example, patients did not realize that they needed to scroll further down the page to find the correct information or buttons in multiple instances. To minimize this problem, the scaling of the app could be optimized to ensure that most buttons are always on the screen. If future feedback suggests patients still have difficulties, a small note or animation could be added to indicate the need to scroll down on first use. Additionally, patients found the font size in the IFU and app too small or did not read the IFU because it was too long. However, the font size on the app is dependent on the default font size settings on the patient's phone and can be changed by the patient, if necessary. Two patients did not understand the terminology used in the app, or struggled with the complexity of the app itself, resulting in confusion and error in completing tasks and navigating the app. However, this was likely due to the short amount of time patients were given to familiarize themselves with the app. It is anticipated that patient understanding of the app's terminology and layout would improve with use, as indicated by the reduction in errors when tasks were attempted for a second time in this study.

Mobile technology apps have been developed to provide information and patient education, facilitate lifestyle changes as a strategy to ameliorate disease, provide medication reminders to promote treatment adherence, and record patient metrics for disease monitoring [12-15]. Previous research suggests that mobile technology apps can help patients feel supported and improve patient satisfaction [12,16]. For example, compared to usual care, a mobile health technology–supported disease management model for atrial fibrillation resulted in improvements in patient quality of life [17]. Patient adherence and disease management may also be improved with the support of treatment management apps [16,18,19]. For example, a 2019 systematic literature review of studies investigating mobile apps as a method to improve treatment adherence identified 11 studies of patients with a variety of diseases [18]. Of these, 7 studies reported an increase in treatment adherence with use of a mobile app [18].

One of the key challenges associated with the use of mobile apps is the high user attrition rate, with the level of mobile app engagement dropping off over time [20]. Understanding the features and characteristics most valued by patients may improve app engagement. Users appreciate features that save time over current methods and identify an app as valuable when it is simple and intuitive to use, provides specific instructions to better manage a condition, shares data with designated individuals, and incorporates motivational "push" factors [20,21]. Increasing age is also associated with reduced use of smartphones for internet browsing, social media, and apps [22]. Consequently, mobile app attrition is likely to be increasingly challenging in the population of patients likely to use mobile apps such as CimplyMe; patients with chronic inflammatory diseases are likely to be older and so are likely to be less comfortable using mobile apps. Designing apps with the user in mind and assessing usability in human factors studies will help ensure successful patient engagement.

Study Limitations

Human factors studies only simulate use in an everyday environment. Patients were asked to behave in the way they thought they would if the scenario was real; however, this may be difficult to judge accurately. Each session was limited to 105 minutes, and, as a result, patients may have felt under some time pressure to complete each task. Time constraints also restricted patients' ability to familiarize themselves with the app. Additionally, each task was completed discretely, which may have interfered with their familiarization of the app, owing to the lack of continuity between tasks. Patients were selected based on their willingness to use mobile apps, which may limit the generalizability of the results; however, the app will be provided as an additional support component and so use by patients will be voluntary. Therefore, the study population is reflective of the final population of patients who would likely use this in clinical practice. Finally, the number of patients participating in the study was small; therefore, the results may not represent the app's usability in a wider patient group or

capture all potential use errors and difficulties when using the app.

Conclusions

In a simulated setting, patients were able to use the companion app successfully. No critical task failures were observed; all patients were able to successfully complete both critical and noncritical tasks within two attempts without any formal training. This study demonstrates that the companion app UI is generally intuitive and easy to use among patients familiar with using apps. The app could support CZP patients to self-manage their treatment.

Acknowledgments

The authors thank the patients, the investigators, and their teams who took part in this study. The authors also acknowledge Susanne Wiegratz, UCB Pharma, Monheim am Rhein, Germany, for publication coordination and Oliver Palmer, BSc (Hons), Emma Phillips, PhD, and Simon Foulcer, PhD, from Costello Medical, United Kingdom, for medical writing and editorial assistance based on the authors' input and direction; their support was funded by UCB Pharma in accordance with Good Publication Practice (GPP3) guidelines (http://www.ismpp.org/gpp3). The development of the CimplyMe app and this study were sponsored by UCB Pharma.

Authors' Contributions

BD, SV, and IM contributed substantially to study conception and design as well as the analysis and interpretation of the data; drafted the article or revised it critically for important intellectual content; and approved the final version of the article to be published.

Conflicts of Interest

BD, SV, and IM are employees of UCB Pharma.

Multimedia Appendix 1
Reasons for critical task difficulties and use errors. Only 12 patients completed critical task 6 ("log an injection date") due to time constraints.
[PNG File , 267 KB - formative_v4i7e17373_app1.png]

References

1. Udalova I, Monaco C, Nanchahal J, Feldmann M. Anti-TNF Therapy. Microbiol Spectr 2016 Aug;4(4). [doi: 10.1128/microbiolspec.MCHD-0022-2015] [Medline: 27726761]
2. Curtis JR, Mariette X, Gaujoux-Viala C, Blauvelt A, Kvien TK, Sandborn WJ, et al. Long-term safety of certolizumab pegol in rheumatoid arthritis, axial spondyloarthritis, psoriatic arthritis, psoriasis and Crohn's disease: a pooled analysis of 11 317 patients across clinical trials. RMD Open 2019 May 31;5(1):e000942 [FREE Full text] [doi: 10.1136/rmdopen-2019-000942] [Medline: 31245056]
3. Cimzia Summary of Product Characteristics. European Medicines Agency. 2018. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/001037/WC500069763.pdf [accessed 2020-06-29]
4. Cimzia Prescribing Information. United Stated Food and Drug Administration. 2018. URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/125160s283lbl.pdf [accessed 2020-06-29]
5. Domańska B, Stumpp O, Poon S, Oray S, Mountian I, Pichon C. Using Patient Feedback to Optimize the Design of a Certolizumab Pegol Electromechanical Self-Injection Device: Insights from Human Factors Studies. Adv Ther 2017 Dec 8;35(1):100-115. [doi: 10.1007/s12325-017-0645-1]
6. Singh K, Drouin K, Newmark L, Rozenblum R, Lee J, Landman A, et al. Developing a Framework for Evaluating the Patient Engagement, Quality, and Safety of Mobile Health Applications. Issue Brief (Commonw Fund) 2016 Feb;5:1-11. [Medline: 26934758]
7. Salimzadeh Z, Damanabi S, Kalankesh L, Ferdousi R. Mobile Applications for Multiple Sclerosis: a Focus on Self-Management. Acta Inform Med 2019 Mar;27(1):12-18 [FREE Full text] [doi: 10.5455/aim.2019.27.12-18] [Medline: 31213737]
8. Mahmood A, Kedia S, Wyant DK, Ahn S, Bhuyan SS. Use of mobile health applications for health-promoting behavior among individuals with chronic medical conditions. Digit Health 2019 Oct 10;5:2055207619882181 [FREE Full text] [doi: 10.1177/2055207619882181] [Medline: 31656632]
9. Lee J, Choi M, Lee SA, Jiang N. Effective behavioral intervention strategies using mobile health applications for chronic disease management: a systematic review. BMC Med Inform Decis Mak 2018 Feb 20;18(1):12 [FREE Full text] [doi: 10.1186/s12911-018-0591-0] [Medline: 29458358]
10. Medical devices Part 1: Application of usability engineering to medical devices. ANSI/AAMI/IEC 62366-1. 2015. URL: https://my.aami.org/aamiresources/previewfiles/6236601_1503_preview.pdf [accessed 2020-04-20]

11. Faulkner L. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. Behavior Research Methods, Instruments, & Computers 2003 Aug;35(3):379-383. [doi: 10.3758/bf03195514]
12. Cahn A, Akirov A, Raz I. Digital health technology and diabetes management. Journal of Diabetes 2017 Nov 06;10(1):10-17. [doi: 10.1111/1753-0407.12606]
13. Santo K, Chow CK, Thiagalingam A, Rogers K, Chalmers J, Redfern J. MEDication reminder APPs to improve medication adherence in Coronary Heart Disease (MedApp-CHD) Study: a randomised controlled trial protocol. BMJ Open 2017 Oct 08;7(10):e017540 [FREE Full text] [doi: 10.1136/bmjopen-2017-017540] [Medline: 28993388]
14. Tabi K, Randhawa AS, Choi F, Mithani Z, Albers F, Schnieder M, et al. Mobile Apps for Medication Management: Review and Analysis. JMIR Mhealth Uhealth 2019 Sep 11;7(9):e13608 [FREE Full text] [doi: 10.2196/13608] [Medline: 31512580]
15. Santo K, Singleton A, Rogers K, Thiagalingam A, Chalmers J, Chow CK, et al. Medication reminder applications to improve adherence in coronary heart disease: a randomised clinical trial. Heart 2018 Aug 27;105(4):323-329. [doi: 10.1136/heartjnl-2018-313479]
16. Torbjørnsen A, Ribu L, Rønnevig M, Grøttland A, Helseth S. Users' acceptability of a mobile application for persons with type 2 diabetes: a qualitative study. BMC Health Serv Res 2019 Sep 06;19(1):641 [FREE Full text] [doi: 10.1186/s12913-019-4486-2] [Medline: 31492176]
17. Guo Y, Chen Y, Lane DA, Liu L, Wang Y, Lip GY. Mobile Health Technology for Atrial Fibrillation Management Integrating Decision Support, Education, and Patient Involvement: mAF App Trial. The American Journal of Medicine 2017 Dec;130(12):1388-1396.e6. [doi: 10.1016/j.amjmed.2017.07.003]
18. Pérez-Jover V, Sala-González M, Guilabert M, Mira JJ. Mobile Apps for Increasing Treatment Adherence: Systematic Review. J Med Internet Res 2019 Jun 18;21(6):e12505. [doi: 10.2196/12505]
19. Faridi Z, Liberti L, Shuval K, Northrup V, Ali A, Katz DL. Evaluating the impact of mobile telephone technology on type 2 diabetic patients' self-management: the NICHE pilot study. J Eval Clin Pract 2008 Jun;14(3):465-469. [doi: 10.1111/j.1365-2753.2007.00881.x]
20. Druce KL, Dixon WG, McBeth J. Maximizing Engagement in Mobile Health Studies. Rheumatic Disease Clinics of North America 2019 May;45(2):159-172. [doi: 10.1016/j.rdc.2019.01.004]
21. Mendiola MF, Kalnicki M, Lindenauer S. Valuable Features in Mobile Health Apps for Patients and Consumers: Content Analysis of Apps and User Ratings. JMIR mHealth uHealth 2015 May 13;3(2):e40. [doi: 10.2196/mhealth.4283]
22. Kumar D, Hemmige V, Kallen M, Giordano T, Arya M. Mobile phones may not bridge the digital divide: a look at mobile phone literacy in an underserved patient population. Cureus 2019 Feb 20;11(2):e4104 [FREE Full text] [doi: 10.7759/cureus.4104] [Medline: 31057998]

Abbreviations

CZP: certolizumab pegol
e-Device: electronic device
EU: European Union
IFU: instructions for use
RA: rheumatoid arthritis
TNF: tumor necrosis factor
UI: user interface

Edited by G Eysenbach; submitted 10.12.19; peer-reviewed by I Contreras, X Ji, N Heynsbergh; comments to author 10.03.20; revised version received 05.05.20; accepted 03.06.20; published 27.07.20.

Please cite as:
Domańska B, Vansant S, Mountian I
A Companion App to Support Rheumatology Patients Treated with Certolizumab Pegol: Results From a Usability Study
JMIR Form Res 2020;4(7):e17373
URL: http://formative.jmir.org/2020/7/e17373/
doi:10.2196/17373
PMID:32716310

Original Paper

A Web-Based Intervention to Prevent Multiple Chronic Disease Risk Factors Among Adolescents: Co-Design and User Testing of the Health4Life School-Based Program

Katrina Elizabeth Champion[1,2*], BAPsych (Hons), BHealth, PhD; Lauren Anne Gardner[1], BPsych (Hons), PhD; Cyanna McGowan[2], MPH; Cath Chapman[1], BA (Hons), PhD; Louise Thornton[1], BPsych (Hons), PhD; Belinda Parmenter[3], PhD; Nyanda McBride[4], PhD; David R Lubans[5], PhD; Karrah McCann[1], BPsych (Hons), MCrim; Bonnie Spring[2], PhD; Maree Teesson[1], BPsych (Hons), PhD; The Health4Life Team[6*]; Nicola Clare Newton[1], BPsych (Hons), PhD

[1]The Matilda Centre for Research in Mental Health and Substance Use, The University of Sydney, Sydney, Australia

[2]Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

[3]Department of Exercise Physiology, School of Medical Sciences, Faculty of Medicine, UNSW Sydney, Sydney, Australia

[4]National Drug Research Institute, Faculty of Health Sciences, Curtin University, Perth, Australia

[5]Priority Research Centre for Physical Activity and Nutrition, Faculty of Education and Arts, University of Newcastle, Callaghan, Australia

[6]See Authors' Contributions section

[*]these authors contributed equally

Corresponding Author:
Katrina Elizabeth Champion, BAPsych (Hons), BHealth, PhD
The Matilda Centre for Research in Mental Health and Substance Use
The University of Sydney
The Matilda Centre, Level 6, Jane Foss Russell Building (G02)
University of Sydney, NSW
Sydney, 2006
Australia
Phone: 61 286279006
Email: katrina.champion@sydney.edu.au

Abstract

Background: Chronic diseases are the leading cause of death worldwide. Addressing key lifestyle risk factors during adolescence is critical for improving physical and mental health outcomes and reducing chronic disease risk. Schools are ideal intervention settings, and electronic health (eHealth) interventions afford several advantages, including increased student engagement, scalability, and sustainability. Although lifestyle risk behaviors tend to co-occur, few school-based eHealth interventions have targeted multiple behaviors concurrently.

Objective: This study aims to summarize the co-design and user testing of the Health4Life school-based program, a web-based cartoon intervention developed to concurrently prevent 6 key lifestyle risk factors for chronic disease among secondary school students: alcohol use, smoking, poor diet, physical inactivity, sedentary recreational screen time, and poor sleep (the *Big 6*).

Methods: The development of the Health4Life program was conducted over 18 months in collaboration with students, teachers, and researchers with expertise relevant to the Big 6. The iterative process involved (1) scoping of evidence and systematic literature review; (2) consultation with adolescents (N=815) via a cross-sectional web-based survey to identify knowledge gaps, attitudes, barriers, and facilitators in relation to the Big 6; (3) content and web development; and (4) user testing of the web-based program with students (n=41) and teachers (n=8) to evaluate its acceptability, relevance, and appeal to the target audience.

Results: The co-design process resulted in a six-module, evidence-informed program that uses interactive cartoon storylines and web-based delivery to engage students. Student and teacher feedback collected during user testing was positive in terms of acceptability and relevance. Commonly identified areas for improvement concerned the length of modules, age appropriateness of language and alcohol storyline, the need for character backstories and links to syllabus information, and feasibility of implementation. Modifications were made to address these issues.

Conclusions: The Health4Life school-based program is the first universal, web-based program to concurrently address 6 important chronic disease risk factors among secondary school students. By adopting a multiple health behavior change approach, it has the potential to efficiently modify the Big 6 risk factors within one program and to equip young people with the skills and knowledge needed to achieve and maintain good physical and mental health throughout adolescence and into adulthood.

(JMIR Form Res 2020;4(7):e19485) doi:10.2196/19485

KEYWORDS

primary prevention; schools; eHealth; chronic disease; mobile phone; health promotion

Introduction

Background

Chronic diseases such as cardiovascular disease, type 2 diabetes, and mental disorders are the leading causes of death globally [1]. It is well established that major chronic diseases share 4 common lifestyle risk factors: poor diet, physical inactivity, smoking, and alcohol use [1,2]. Emerging evidence has also linked these established chronic disease risk factors with 2 additional risk behaviors: sedentary behavior (ie, sitting and recreational screen time) [3,4] and poor sleep (ie, long or short duration and/or poor quality) [5]. Together, these 6 risk factors (the *Big 6*) are important targets for prevention and early intervention programs to reduce chronic disease.

A life course approach to prevention includes intervening early to reduce risk factors *before the onset* of chronic disease [6]. Early adolescence is a critical period to intervene, as it coincides with the emergence of many health risk factors and is a time when young people acquire greater autonomy over their lifestyle choices [7]. To date, most prevention approaches have focused on changing single behaviors, despite risk factors commonly co-occurring [8]. For example, watching television is associated with high-fat snacking [9], and young people who engage in risky alcohol and other drug use are also more likely to eat poorly and be sedentary [7]. In recognition of this clustering, multiple health behavior change interventions [10] have been developed to address risk factors together, rather than in isolation. These interventions are underpinned by the *transfer theory* [11], whereby skills and knowledge learned about one behavior are thought to transfer to other contexts [12], resulting in improvements across multiple behaviors without additional intervention [13]. For example, an intervention targeting physical activity was also shown to improve eating habits [14].

Schools are ideal locations to deliver healthy lifestyle interventions, as they provide an opportunity to engage large numbers of students during this critical time period, and in Australia, schools are mandated to teach health education. Furthermore, outside of the family environment, the school is the primary setting within which the development of children and young people can be directed and shaped [15]. As teaching time is often limited, multiple health behavior change interventions that can simultaneously address multiple risk factors are particularly advantageous. However, evidence-based prevention programs are typically poorly implemented, often because of limited resources, a lack of teacher training and support, and program adaptations that undermine efficacy [16,17]. In fact, it has been estimated that only 14% of programs delivered in schools have the correct content and modes of delivery [18], and more recently, that teachers adapt between 50% and 68% of content in prevention interventions [17,19]. This may help to explain why interventions targeting health behaviors, such as physical activity, among adolescents have been largely unsuccessful [20]. Web-based interventions have the potential to overcome implementation barriers encountered by face-to-face interventions and offer several advantages, including increased student engagement, scalability, and program fidelity [21]. For example, core content is preprogrammed on the web, and completion is self-directed by students; therefore, delivery is not dependent on teacher training or skills, which promotes faithful delivery of key program components [22,23]. Research indicates that web-based interventions can be an effective means of preventing and reducing substance use among adolescents [21,24], and computer-based education has been shown to be superior to a generic classroom curriculum in increasing physical activity and improving diet [25]. Furthermore, schools are increasingly looking for digital web-based platforms to deliver key curriculum content as an alternative or adjunct to face-to-face teaching. However, there are no existing web-based interventions that adopt a multiple health behavior change approach to address the Big 6 among school students.

The effective *Climate Schools* prevention programs [26-28] use interactive, web-based cartoons about a group of teenagers to engage students; summaries and optional class activities to reinforce key content; and principles of social influence theory, namely, information provision, normative education, and resistance skills training [29], to prevent risky substance use and associated harms among adolescents [26-28]. Using the *Climate Schools* programs as a model, we developed the Health4Life school-based program to address gaps in the field. Underpinned by a multiple health behavior change approach [10], Health4Life aims to empower adolescents to reduce chronic disease risk and improve current physical and mental health by providing simultaneous web-based education about the Big 6 lifestyle risk factors. The intervention consists of 3 components: (1) a school-based program delivered to all grade 7 students (aged approximately 12 years) during class time, regardless of risk (universal prevention); (2) an accompanying smartphone app, available to all students regardless of risk (universal prevention); and (3) additional web and app content delivered outside of school to students who remain at risk of chronic disease 1 and 2 years after initial intervention delivery (selective prevention).

Objectives

This study aims to describe the co-design and user testing of the universal in-class school-based component of the

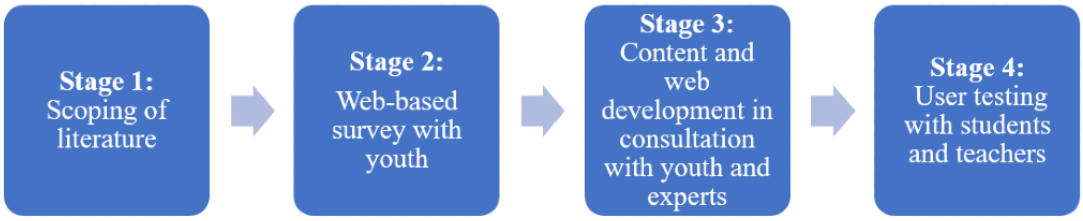
Health4Life intervention to aid in the development of future web-based prevention resources.

Methods

Co-Design

The Health4Life school-based program was developed using an iterative co-design process consisting of 4 key stages (Figure 1). A co-design approach was used to engage young people in the co-creation and refinement of the program. Co-design allows for a richer understanding of user needs and may be particularly important in fostering the engagement and satisfaction needed for web-based interventions to succeed among young people [30]. The Health4Life school-based program was co-designed with youth and teachers, with development coordinated by a research team comprising experts in addiction, physical activity, exercise physiology, sleep, dietetics, mental health, electronic health (eHealth) interventions, and behavior change.

Figure 1. Co-design of the Health4Life school-based program.



Stage 1: Scoping of the Literature

Although guided by an overarching multiple health behavior change approach [10], scoping of the literature was conducted to identify evidence-based prevention principles and best practice for the prevention of each of the Big 6 behaviors. Specifically, members of the research team with expertise in each of the 6 health behaviors were consulted to identify key behavioral theories and seminal school-based prevention papers and systematic reviews concerning each behavior. In addition, searches were conducted to identify national guidelines or recommendations, along with any further school-based prevention papers or systematic reviews and behavioral theories that could guide development. The key behavioral theories and evidence-based prevention principles that were identified and embedded in the Health4Life school-based program are outlined in Table 1.

Table 1. Key behavioral theories and evidence-based prevention principles for each of the Big 6 behaviors.

Big 6 behavior	Key behavioral theories	Evidence-based prevention principles
Alcohol use and smoking	• Social influence theory • Social learning theory • Social cognitive theory	• Normative education • Resistance skills training • Information and knowledge provision • Using peer leaders • Harm minimization • Life skills training, for example, decision making, problem-solving, coping skills, refusal and assertion skills, self-esteem, and self-control
Physical activity	• Self-determination theory • Social cognitive theory	• Development of competence, relatedness and social connection, and autonomy • Promotion of autonomous motivation (engaging in behavior that is valued, personally relevant, and enjoyable) • Self-regulatory skill development, for example, goal setting, self-monitoring, and decision making
Sedentary recreational screen time	• Self-determination theory • Social cognitive theory	• Development of competence, relatedness and social connection, and autonomy • Self-regulatory skill development, for example, goal setting, self-monitoring, and decision making
Sleep	• Two-process model of sleep • Social cognitive theory	• Teaching biological contributing factors to sleep timing and duration • Maintaining regular sleep patterns and identifying sleep problems • Self-regulatory skill development, for example, goal setting, self-monitoring, and decision making
Diet	• Social cognitive theory	• Self-regulatory skill development, for example, goal setting, self-monitoring, and decision making

In addition, to better understand the effectiveness of interventions targeting multiple risk behaviors, we conducted a systematic review and meta-analysis of eHealth school-based multiple health behavior change interventions targeting two or more of the Big 6 risk factors [31]. A total of 22 publications assessing 16 interventions were included in the review. Most of the studies assessed interventions to prevent alcohol use and smoking or improve diet and physical activity. Few studies targeted screen or sitting time, and none addressed sleep. Pooled findings supported the effectiveness of eHealth school-based

multiple health behavior change interventions for improving both accelerometer- (standard mean difference [SMD] 0.33; 95% CI 0.05 to 0.61) and self-report–(SMD 0.14; 95% CI 0.05 to 0.23) measured physical activity, screen time (SMD −0.09; 95% CI −0.17 to −0.01), and fruit and vegetable intake (SMD 0.11; 95% CI 0.03 to 0.19). However, the effects were small and short lived. No effects were seen for alcohol or smoking, fat, or sugar-sweetened beverage or snack consumption.

Interventions varied greatly in their content and delivery. Few studies described specific behavior change techniques, and none of the studies referred to an established behavior change taxonomy [32], making it difficult to tease apart effective intervention components that could be embedded into future programs. However, effective programs tended to use computer-based tailored feedback, whereby students were provided with individualized, normative, or stage-matched feedback based on self-reported responses to computer-based assessments, suggesting that this may be an important component to include in future interventions.

Results from the systematic review highlighted the need to develop multiple health behavior change interventions that address sleep problems among youth, as none of the included interventions addressed sleep. This is despite the growing recognition that physical inactivity, sedentary behavior, and sleep are codependent [33,34] and the fact that many young people report sleep problems [35]. Generally, ineffective interventions for alcohol use and smoking were brief, primarily used the transtheoretical model [36], and did not provide sufficient opportunities for skill building. It was concluded that future eHealth multiple health behavior change interventions targeting alcohol and tobacco use alongside other risk behaviors should be guided by principles of effective substance use prevention (eg, normative education and life skills training). Finally, the findings indicated that eHealth multiple behavior change interventions were only effective in increasing physical activity among students aged 13 years or older, and boys had a greater increase in physical activity than girls. This suggests that such interventions might not adequately engage girls or young teenagers (aged <13 years) and that formative research is required to better understand the beliefs, attitudes, and motivations of these groups regarding physical activity.

Stage 2: Web-Based Survey With Adolescents

A web-based self-report survey was conducted to understand young people's knowledge about the Big 6, their current engagement with health behaviors, their beliefs and attitudes about health, and barriers and facilitators to achieving good health.

Participants and Procedure

A total of 7 independent secondary schools in New South Wales and the Australian Capital Territory of Australia were invited to participate. Of the 7 schools, 3 agreed to participate. Participating schools were asked to distribute information and consent forms to the parents or guardians of their grade 7 to 9 students. Passive parental consent and active student consent were required for youth to be eligible to participate in the study (99% consent rate). Students completed an anonymous web-based survey in a supervised classroom setting between August 2018 and September 2018. Participants who completed the survey were entered into the draw to win a Fitbit valued at Aus $450 (US $315), with one prize given per school. Ethics approval was obtained from the University of New South Wales Human Research Ethics Committee (HREC; HC180224).

Measures

Demographic data collected included age, sex, self-reported height and weight, and postcode. Self-reported health status was measured using a single item, "How would you rate your own health?," with responses made on a 5-point Likert scale ranging from "Poor" to "Excellent."

Physical Activity

To assess moderate-to-vigorous physical activity (MVPA) students were asked, "Over the past 7 days, how many days did you do moderate or vigorous intensity physical activity for at least 60 minutes per day?" [37]. Participants were provided with a written description of what constitutes MVPA.

Screen Time

Using a modified version of the Adolescent Sedentary Activity Questionnaire [38], students were asked to recall the amount of time (hours and minutes) typically spent on recreational screen time on weekdays and weekend days.

Fruit and Vegetable Consumption

Using validated short items commonly used in health research [39,40], fruit intake was assessed via a single item: "About how many serves of fruit do you usually have each day?" ("don't eat fruit" to "6 serves or more"). A similar item measured vegetable intake. Participants were provided with information about what constitutes one serve of fruit or vegetables.

Sleep

Students were asked to report their usual bedtime and wake time on school nights and weekend nights. Sleep duration (in hours) was calculated as the difference between wake time and bedtime. Self-reported estimates of bedtime, wake time, and sleep duration have been shown to be reliable and valid [41].

Alcohol and Tobacco Use

Students were asked to report if they had ever tried alcohol or tobacco, based on validated items used in previous school-based trials [28,42].

Attitudes, Knowledge, Barriers, and Facilitators

A series of open-ended questions were asked to understand attitudes toward health, for example, "What does good health mean to you?" Students were asked 5 multiple-choice questions to test their knowledge about age-appropriate national recommendations for the Big 6 risk behaviors. These items were based on national guidelines for each behavior [34,43-45]. As there are no official national guidelines for smoking, participants' knowledge could not be assessed. Open-ended questions were used to assess key motivational factors, barriers, and facilitators of health, for example, "What gets in the way, or stops you, from being more physically active?"

Analysis

Data were collated and analyzed using IBM SPSS Statistics 24 (IBM Corp). Descriptive analyses were conducted to illustrate the sample characteristics, prevalence rates of the Big 6, and knowledge. For the collected open-ended data, we selected a subsample of student responses (20%-25%) and carried out a qualitative analysis on these responses until no new themes emerged (ie, data saturation was reached). The sample was stratified by age and year group, and a random subsample was selected to ensure balanced representation across age and year groups. Using an inductive approach, one author (LG) coded responses from the subsample, examined the data for frequent or significant responses, and grouped them according to key themes or categories.

Findings

Health Behaviors and Knowledge

A total of 815 students (mean age 13.89 years, SD 0.89; 84.3% [687/815] female) from 3 schools (2 coeducational and 1 female only) completed the survey. The majority of participants perceived their health to be "very good" (334/801, 41.7%) or "good" (274/801, 34.2%); however, adherence to national guidelines for the Big 6 was mixed. Only 12.9% (101/784) of participants achieved the recommended amount of MVPA, and only 11.8% (92/779) of participants reported eating enough vegetables per day. The majority of students (622/779, 79.8%) reported eating sufficient fruit serves, and approximately half of the participants (382/779, 49.0%) met guidelines for screen time on weekdays; however, only 23.0% (179/779) of participants met guidelines on the weekend. A small proportion of students (29/778, 3.7%) had tried tobacco in their lifetime, and 64.0% (498/779) of students had used alcohol in their lifetime (including a taste or sip). Among students aged 14 years or older, 70.1% (262/374) met the guidelines of sleeping 8 to 10 hours; however, only 53.0% (222/419) of students aged 12 to 13 years met their guideline of 9 to 11 hours. Knowledge of the recommended guidelines was poor for physical activity, diet, and sleep (Table 2); however, most students (719/786, 91.5%) correctly identified that for adolescents, the safest option is not to drink alcohol at all, and 98.1% (771/786) of students agreed that smoking is harmful.

Table 2. Percentage of students who correctly identified national guidelines.

Risk behavior and Australian guidelines for young people	Value, n (%)
Physical activity (n=798)	
Accumulate 60 min or more of moderate-to-vigorous physical activity per day	212 (26.6)
Sedentary recreational screen time (n=792)	
No more than 2 hours per day	394 (49.7)
Sleep	
9-11 hours of uninterrupted sleep per night for those aged 5-13 years (n=405)	77 (19.0)
8-10 hours per night for those aged 14-17 years (n=351)	210 (59.8)
Diet (n=789)	
Minimum of 2 serves of fruit per day	279 (35.4)
Minimum of 5 serves of vegetables per day	266 (33.7)
Alcohol use (n=786)	
For children and young people aged younger than 18 years, not drinking alcohol is the safest option	719 (91.5)

Attitudes, Barriers, and Facilitators

The key themes extracted from the open-ended responses are illustrated in Table 3. Many of the student responses aligned with the key behavioral theories and prevention principles outlined in stage 1. For example, consistent with self-determination theory, key facilitators of physical activity included social connection and intrinsic motivation. In addition, in line with social influence and social learning theories, the role of peers and older influences was prominent when considering reasons to drink alcohol and smoke cigarettes. These findings further shaped the storylines and guided the core content in the Health4Life program (more information is provided in the *Stage 3: Content and Web Development* section).

Table 3. Summary of key themes extracted from the open-ended responses of students.

Theme	Example
The meaning of good health	
Physical health factors such as being physically fit or active and having a good diet	"Eating well and just keeping active." (Male, 14 years)
Mental health factors such as having a positive mental state and feeling well	"Being active and eating healthy, but also to have a positive and healthy mindset." (Female, 13 years)
Emotional health factors such as feeling happy	"Good health to me means being happy. It also means being healthy by exercising and eating the right foods. But mainly I think it means being happy." (Female, 13 years)
Physical activity motivators	
Outcomes of being active such as the positive emotions felt afterward and maintenance of fitness and health	"I know that when I do exercise it puts me in a better mood which motivates me because who doesn't want to be in a good mood." (Female, 14 years)
Social factors such as friends, family, and teammates	"My friends and family. They help encourage me to take part in sport and also get me outside of the house for a run." (Female, 12 years)
Enjoyment and fun derived from physical activity and sports	"What motivates me is my liking and love for sports. I find most sports extremely fun to play." (Female, 12 years)
Physical activity barriers	
Lack of time because of school, homework, or assignments and other commitments	"My schedule is full with other work [homework, study for exams] and also I have commitments like family and community work so I don't always have enough time to work out." (Male, 14 years)
Laziness or a lack of motivation and feeling tired	"Things such as not having enough sleep, not being motivated enough and being too tired can contribute." (Female, 13 years)
Reasons for young people smoking	
To look or feel cool	"It is seen as the cool thing to do or a way to make you seem older than you are." (Male, 14 years)
Peer pressure or to fit in	"Peer pressure from people either older than them or their age." (Female, 12 years)
Curiosity or just to try it	"Some people my age smoke probably because they want to seem like a 'cool' kid and also try to experience what it tastes like." (Female, 14 years)
No reason as people do not smoke at that age	"I don't think they smoke." (Female, 12 years)
Reasons for young people drinking alcohol	
To look or feel cool	"I think people my age drink alcohol because they want to seem cool and mature, as alcohol is an adult thing to do." (Female, 13 years)
Peer pressure or to fit in	"Because of peer pressure, thinks it's cool and will achieve friends loyalty." (Male, 13 years)
Curiosity	"People my age drink alcohol mostly out of curiosity and because they want to see what it tastes like." (Female, 12 years)
No reason as people do not drink alcohol at that age	"I don't think people my age drink alcohol." (Female, 13 years)
Barriers to getting enough sleep	
Using technology and devices	"...because they are on their technology just before bed, making it hard to get to sleep." (Female, 13 years)
Other commitments taking their time such as homework or extracurricular activities	"I think that people my age do not sleep because they stay awake doing homework or schoolwork. Assignments, projects." (Male, 14 years)

Stage 3: Content and Web Development

Background and Content Development

The development of the Health4Life school-based program was based on the effective *Climate Schools* substance use prevention programs [26-28], which use web-based cartoon storylines about a group of teenagers to deliver key prevention messages and engage students. Underpinned by a multiple health behavior change approach [10] and guided by the key behavioral theories and prevention principles for each of the Big 6 behaviors identified in stage 1, the findings from the web-based youth survey conducted in stage 2, and consultation with researchers with expertise in the Big 6, the core content areas were developed, which formed the basis for the 6-module Health4Life program (Table 4).

Table 4. Summary of Health4Life program content.

Module	Key messages
1	• Guidelines for eating healthily and benefits of a healthy diet • Tips for increasing water intake • Sleep needs of adolescents and the benefits of sleeping well • Guidelines for recreational screen time and the benefits of limiting screen use
2	• Prevalence and patterns of alcohol and tobacco use among adolescents • Australian guidelines to reduce health risks from drinking alcohol • Identifying reasons why teenagers choose to, or not to, drink alcohol • Benefits of being physically active • Finding physical activities that you enjoy
3	• Short- and long-term consequences of alcohol and tobacco use • Consequences of excessive sedentary recreational screen time • Strategies for reducing sedentary recreational screen time • Responsible use of social media
4	• Social, financial, and legal consequences of alcohol and tobacco use • Assertive communication skills and refusal skills • Australian guidelines for physical activity and sedentary behavior • Tips for setting specific, measurable, achievable, relevant, time-bound (SMART) goals
5	• Understanding food labels • Limiting sugar-sweetened beverage consumption • Improving sleep hygiene • Avoiding too much sleep on weekends (“social jet lag”)
6	• Associations and interrelations between health habits • Relationships between the Big 6 and mental health • Physical, social, and emotional benefits of health and well-being • The Big 6 and long-term health

Character Development and Script Writing

The first step in designing the cartoon modules was to develop character profiles. The characters were purposely designed to be of similar age to the target audience (ie, grade 7 students, aged approximately 12 years) and to have different strengths and weaknesses and various health behavior clusters. Basic character profiles were developed with and reviewed by young people (n=7; aged 12-15 years; 2 males and 5 females) recruited via personal and professional networks. The youth reviewers were asked to make suggestions about character names, appearance, personalities, and health habits and to provide ideas for age-appropriate scenarios to be embedded into the storylines. Next, based on these characters and the findings from the web-based youth survey (stage 2), initial cartoon scripts for the 6 modules were written by members of the research team. The scripts aimed to provide students with information and skills about the Big 6 through the context of a teenage drama. The script writing was an iterative process, undergoing review by the expert researchers (n=22) and young people (n=9) before being sent for animation. Youth reviewers were asked to comment on the language used and relevance of the storylines, whereas expert reviewers commented on the accuracy of the content and prevention messages. Scripts were modified accordingly and then sent to animators to develop into the cartoon animations.

Module Summaries and Activities

Consistent with the *Climate Schools* programs [26-28], teacher and student summaries and optional class activities were developed, in addition to the cartoon storylines, for each of the 6 modules. Module summaries were developed to reinforce the key content taught within the cartoons. The summaries were written by members of the research team, drawing on the latest available evidence, including national prevalence data, for example [46,47], national guidelines, for example [48], and recent scientific literature relating to the Big 6 risk factors, for example, benefits of engaging in healthy behaviors and consequences of engaging in risky behaviors. In addition, a set of 4 to 5 activities was developed for each module, including points for class discussions, worksheets, quizzes, and homework tasks. Activities were designed to develop key self-management and interpersonal skills, such as communication, decision making, and problem-solving, and to meet key outcomes from the stage 4 New South Wales Personal Development, Health and Physical Education syllabus, the Western Australian Year 7 Health and Physical Education syllabus, and the years 7 and 8 Australian Health and Physical Education curriculum (for Queensland). Each module includes one web-based activity, such as interactive worksheets, games, or quizzes, to provide an engaging and immersive learning experience. Summaries and activities for each module were reviewed by members of the wider Health4Life team, including those with expertise relevant to the Big 6 behaviors.

Stage 4: User Testing

User testing was conducted in 2019 with year 7 students and teachers to gain feedback on the acceptability and feasibility of the Health4Life school-based program before implementation. Ethical approval was obtained from the University of Sydney HREC (2018/882).

Student Testing: Participants and Procedure

A selection of secondary schools in Sydney that had previously expressed interest in participating in research were invited to participate in user testing of the Health4Life program, of which 2 agreed to participate. A total of 2 focus groups (45-60 min) were conducted with grade 7 students in 2 independent secondary schools in Sydney: one all male (n=21; mean age 12.05 years, SD 0.38) and one all female (n=20; mean age 12.10 years, SD 0.31). Consent forms were distributed to parent or guardians, with passive parental consent and active student consent required. Facilitated by 2 researchers, the focus group consisted of students watching a video presentation of modules 1 and 2 cartoons as a group. Following each cartoon, the facilitators led a class discussion, prompting students with questions regarding the likability and relatability of the characters, storylines, and language used in the cartoons. The focus groups were audio recorded and later transcribed. At the end of the discussion, students were also asked to complete a brief, 10-min questionnaire to provide their written feedback about the acceptability and relevance of the Health4Life modules. Students were remunerated a Aus $30 (US $20) gift voucher for their time spent in the focus group. Students were also invited to provide feedback on the remaining 4 modules via a web-based survey for an additional reimbursement of one Aus $15 gift (US $10) voucher per module.

Teacher Testing: Participants and Procedure

All grade 7 health education teachers at the same 2 participating schools were invited to participate in the user testing process. Additional teachers were recruited via social media and personal and professional networks. A total of 8 teachers from 5 schools agreed to participate. After providing informed consent, teachers were invited to view 2 Health4Life modules (including cartoons, module summaries, and activities) and asked to complete a web-based survey to comment on the acceptability and suitability of the Health4Life school-based program. The teachers were reimbursed Aus $50 (US $35) for each module they reviewed. In addition, health education curriculum experts (n=6) from New South Wales, Western Australia, and Queensland were engaged to provide expert advice on the alignment of the Health4Life content with the national- and state-based health education curricula.

Analysis

Student and teacher questionnaire data were analyzed using IBM SPSS Statistics 24. Descriptive statistics were conducted to calculate the percentage agreement for each survey item. The focus group data were transcribed by one researcher (EH). Individual comments and specific recommendations were extracted to inform refinements and modifications to the Health4Life program.

Results

User Testing Results

A total of 41 students (mean age 12 years, SD 0.35; 49% [20/41] female) provided feedback about the Health4Life school-based program. As the review of modules 3 to 6 occurred outside of the classroom and was optional, the response rate ranged from 39% (16/41) to 100% (41/41) across modules. Overall, students rated the cartoon modules favorably, providing positive feedback about the likability and believability of the storylines and characters (Table 5). An example of open-ended feedback provided by students is summarized in Multimedia Appendix 1.

Table 5. Summary of student questionnaire data.

Question	Module 1 (n=41), n (%)	Module 2 (n=41), n (%)	Module 3 (n=31), n (%)	Module 4 (n=26), n (%)	Module 5 (n=18), n (%)	Module 6 (n=16), n (%)
Overall rating (% good/very good)	36 (88)	37 (90)	30 (97)	24 (92)	16 (89)	14 (88)
How much did you like the storylines? (% liked a little/lot)	35 (85)	38 (93)	30 (97)	24 (92)	18 (100)	13 (81)
How much did you like the characters? (% liked a little/lot)	29 (71)	34 (83)	28 (90)	24 (92)	17 (94)	12 (75)
How believable and realistic were the storylines? (% completely/somewhat)	31 (78)	32 (82)	26 (87)	23 (88)	16 (89)	13 (81)
Do you think that other year 7 students will understand the information in the lessons? (% strongly agree/agree)	30 (77)	—[a]	22 (73)	22 (85)	14 (78)	14 (88)
Do you think other year 7 students will like the characters? (% strongly agree/agree)	28 (72)	—	20 (67)	20 (77)	14 (78)	11 (69)
Do you think that other students will find the cartoons an engaging way to learn? (% strongly agree/agree)	32 (82)	—	21 (70)	20 (77)	11 (61)	10 (63)

[a]Students were asked to respond to questions 5-7 about both modules 1 and 2, so separate data for each module were not available.

A total of 8 teachers (75% male) participated in user testing: 7 taught at coeducational schools and 1 taught at a single-sex school. On average, participants had been teaching for 11.5 (SD 10.09) years. Table 6 summarizes the feedback provided by teachers via the web-based questionnaire. Overall, the teachers rated the Health4Life program positively, with the percentage of teachers rating the modules as good or very good ranging from 88% to 100%. The vast majority of teachers also indicated that the modules were a good fit with the health education syllabus and that the storylines were believable, age appropriate, and understandable to students. However, some teachers had concerns about the modules adequately covering content, having sufficient lesson time, and feasibility of implementation. Multimedia Appendix 2 provides key themes and examples extracted from the open-ended teacher feedback.

Table 6. Summary of teacher questionnaire feedback.

Feedback item	Module 1 (n=8), n (%)	Module 2 (n=8), n (%)	Module 3 (n=7), n (%)	Module 4 (n=7), n (%)	Module 5 (n=7), n (%)	Module 6 (n=7), n (%)
Overall rating (% good/very good)	7 (88)	7 (88)	7 (100)	7 (100)	7 (100)	7 (100)
Fit with syllabus (% very well/extremely well)	6 (75)	6 (75)	6 (86)	6 (86)	7 (100)	7 (100)
Believable and valid storylines (% yes)	7 (88)	8 (100)	7 (100)	7 (100)	6 (86)	7 (100)
Age-appropriate content (% yes)	8 (100)	7 (88)	7 (100)	7 (100)	7 (100)	7 (100)
Concepts understood/remembered by students (% yes)	8 (100)	8 (100)	6 (86)	6 (86)	7 (100)	7 (100)
Lesson length adequately covers content (% yes)	4 (50)	5 (63)	3 (43)	4 (57)	4 (57)	3 (43)
Language acceptable to youth (% yes)	8 (100)	7 (88)	7 (100)	7 (100)	7 (100)	7 (100)
Concerns about language used (% yes)	1 (13)	1 (13)	1 (14)	1 (14)	0 (0)	0 (0)
Effectiveness of program improving students' health behaviors (% somewhat effective/very effective)	8 (100)	8 (100)	7 (100)	7 (100)	7 (100)	7 (100)
Likely to have implementation problems (% yes)	5 (63)	5 (63)	4 (57)	4 (57)	1 (14)	1 (14)
Sufficient time to deliver modules (% yes)	6 (75)	6 (75)	4 (57)	4 (57)	4 (57)	4 (57)
Sufficient computers for students (% yes)	7 (88)	7 (88)	6 (86)	6 (86)	5 (71)	5 (71)

Key Changes Made to the Health4Life Program After User Testing

On the basis of feedback obtained from students and teachers, several modifications were made to the Health4Life program, including changes to the language, a reduction in cartoon text and slides, the addition of a backstory module and links to syllabus overviews, and the provision of alternate delivery methods (Table 7).

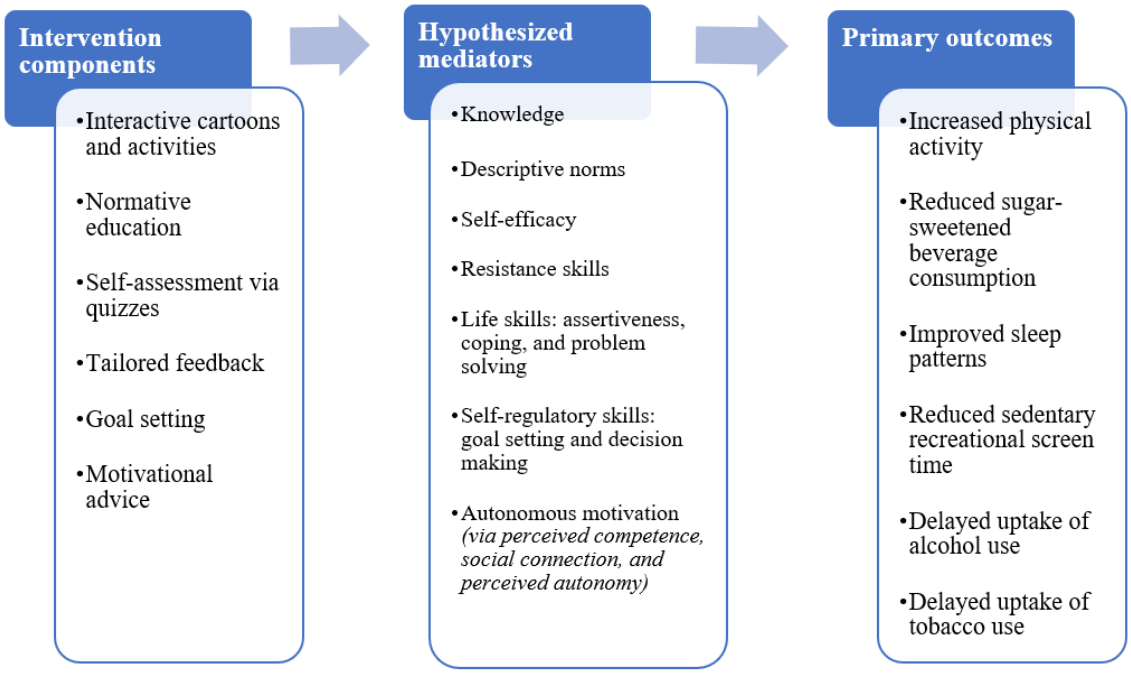
Table 7. Examples modifications made to the Health4Life program.

Key issues identified	Modifications
Lesson length • Text-heavy cartoons • Too many cartoon slides • Concern over fitting the cartoons and activities into one lesson	• Scripts were revised to make wording more succinct, and cartoon slides were removed where possible. • It was made clear to teachers that the class activities are optional and that a range of activities are provided so that the most suitable and feasible options can be selected.
Cartoon content • Age appropriateness of language • Need for a backstory • Relatability of alcohol storyline for year 7 students	• Language was refined based on student suggestions (eg, changing "okay" to "kk" in text messages). • An introductory cartoon module was added to provide a backstory for each of the main characters. • The script was adapted to make it clear that, aside from Xavier (year 7), the only other characters drinking alcohol are older (year 10).
Linking content to the health and physical education curricula	• Expert curriculum consultants were engaged to develop or review unit overviews that map the cartoon content and activities to the relevant syllabus outcomes and descriptors for each state. • An outline of how the components of the Health4Life program align with the different stages of learning (knowledge, understanding, skills, and application) and the 5 propositions (taking a strengths-based approach, focusing on educative purposes, valuing movement, developing health literacy, and including a critical inquiry approach) that shape the health and physical education curricula was provided.
Implementation feasibility for schools with limited access to computer rooms or devices	• Teachers were provided with alternative delivery options, such as using a smart board to go through the modules as a class or using hard copies of the cartoons and activities.

The Final Health4Life School-Based Program

The Health4Life school-based program is underpinned by a multiple health behavior change approach [10]. It is anticipated that by providing students with concurrent education about the Big 6, while also highlighting the associations and interrelations between health behaviors, the program will efficiently facilitate change across multiple behaviors. The program uses principles of social influence [29], social cognitive [49], social learning [50] and self-determination theories [51], and the two-process model of sleep [52,53] (Table 1 and Figure 2), with key behavior change components and mechanisms for change woven into the storyline and accompanying activities. The Health4Life school-based program was developed to meet outcomes from the stage 4 New South Wales Personal Development, Health and Physical Education syllabus, the Western Australian Year 7 Health and Physical Education syllabus, and the years 7 and 8 Australian Health and Physical Education curriculum (for Queensland). Teachers are provided with unit overviews that map the cartoon content and activities to the state-based syllabus outcomes and descriptors, along with an outline of how the components of the Health4Life program align with the different stages of learning (knowledge, understanding, skills, and application) and the 5 propositions (taking a strengths-based approach, focusing on educative purposes, valuing movement, developing health literacy, and including a critical inquiry approach) that shape the health and physical education curriculum. An overview of the program content is provided in Table 4.

Figure 2. Health4Life Conceptual model.



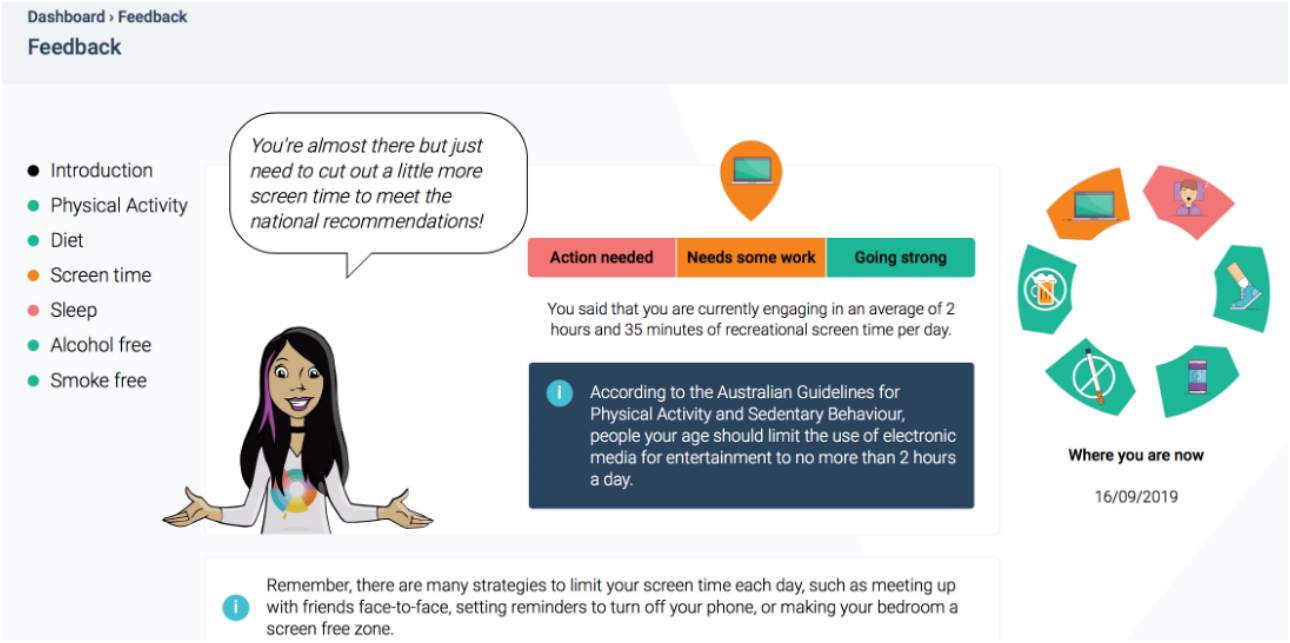
The Health4Life school-based program consists of 6 modules, ideally delivered once per week, each comprising the following:

1. A 20-min web-based cartoon completed individually by students (Figure 3): The cartoons follow the story of a group of teenagers, of similar ages to grade 7 students, to impart key prevention messages while engaging students and maintaining interest. The cartoon modules aim to provide evidence-based information about the Big 6, improve resistance skills, modify existing norms, and increase autonomous motivation. Short web-based quizzes are embedded at the end of each cartoon module to test knowledge.
2. Module summaries: teachers and students are provided with PDF factsheets for each module to provide additional details and reinforce key messages.
3. Optional activities: teachers are provided with a selection of 4 activities (eg, worksheets, group discussions, and homework tasks) to implement with their students after the cartoon in remaining lesson time (approximately 20 min). Teachers can select which and how many activities they implement to best suit the needs of their class. A self-directed interactive web-based activity (eg, web-based worksheet and game) is also included after each cartoon module to accommodate students completing the cartoon component at different speeds.

On the basis of the findings of the systematic review conducted in stage 1, the final component of the Health4Life school-based program is web-based tailored feedback. Students are provided color-coded strengths-based feedback about their adherence to national health guidelines for each of the Big 6 immediately after completing a self-report web-based assessment, via the study website (Figure 4). The feedback aims to help students identify behaviors to increase, decrease, or maintain. Using a traffic light system, *green* notifications are provided to students who are currently meeting national guidelines (*Going Strong*), *orange* notifications are given to students who are not yet meeting guidelines but are close (*Needs some work*), and *red* notifications are provided to students who are not meeting guidelines (*Action needed*).

Figure 3. Example cartoons from the Health4Life modules.

Figure 4. Example web-based tailored feedback.



Program Implementation

The Health4Life school-based program is a universal intervention designed to be delivered to all grade 7 students, regardless of their level of risk for chronic disease. It was designed for delivery during health education classes approximately once per week over 6 weeks. All materials are available on the web and accessed via student and teacher portals. Teachers are provided with implementation guidelines and links to the syllabus; however, no teacher training is required.

Discussion

Principal Findings

This study aimed to describe the comprehensive, formative research and development process of the school-based component of the *Health4Life* initiative. The program was developed in collaboration with students, teachers, and health professionals and is grounded in theory and the best available scientific evidence and is aligned closely with the latest health education curricula in Australia.

Strengths and Limitations

A key strength of the Health4Life school-based program is that it is the first eHealth intervention to concurrently address the Big 6 lifestyle risk factors for chronic disease—physical inactivity, poor diet, smoking, alcohol use, poor sleep, and sedentary recreational screen time—among school students. By adopting a multiple health behavior change approach, it has the potential to efficiently modify the Big 6 within one program and to equip young people with the skills and knowledge needed to maintain good physical and mental health throughout adolescence and into adulthood. Although not the focus of this paper, the Health4Life school-based program is supplemented by a universal smartphone app that prompts students to monitor their behaviors, track their progress, and set goals. In addition, a selective intervention based on cognitive behavioral and motivational enhancement principles provides youth who remain at risk of chronic disease when they are in grades 8 and 9, with additional web and app content to assist them in developing coping strategies and skills to facilitate healthy behavior change. Detailed information about these additional components has been published elsewhere [54]. Together, these 3 components span both universal and selective prevention, maximizing outcomes for students throughout their secondary school years. Finally, as the Health4Life program is delivered to students via web-based technology with interactive components and no teacher training is required for implementation, student engagement and program fidelity is likely to be increased [21] (assessment of these outcomes is reported on elsewhere [54]). Importantly, web-based interventions also typically lead to improved scalability and sustainability, which enhances the dissemination potential and reach of the Health4Life program. A notable limitation of the co-design process is that students and teachers were only recruited from independent schools in Sydney for user testing and, although the web-based survey sample in stage 2 spanned across 2 states, participants were predominantly female. Nonetheless, this study was successful in engaging end users at various stages of the development process, and the final program was rated favorably and deemed acceptable and relevant by young people.

Future Directions

The next important step is to evaluate the effectiveness of the Health4Life intervention. A large cluster randomized controlled trial is currently underway in 71 Australian schools (>6600 students) to evaluate whether Health4Life is more effective than health education as usual in delaying the uptake of alcohol and tobacco use, reducing sedentary recreational screen time, reducing the decline in MVPA, reducing consumption of sugar-sweetened beverages, and improving sleep [54].

Conclusions

Ultimately, this study has the potential to make a substantial public health impact by concurrently addressing 6 key risk factors, thereby reducing the incidence of chronic disease and minimizing the associated costs, disability, and early mortality.

Acknowledgments

The authors acknowledge Professor Michael Gradisar, the Association of Independent Schools of New South Wales, and Dominque Sidaros for providing expert advice and feedback throughout the development process. The authors would also like to thank Evelyn Heap and Courtney Maher for their assistance with data collection and the many students and teachers who participated in the development and testing process. Finally, the authors acknowledge Warren Salley and John Horvarth at Big Q productions for producing the cartoon animations and Netfront Pty Ltd for programming and web development. This study was funded by the Paul Ramsay Foundation and the Australian National Health and Medical Research Council via Fellowships (KC, APP1120641; MT, APP1078407; and NN, APP1166377) and via the Centre of Research Excellence in the Prevention and Early Intervention in Mental Illness and Substance Use (APP11349009). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

The Health4Life Team includes: Frances Kay-Lambkin, Tim Slade, Katherine Mills, Matthew Sunderland, Steve Allsop, Leanne Hides, Emma Barrett, Louise Mewton, Lexine Stapinski, and Louise Birrell.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Student open-ended feedback.
[DOCX File , 17 KB - formative_v4i7e19485_app1.docx]

Multimedia Appendix 2
Teacher open-ended feedback.
[DOCX File , 18 KB - formative_v4i7e19485_app2.docx]

References

1. Understanding Health and Illness. Australian Institute of Health and Welfare. 2014. URL: https://www.aihw.gov.au/getmedia/d2946c3e-9b94-413c-898c-aa5219903b8c/16507.pdf.aspx?inline=true [accessed 2020-04-20]
2. Ezzati M, Riboli E. Behavioral and dietary risk factors for noncommunicable diseases. N Engl J Med 2013 Sep 5;369(10):954-964. [doi: 10.1056/NEJMra1203528] [Medline: 24004122]
3. Ding D, Rogers K, van der Ploeg H, Stamatakis E, Bauman AE. Traditional and emerging lifestyle risk behaviors and all-cause mortality in middle-aged and older adults: evidence from a large population-based Australian cohort. PLoS Med 2015 Dec;12(12):e1001917 [FREE Full text] [doi: 10.1371/journal.pmed.1001917] [Medline: 26645683]
4. Lynch BM, Owen N. Too much sitting and chronic disease risk: steps to move the science forward. Ann Intern Med 2015 Jan 20;162(2):146-147. [doi: 10.7326/M14-2552] [Medline: 25599352]
5. Cappuccio FP, Cooper D, D'Elia L, Strazzullo P, Miller MA. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. Eur Heart J 2011 Jun;32(12):1484-1492. [doi: 10.1093/eurheartj/ehr007] [Medline: 21300732]
6. Wilcox S. Chronic Diseases in Australia: Blueprint For Preventive Action. Victoria University: Melbourne Australia. 2015. URL: https://www.vu.edu.au/sites/default/files/AHPC/pdfs/Chronic-Diseases-in-Australia-Blueprint-for-preventive-action.pdf [accessed 2020-04-20]
7. de la Haye K, D'Amico EJ, Miles JN, Ewing B, Tucker JS. Covariance among multiple health risk behaviors in adolescents. PLoS One 2014;9(5):e98141 [FREE Full text] [doi: 10.1371/journal.pone.0098141] [Medline: 24858838]
8. Spring B, Moller AC, Coons MJ. Multiple health behaviours: overview and implications. J Public Health (Oxf) 2012 Mar;34(Suppl 1):i3-10 [FREE Full text] [doi: 10.1093/pubmed/fdr111] [Medline: 22363028]
9. Marsh S, Ni Mhurchu C, Maddison R. The non-advertising effects of screen-based sedentary activities on acute eating behaviours in children, adolescents, and young adults. A systematic review. Appetite 2013 Dec;71:259-273. [doi: 10.1016/j.appet.2013.08.017] [Medline: 24001394]
10. Prochaska JJ, Spring B, Nigg CR. Multiple health behavior change research: an introduction and overview. Prev Med 2008 Mar;46(3):181-188 [FREE Full text] [doi: 10.1016/j.ypmed.2008.02.001] [Medline: 18319098]
11. Lippke S, Nigg CR, Maddock JE. Health-promoting and health-risk behaviors: theory-driven analyses of multiple health behavior change in three international samples. Int J Behav Med 2012 Mar;19(1):1-13 [FREE Full text] [doi: 10.1007/s12529-010-9135-4] [Medline: 21234735]
12. Fleig L, Lippke S, Pomp S, Schwarzer R. Intervention effects of exercise self-regulation on physical exercise and eating fruits and vegetables: a longitudinal study in orthopedic and cardiac rehabilitation. Prev Med 2011 Sep;53(3):182-187. [doi: 10.1016/j.ypmed.2011.06.019] [Medline: 21784096]
13. Johnson SS, Paiva AL, Mauriello L, Prochaska JO, Redding C, Velicer WF. Coaction in multiple behavior change interventions: consistency across multiple studies on weight management and obesity prevention. Health Psychol 2014 May;33(5):475-480 [FREE Full text] [doi: 10.1037/a0034215] [Medline: 24274806]
14. Mata J, Silva MN, Vieira PN, Carraça EV, Andrade AM, Coutinho SR, et al. Motivational 'spill-over' during weight control: increased self-determination and exercise intrinsic motivation predict eating self-regulation. Health Psychol 2009 Nov;28(6):709-716. [doi: 10.1037/a0016764] [Medline: 19916639]
15. National Research Council (US) and Institute of Medicine (US) Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth, and Young Adults: Research Advances and Promising Interventions. Family, School, and Community Interventions. In: O'Connell ME, Boat T, Warner KE, editors. Preventing Mental, Emotional, and Behavioral Disorders Among Young People: Progress and Possibilities. Washington, DC: National Academies Press; 2009:157-190.
16. Dusenbury L, Brannigan R, Hansen WB, Walsh J, Falco M. Quality of implementation: developing measures crucial to understanding the diffusion of preventive interventions. Health Educ Res 2005 Jun;20(3):308-313. [doi: 10.1093/her/cyg134] [Medline: 15522898]
17. Miller-Day M, Pettigrew J, Hecht ML, Shin Y, Graham J, Krieger J. How prevention curricula are taught under real-world conditions: types of and reasons for teacher curriculum adaptations. Health Educ (Lond) 2013;113(4):324-344 [FREE Full text] [doi: 10.1108/09654281311329259] [Medline: 26290626]

18. Ennett ST, Ringwalt CL, Thorne J, Rohrbach LA, Vincus A, Simons-Rudolph A, et al. A comparison of current practice in school-based substance use prevention programs with meta-analysis findings. Prev Sci 2003 Mar;4(1):1-14. [doi: 10.1023/a:1021777109369] [Medline: 12611415]
19. Pettigrew J, Graham JW, Miller-Day M, Hecht ML, Krieger JL, Shin YJ. Adherence and delivery: implementation quality and program outcomes for the seventh-grade keepin' it REAL program. Prev Sci 2015 Jan;16(1):90-99 [FREE Full text] [doi: 10.1007/s11121-014-0459-1] [Medline: 24442403]
20. Borde R, Smith JJ, Sutherland R, Nathan N, Lubans DR. Methodological considerations and impact of school-based interventions on objectively measured physical activity in adolescents: a systematic review and meta-analysis. Obes Rev 2017 Apr;18(4):476-490. [doi: 10.1111/obr.12517] [Medline: 28187241]
21. Champion KE, Newton NC, Barrett EL, Teesson M. A systematic review of school-based alcohol and other drug prevention programs facilitated by computers or the internet. Drug Alcohol Rev 2013 Mar;32(2):115-123. [doi: 10.1111/j.1465-3362.2012.00517.x] [Medline: 23039085]
22. Backer TE. Finding the Balance: Program Fidelity and Adaptation in Substance Abuse Prevention: A State-of-the-Art Review. California State University, Northridge (CSUN). 2001. URL: https://www.csun.edu/sites/default/files/FindingBalance1.pdf [accessed 2020-07-09]
23. Pankratz MM, Jackson-Newsom J, Giles SM, Ringwalt CL, Bliss K, Bell ML. Implementation fidelity in a teacher-led alcohol use prevention curriculum. J Drug Educ 2006;36(4):317-333. [doi: 10.2190/H210-2N47-5X5T-21U4] [Medline: 17533804]
24. Champion KE, Newton NC, Teesson M. Prevention of alcohol and other drug use and related harm in the digital age: what does the evidence tell us? Curr Opin Psychiatry 2016 Jul;29(4):242-249. [doi: 10.1097/YCO.0000000000000258] [Medline: 27153124]
25. de Bourdeaudhuij I, van Cauwenberghe E, Spittaels H, Oppert J, Rostami C, Brug J, et al. School-based interventions promoting both physical activity and healthy eating in Europe: a systematic review within the HOPE project. Obes Rev 2011 Mar;12(3):205-216. [doi: 10.1111/j.1467-789X.2009.00711.x] [Medline: 20122137]
26. Newton NC, Teesson M, Vogl LE, Andrews G. Internet-based prevention for alcohol and cannabis use: final results of the climate schools course. Addiction 2010 Apr;105(4):749-759. [doi: 10.1111/j.1360-0443.2009.02853.x] [Medline: 20148791]
27. Newton NC, Vogl LE, Teesson M, Andrews G. CLIMATE schools: alcohol module: cross-validation of a school-based prevention programme for alcohol misuse. Aust N Z J Psychiatry 2009 Mar;43(3):201-207. [doi: 10.1080/00048670802653364] [Medline: 19221908]
28. Teesson M, Newton NC, Slade T, Carragher N, Barrett EL, Champion KE, et al. Combined universal and selective prevention for adolescent alcohol use: a cluster randomized controlled trial. Psychol Med 2017 Jul;47(10):1761-1770. [doi: 10.1017/S0033291717000198] [Medline: 28222825]
29. Botvin GJ. Preventing drug abuse in schools: social and competence enhancement approaches targeting individual-level etiologic factors. Addict Behav 2000;25(6):887-897. [doi: 10.1016/s0306-4603(00)00119-2] [Medline: 11125777]
30. Thabrew H, Fleming T, Hetrick S, Merry S. Co-design of ehealth interventions with children and young people. Front Psychiatry 2018;9:481 [FREE Full text] [doi: 10.3389/fpsyt.2018.00481] [Medline: 30405450]
31. Champion KE, Parmenter B, McGowan C, Spring B, Wafford QE, Gardner LA, et al. Effectiveness of school-based eHealth interventions to prevent multiple lifestyle risk behaviours among adolescents: a systematic review and meta-analysis. Lancet Digit Health 2019 Sep;1(5):e206-e221. [doi: 10.1016/s2589-7500(19)30088-3]
32. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
33. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, SBRN Terminology Consensus Project Participants. Sedentary behavior research network (SBRN) - terminology consensus project process and outcome. Int J Behav Nutr Phys Act 2017 Jun 10;14(1):75 [FREE Full text] [doi: 10.1186/s12966-017-0525-8] [Medline: 28599680]
34. Australian 24-Hour Movement Guidelines For Children and Young People (5 to 17 Years). Australian Government Department of Health. 2019. URL: https://www1.health.gov.au/internet/main/publishing.nsf/Content/AC46EFBE3130BC97CA2583D00023E7B0/$File/Australian%2024-Hour%20Movement%20Guidelines%20for%20Children%20and%20Young%20People%205-17%20years.pdf [accessed 2020-04-20]
35. Matricciani L, Olds T, Petkov J. In search of lost sleep: secular trends in the sleep time of school-aged children and adolescents. Sleep Med Rev 2012 Jun;16(3):203-211. [doi: 10.1016/j.smrv.2011.03.005] [Medline: 21612957]
36. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [doi: 10.1037//0022-006x.51.3.390] [Medline: 6863699]
37. Ridgers ND, Timperio A, Crawford D, Salmon J. Validity of a brief self-report instrument for assessing compliance with physical activity guidelines amongst adolescents. J Sci Med Sport 2012 Mar;15(2):136-141. [doi: 10.1016/j.jsams.2011.09.003] [Medline: 22051688]
38. Hardy LL, Booth ML, Okely AD. The reliability of the adolescent sedentary activity questionnaire (ASAQ). Prev Med 2007 Jul;45(1):71-74. [doi: 10.1016/j.ypmed.2007.03.014] [Medline: 17532371]

39. Nguyen B, Ding D, Mihrshahi S. Fruit and vegetable consumption and psychological distress: cross-sectional and longitudinal analyses based on a large Australian sample. BMJ Open 2017 Mar 15;7(3):e014201 [FREE Full text] [doi: 10.1136/bmjopen-2016-014201] [Medline: 28298322]
40. Rustihauser I, Webb K, Abraham B, Allsopp R. Evaluation of Short Dietary Questions From the 1995 NNS. Canberra, Australia: Commonwealth Department of Health and Aged Care; 2001.
41. Golley RK, Maher CA, Matricciani L, Olds TS. Sleep duration or bedtime? Exploring the association between sleep timing behaviour, diet and BMI in children and adolescents. Int J Obes (Lond) 2013 Apr;37(4):546-551. [doi: 10.1038/ijo.2012.212] [Medline: 23295498]
42. Teesson M, Newton NC, Slade T, Chapman C, Birrell L, Mewton L, et al. Combined prevention for substance use, depression, and anxiety in adolescence: a cluster-randomised controlled trial of a digital online intervention. Lancet Digit Health 2020 Feb;2(2):e74-e84. [doi: 10.1016/s2589-7500(19)30213-4]
43. National Health and Medical Research Council. Australian Dietary Guidelines. Canberra, Australia: NHMRC; 2013.
44. Hirshkowitz M, Whiton K, Albert SM, Alessi C, Bruni O, DonCarlos L, et al. National sleep foundation's sleep time duration recommendations: methodology and results summary. Sleep Health 2015 Mar;1(1):40-43. [doi: 10.1016/j.sleh.2014.12.010] [Medline: 29073412]
45. National Health and Medical Research Council. Australian Guidelines to Reduce Health Risks From Drinking Alcohol. Canberra, Australia: Commonwealth of Australia; 2009.
46. Guerin N, White V. Secondary School Students' Use of Tobacco, Alcohol and Other Drugs in 2017. Australian Government Department of Health. 2018. URL: https://www.health.gov.au/resources/publications/secondary-school-students-use-of-tobacco-alcohol-and-other-drugs-in-2017 [accessed 2020-04-20]
47. Muscular Fitness: It's Time for a Jump Start. Active Healthy Kids Australia. 2018. URL: https://www.activehealthykidsaustralia.com.au/siteassets/documents/2018/ahka-report-card-long-form-2018-final-for-web.pdf [accessed 2020-04-20]
48. Australian 24-Hour Movement Guidelines for Children and Young People (5-17 years) – An Integration of Physical Activity, Sedentary Behaviour and Sleep. Australian Government Department of Health. 2019. URL: https://www1.health.gov.au/internet/main/publishing.nsf/Content/health-24-hours-phys-act-guidelines [accessed 2020-07-09]
49. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
50. Kelman HC. Compliance, identification, and internalization three processes of attitude change. J Conflict Resolut 2016 Jul;2(1):51-60. [doi: 10.1177/002200275800200106]
51. Ryan RM, Deci EL. Intrinsic and extrinsic motivations: classic definitions and new directions. Contemp Educ Psychol 2000 Jan;25(1):54-67. [doi: 10.1006/ceps.1999.1020] [Medline: 10620381]
52. Borbély AA. A two process model of sleep regulation. Hum Neurobiol 1982;1(3):195-204. [Medline: 7185792]
53. Daan S, Beersma DG, Borbély AA. Timing of human sleep: recovery process gated by a circadian pacemaker. Am J Physiol 1984 Feb;246(2 Pt 2):R161-R183. [doi: 10.1152/ajpregu.1984.246.2.R161] [Medline: 6696142]
54. Teesson M, Champion KE, Newton NC, Kay-Lambkin F, Chapman C, Thornton L, Health4Life Team. Study protocol of the Health4Life initiative: a cluster randomised controlled trial of an eHealth school-based program targeting multiple lifestyle risk behaviours among young Australians. BMJ Open 2020 Jul 13;10(7):e035662 [FREE Full text] [doi: 10.1136/bmjopen-2019-035662] [Medline: 32665344]

Abbreviations

eHealth: electronic health
HREC: human research ethics committee
MVPA: moderate-to-vigorous physical activity
SMD: standard mean difference

Edited by G Eysenbach; submitted 20.04.20; peer-reviewed by H Thabrew, R Pine; comments to author 12.06.20; revised version received 22.06.20; accepted 25.06.20; published 28.07.20.

Please cite as:
Champion KE, Gardner LA, McGowan C, Chapman C, Thornton L, Parmenter B, McBride N, Lubans DR, McCann K, Spring B, Teesson M, The Health4Life Team, Newton NC
A Web-Based Intervention to Prevent Multiple Chronic Disease Risk Factors Among Adolescents: Co-Design and User Testing of the Health4Life School-Based Program
JMIR Form Res 2020;4(7):e19485
URL: http://formative.jmir.org/2020/7/e19485/
doi:10.2196/19485
PMID:32720898

Original Paper

Feasibility of Assessing Economic and Sexual Risk Behaviors Using Text Message Surveys in African-American Young Adults Experiencing Homelessness and Unemployment: Single-Group Study

Larissa Jennings Mayo-Wilson[1,2], MHS, PhD; Nancy E Glass[3], RN, MPH, PhD; Alain Labrique[2], MHS, MS, PhD; Melissa Davoust[2,4], MSc; Fred M Ssewamala[5], MSW, PhD; Sebastian Linnemayr[6], MA, MPhil, PhD; Matthew W Johnson[7], MA, PhD

[1]Department of Applied Health Science, Indiana University School of Public Health, Bloomington, IN, United States
[2]Department of International Health, Johns Hopkins University School of Public Health, Baltimore, MD, United States
[3]Johns Hopkins University School of Nursing, Baltimore, MD, United States
[4]Department of Health Law, Policy and Management, Boston University School of Public Health, Boston, MA, United States
[5]The Brown School, Washington University in St Louis, St Louis, MO, United States
[6]RAND Corporation, Santa Monica, CA, United States
[7]Behavioral Pharmacology Research, John Hopkins University School of Medicine, Baltimore, MD, United States

Corresponding Author:
Larissa Jennings Mayo-Wilson, MHS, PhD
Department of Applied Health Science
Indiana University School of Public Health
1025 East 7th Street
Bloomington, IN
United States
Phone: 1 812 856 0902
Email: ljmayowi@iu.edu

Abstract

Background: Text messages offer the potential to better evaluate HIV behavioral interventions using repeated longitudinal measures at a lower cost and research burden. However, they have been underused in US minority settings.

Objective: This study aims to examine the feasibility of assessing economic and sexual risk behaviors using text message surveys.

Methods: We conducted a single-group study with 17 African-American young adults, aged 18-24 years, who were economically disadvantaged and reported prior unprotected sex. Participants received a text message survey once each week for 5 weeks. The survey contained 14 questions with yes-no and numeric responses on sexual risk behaviors (ie, condomless sex, sex while high or drunk, and sex exchange) and economic behaviors (ie, income, employment, and money spent on HIV services or products). Feasibility measures were the number of participants who responded to the survey in a given week, the number of questions to which a participant responded in each survey, and the number of hours spent from sending a survey to participants to receiving their response in a given week. One discussion group was used to obtain feedback.

Results: Overall, 65% (n=11/17) of the participants responded to at least one text message survey compared with 35% (n=6/17) of the participants who did not respond. The majority (n=7/11, 64%) of the responders were women. The majority (n=4/6, 67%) of nonresponders were men. An average of 7.6 participants (69%) responded in a given week. Response rates among ever responders ranged from 64% to 82% across the study period. The mean number of questions answered each week was 12.6 (SD 2.7; 90% of all questions), ranging from 72% to 100%. An average of 6.4 participants (84%) answered all 14 text message questions in a given week, ranging from 57% to 100%. Participants responded approximately 8.7 hours (SD 10.3) after receiving the survey. Participants were more likely to answer questions related to employment, condomless sex, and discussions with sex partners. Nonresponse or *skip* was more often used for questions at the end of the survey relating to sex exchange and money

spent on HIV prevention services or products. Strengths of the text message survey were convenience, readability, short completion time, having repeated measures over time, and having incentives.

Conclusions: Longitudinal text message surveys may be a valuable tool for assessing HIV-related economic and sexual risk behaviors.

Trial Registration: ClinicalTrials.gov NCT03237871; https://clinicaltrials.gov/ct2/show/NCT03237871

(JMIR Form Res 2020;4(7):e14833) doi:10.2196/14833

KEYWORDS

HIV; sexual risk behaviors; homelessness; text messages; young adults; economic; mobile phones

Introduction

Prior research has found that text messaging may be a promising strategy for involving young adults in research [1-7], as young adults are among the largest consumers of digital communication technologies [8-10]. More than 8 billion text messages are sent in the United States each day [8,11], and 97% of US young adults, aged 18 to 29 years, report using text messages at least once a day [8,12]. According to smartphone user data in the United States, young adults send and receive as many as 75 text messages per day, regardless of socioeconomic status [11].

In recent years, two-way text messages in the form of text message questionnaires (or surveys) have been used to obtain real-time data in health research settings [3,5,7,13-15]. Text messaging as an assessment tool has been valued given that it can be easily integrated into the lives of study participants, who often carry cell phones throughout the day, without being intrusive or requiring additional travel or study visitation time [16]. For researchers, data collection has a rapid turnaround time, is scalable to large groups, and is relatively inexpensive [3,4,16,17]. Participants may also be more responsive to the convenience of text messaging [16], and there is an additional benefit of anonymity when reporting sensitive behaviors, such as sexual activity, drug use, or housing instability [2]. In fact, one study found that text message responses from participants were more candid than responses from voice interviews [18]. Text message surveys yield data that are comparable with other paper and online assessment tools [19,20], while overcoming many of the limitations of these traditional approaches (ie, interviews, computer-assisted surveys, and school-based assessments) [4,7,21,22]. For example, real-time text message data can reduce recall biases inherent in costly assessments that may be several months or years apart [5,23]. More frequent text message surveys, which are administered daily or weekly over the life of a study, may also provide a more detailed picture of how behaviors change over time [1,5,23]. Obtaining data in real-time can also enable researchers to address any issues related to measurement or nonresponse promptly [24,25]. Text message surveys may also result in more representative research data by better engaging out-of-school individuals or individuals living in lower-income and underserved communities, who might otherwise be missed when using school- and clinic-based assessments [4,21].

Text message surveys have been used in many health areas, including diet and obesity [6,7], asthma [7], teen pregnancy [26], and depression [1]. However, with the exception of measuring medication adherence [3,5,24,25,27], two-way text message surveys have rarely been used in HIV prevention research. Sexual risk behaviors, such as unprotected sex, sex while intoxicated, or sex with multiple concurrent partners, are known to contribute to the spread of HIV [28-30]. As a result, reducing sexual risk-taking is a hallmark of many HIV behavioral prevention strategies, particularly among African American young adults who are disproportionately impacted by HIV [31-33]. Yet, despite the high rates of cell phone usage and the alarmingly high rates of HIV in African American young adults, few studies have used text messages to collect data on sexual behaviors [2]. Commonly used methods to collect data on sexual behaviors, such as those mentioned earlier (ie, interviews, computer-assisted surveys, and clinic visits), are less likely to measure behaviors in the most recent hours or days prior [5]. In addition, the economic drivers of HIV are rarely assessed using repeated measures. Prompting young adults to provide a text message reply regarding the frequency and type of sex they engaged in, in addition to other socioeconomic factors, may be a viable means of data collection, provided it is feasible, acceptable, and reliable.

The aim of this study was to examine the feasibility of assessing sexual and economic behaviors using text message surveys in African American young adults who were out-of-school and experiencing homelessness and unemployment in Baltimore, Maryland. The majority (82%) of HIV diagnoses in Baltimore is found among African Americans, with young adults, aged 20 to 29 years, representing the highest proportion [34]. Young adults in the city make up an increasing proportion of the homeless and unemployed [35,36]. Young adults experiencing homelessness are 6-12 times more likely to become infected with HIV than housed young adults, with prevalence rates as high as 12% [37-39]. HIV prevalence among African Americans in Baltimore is 3.1%, which is more than 10 times the national HIV prevalence in the United States (0.3%) and which exceeds the Joint United Nations Programme on HIV and AIDS's definition of a generalized epidemic (HIV prevalence >1%) [34,40,41]. Specifically, this manuscript describes the process, challenges, and solutions regarding text message survey responsiveness and utility. In addition, it discusses the implications of using text message surveys in future HIV behavioral intervention trials.

Methods

Design

A single-group cohort study was used to examine the feasibility of assessing economic and sexual risk behaviors using weekly text message surveys. Participants were invited to respond to a text message survey sent to their cell phone every Monday at 9 AM for 5 weeks.

Recruitment and Enrollment

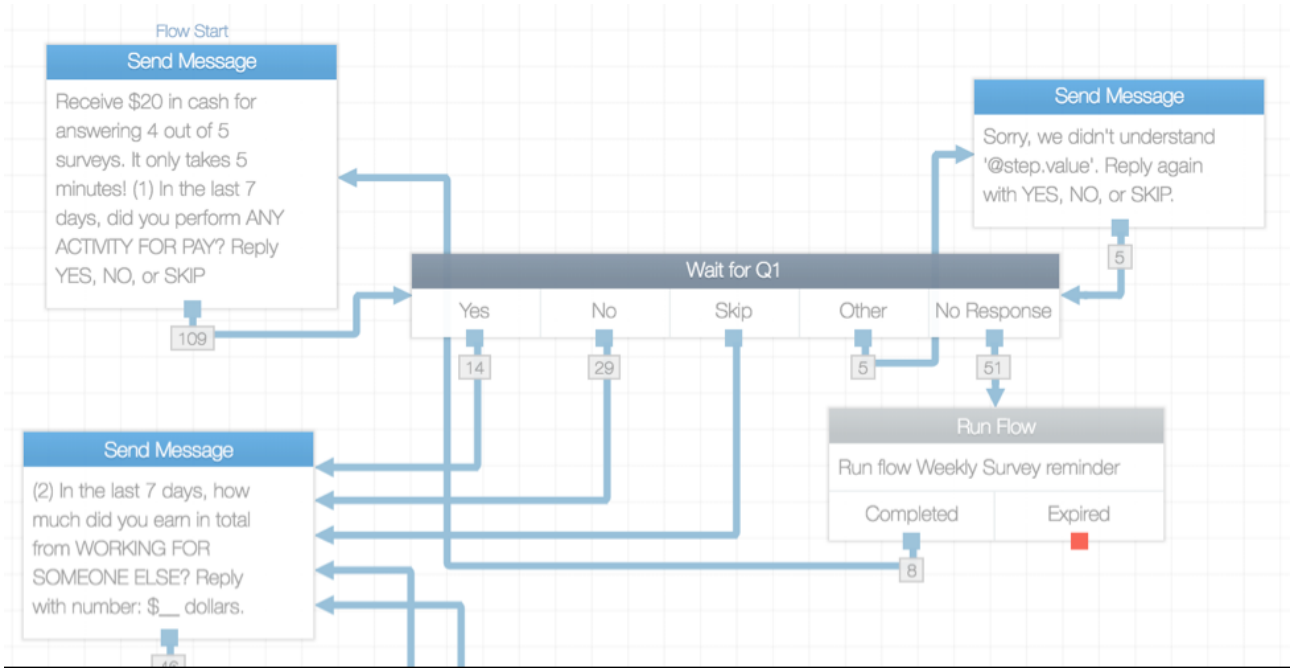
Potential participants were recruited onsite from 2 community-based organizations (CBOs) providing emergency and supportive residential services to young adults in Baltimore, Maryland. A recruitment flyer was posted in the main building of both CBOs. Designated CBO staff introduced potential participants to the study team on scheduled visit days. Study eligibility was determined using a paper-based screening questionnaire that was administered by a trained research assistant. Individuals were eligible to participate if, at the time of enrollment, they were African American, aged 18-24 years, living in Baltimore, experiencing homelessness within the last 12 months (ie, defined as reporting any episode in which a person lacked a regular or adequate nighttime residence, such as a hotel/motel, vehicle, shelter, or friend's home, and was living primarily on their own, apart from a parent or guardian), unemployed or underemployed (≤10 hours per week), out-of-school, reporting one or more episodes of unprotected sex in the last 12 months, and having a cell phone that could send and receive text messages. Eligible participants were then introduced to the study, and informed consent was obtained.

As part of the enrollment process, we invited participants to register their cell phone number with the text message survey app. Participants sent a text message with the word *join* to the study phone number to register. Each person then received a brief orientation regarding the survey's content, timing, and payment incentive (US $20 in cash for answering 4 out of 5 weekly surveys). We also provided snacks and beverages. In the presence of a trained research assistant, the participant also completed a mock but identical version of the 14-question text message survey on his or her cell phone. This was done to confirm readability and understanding of the text message questions and prompts and to clarify any points of confusion. Participants were also advised on how to opt out of the survey by sending a text with the word *leave* at any time. As a final orientation step, participants were provided an informational sheet and advised on how to increase privacy during the study period and avoid unintended loss of confidentiality, such as activating cell phone passwords, deleting all text message surveys, responding only to the study's phone number, and answering in a quiet and private space. Participants were also informed of the study's security protocol that included separating cell phone numbers from identifying information; selecting a platform, such as TextIt.In, that did not require handing over participants' names, addresses, or other identifying information to a mobile database software company; anonymizing phone numbers with a random code at the end of the study to render numbers invalid for future use; and using encrypted and compliant channels of TextIt.In.

Text Survey Design

We used *TextIt.In* to create, send, and receive text messages from participants. TextIt.In is an online service for building text messaging apps using a visual and interactive flow. The text message survey was powered by *Twilio*, a cloud communications platform, using a study-sponsored phone number. We then developed an online logic tree to order how survey questions would be texted to the participants. Figure 1 shows an excerpt of the branch logic used in the question tree. Each of the 14 questions was sent sequentially and in the same order as a single text message after the prior question had been answered. To facilitate responsiveness and data quality, the text messaging app included automated reminders and quality check prompts. Participants had 24 hours to complete each weekly text message survey. One automated text message reminder was sent to participants who did not initiate responding to the survey or to those who started but did not complete the survey within the first 24 hours. Reminder text messages included the name of the study, the payment incentive, and a reminder to respond to the survey within the next 24 hours. In addition, participants who responded with ineligible words or numbers outside of preset ranges received a text message query asking them to re-enter a valid response. All completed surveys generated 1 automatic text message that thanked the participant for his or her time.

Figure 1. Screen shot of TextIt.In weekly survey flow (excerpt only).



Measures

Data were collected in August and September 2017. Participants received the same 14 questions as text messages each week, regardless of their responses to the week's prior text message survey. Table 1 lists the questions used in the survey of this formative study. The set of questions was adapted from previous studies of economic empowerment and costs associated with HIV preventive and treatment services [42-51] and included questions developed by the study team, specifically for African American young adults experiencing homelessness and unemployment [52-54]. All questions were in English and reviewed by the study team, piloted with young adults, and revised as necessary prior to utilization. The questions referred to the last 7 days, equivalent to the prior week. Eligible responses were yes/no or number of units (ie, dollars, episodes, and people). There were 7 economic questions relating to involvement in any type of paid work, the amount of cash earned from a job, the amount of cash earned from one's own business, the amount of cash deposited into a savings account, the occurrence of loss of housing, the occurrence of requesting for cash to meet living expenses, and the amount of cash spent on any HIV preventive services or products (ie, condom or lubricant purchases, insurance copays for HIV testing or antiviral medications, and travel expenses to HIV educational sessions). An additional 7 sexual behavioral questions inquired about the number of sex partners, engagement in sex while high or drunk, frequency of condomless sex, utilization of other noncondom HIV preventive methods, frequency of sex exchange, discussion of HIV testing with sex partners, and receipt of HIV testing.

The primary feasibility measures of this study were: number of participants who responded to the survey in a given week, number of questions to which a participant responded in a given week, and number of hours from sending a survey to participants to receiving their response in a given week. We calculated the number, mean, and proportion of participants who responded to each question in each of the weekly surveys over the 5-week study period. A participant was categorized as responding to the question if he or she provided a valid response such as *yes/no*, a numerical response, a free-form text, or a *skip* response to proceed to the next question. A participant was categorized as responding to the survey if he or she provided a valid response to at least one question of the 14-question survey. Ever responders were defined as enrolled participants who responded to at least one text message survey over the course of the study period. Nonresponders were defined as enrolled participants who did not return a text message response to any of the text message surveys over the course of the study period. We considered the study to be feasible if we were able to identify and recruit >15 eligible participants within the study period. The study was also considered feasible if a mean question response rate of 70% or more among ever responders was achieved and if the mean response time was 24 hours or less. Feasibility was based on ever responders to account for any initial run-off of participants who signed up for the study but did not participate once the text message survey was initiated.

As part of the study's process evaluation, 1 discussion group with 5 responders was used to obtain feedback. A recruitment flyer was posted in the main building of both CBOs, and all participants, including nonresponders, received a text message regarding the day, time, and location of the focus group discussion. The group discussion was moderated by the study PI who used a focus group discussion guide. Participants were asked to describe what they liked or disliked about the text message survey, what they considered to be barriers and facilitators to responding, and what changes they would recommend regarding the survey design (ie, questions, timing, and frequency), including suggestions for additional information or questions they would have liked to receive. Participant responses were documented using memo field notes that were expanded upon immediately after the discussion.

Table 1. List of questions included in the weekly text message survey with response options.

Indicator	Text message question	Response options[a]
Employment	In the last 7 days, did you perform any activity for pay?	01 Yes; 02 No
Income earned from job	In the last 7 days, how much did you earn in total from working for someone else?	US $
Income earned from own business	In the last 7 days, how much did you earn from being self-employed or from your own business?	US $
Savings	In the last 7 days, how much did you deposit into a savings account?	US $
Housing stability	In the last 7 days, have you been without a place to stay?	01 Yes; 02 No
Financial distress	In the last 7 days, did you ask someone for money to meet your food, housing, or other living expenses?	01 Yes; 02 No
Money spent on HIV prevention	How much have you spent in the last 7 days on HIV prevention?	US $
Sex partners	In the last 7 days, how many people have you had sex with?	__ # people
Sex while drunk or high	With any of these people, were you drunk or high while having sex?	01 Yes; 02 No
Condomless sex	In the last 7 days, how many times did you have sex without a condom?	__ # of times
Noncondom HIV preventive methods	Not including a condom, what other method(s) did you use to prevent HIV in the last 7 days?	Free text
Sex exchange	In the last 7 days, how many times did you receive money, food, or drugs in exchange for having sex?	__ # of times
Discussion of HIV testing	In the last 7 days, did you discuss HIV testing with your sex partner(s)?	01 Yes; 02 No
Uptake of HIV testing	In the last 7 days, did you get tested for HIV?	01 Yes; 02 No

[a]All questions included a response option of *skip*.

Analysis

To analyze the results of the text message survey, we first created a database in Excel that included: a cell phone number for each participant; a participant unique study ID; demographic data relating to participants' age, gender, education level, years living in Baltimore, number of hours worked per week, and number of children living in and out of the household, the date and time of study enrollment, the date and time of all outgoing and/or incoming messages, and the numerical, textual, or free-from text message response to each of the 14 text message questions each week. Secondly, we calculated the study's primary feasibility measures, as listed above. Third, we calculated the frequencies of sexual and economic behaviors per the specific responses for each weekly question. Finally, lessons learned from the study's process evaluation were analyzed over 5 implementation domains: acceptability, enrollment and registration, responsiveness, data quality, and data access. This process involved a close reading of the study's field notes, coding lessons learned by each implementation domain, and discussing findings with the study team.

Ethics Approval

This study received ethics approval from the Johns Hopkins Bloomberg School of Public Health institutional review board (IRB#00007563).

Availability of Data and Materials

The dataset analyzed during this study is available in the Mendeley repository [55].

Results

Sample Characteristics

A total of 17 participants were enrolled in the study, accounting for 1 of the study's 3 feasibility criteria. Table 2 describes the sample's demographic characteristics. All participants (n=17/17, 100%) were African American (per inclusion criteria), living in Baltimore, and recruited from 1 of the 2 community organizations providing support to young adults experiencing homelessness. The mean age was 21.2 years (SD 2.1). 53% (n=9/17) of participants were female and 47% (n=8/17) of the participants were male. About half of participants (n=9/17, 53%) had not received a high school diploma or equivalent. None were currently enrolled in school. The majority of participants were unemployed (n=13/17, 76%), and 24% (n=4/17) were working part-time. 29% (n=5/17) were parents with children living in or outside of their household.

Table 2. Demographic characteristics of participants.

Sample characteristics	Response group		Total
	Ever responders	Nonresponders	
Number of participants, n (%)	11 (65%)	6 (35%)	17 (100%)
Age (years), mean (SD)	21.0 (1.9)	21.5 (2.5)	21.2 (2.1)
African American[a], n (%)			
Yes	11 (100)	6 (100)	17 (100)
No	0 (0)	0 (0)	0 (0)
Gender, n (%)			
Male	4 (36)	4 (67)	8 (47)
Female	7 (64)	2 (33)	9 (53)
Recruited from community-based organizations for homeless young adults[a], n (%)			
Yes	11 (100)	6 (100)	17 (100)
No	0 (0)	0 (0)	0 (0)
Highest level of education, n (%)			
<12th grade	5 (45)	4 (67)	9 (53)
High school diploma or equivalent	6 (55)	2 (33)	8 (47)
Post-baccalaureate	0 (0)	0 (0)	0 (0)
Currently enrolled in school[a], n (%)			
Yes	0 (0)	0 (0)	0 (0)
No	11 (100)	6 (100)	17 (100)
Employment status[a], n (%)			
Unemployed	8 (73)	5 (83)	13 (76)
Employed part-time	3 (27)	1 (17)	4 (24)
Employed full-time	0 (0)	0 (0)	0 (0)
Number of years living in Baltimore, mean (SD)	17.5 (6.1)	21.5 (2.5)	18.9 (5.4)
Are parents, n (%)			
Yes	4 (36)	1 (17)	5 (29)
No	7 (64)	5 (83)	12 (71)

[a]Per inclusion criteria.

Text Survey Responsiveness

Table 3 describes additional feasibility measures of the study. 65% (n=11/17) of participants responded to at least 1 text message survey over the 5-week study period compared with 35% (n=6/17) of participants who never responded. The majority (n=7/11, 64%) of ever responders were young women. The majority (n=4/6, 67%) of never responders were young men (Table 2). Among those who ever responded, an average of 7.6 participants responded to the text message survey in any given week (69% response rate; Table 3). Response rates among ever responders ranged from 64% to 82% across the 5-week study period, representing 62.7% of all survey questions in all 5 weeks. When participants responded in a given week, they also answered the majority of the 14 survey questions. The mean number of answered questions for responders in a given week was 12.6 (SD 2.7; 90% of all questions), ranging from 72% to 100% of all questions. This met the study's feasibility criteria of an average weekly question response rate of 70% or more among ever responders. An average of 6.4 participants (84%) answered all 14 text message survey questions in a given week, ranging from 57% to 100%. Participants responded on average 8.7 hours (SD 10.3) after receiving the survey. In week 1, participants responded the fastest with an average of 1.7 hours (SD 2.2). The slowest mean time to response was 12.6 hours (SD 13.2) in week 3. This met the study's feasibility criteria of a mean response time of <24 hours.

Table 4 presents the number of responders per question per week among ever-responding participants. The questions at the beginning of the survey had the highest response rates. Response rates were comparable across all questions in weeks 2 and 5 but tapered at the end of the survey in weeks 1, 3, and 4. In week 5, 1 participant answered the first question of the survey but omitted answering any further questions. Participants were most

responsive to questions about employment, condomless sex, and discussions with sex partners. Nonresponse was highest for questions relating to sex exchange and money spent on HIV prevention products or services. The text messaging app successfully sent and received 1289 text messages with few errors (0.1%), indicating relative efficiency and reliability (Table 3).

Table 3. Feasibility measures among all and ever-responding participants by week and in total.

Study subgroups	Week					Total
	1	2	3	4	5	
All participants (n=17), n						
Number of weekly text message surveys sent out	17	17	17	17	17	85
Number of text messages sent and received (includes all welcome, survey, reminder, correction, and thank you messages)	197	291	298	236	267	1289
Ever-responding participants (n=11)						
Participants[a] who responded to the survey each week, n (%)	7 (64)	8 (72)	9 (82)	7 (64)	7 (64)	7.6 (69)
Questions participants[a] responded to each week, mean (SD) (% out of 14)	10.1, 5.2 (72)	14.0, 0.0 (100)	13.6, 1.3 (97)	13.3, 1.9 (95)	12.1, 4.9 (86)	12.6, 2.7 (90)
Hours from sending survey to receiving participants'[a] response each week, mean (SD)	1.7 (2.2)	10.2 (12.2)	12.6 (13.2)	4.9 (8.6)	13.9 (15.4)	8.7 (10.3)
Participants[a,b] who responded to all 14 questions each week, n (%)	4 (57)	8 (100)	8 (89)	6 (86)	6 (86)	6.4 (84)

[a]Never responders are excluded.

[b]Nonresponders for the specific week are excluded.

Table 4. Number of responders per question per week.

Question number	Week				
	1, n	2, n	3, n	4, n	5, n
1	7	8	9	7	7
2	7	8	9	7	6
3	6	8	9	7	6
4	6	8	9	7	6
5	6	8	9	7	6
6	6	8	9	7	6
7	5	8	9	7	6
8	5	8	9	7	6
9	5	8	9	7	6
10	4	8	9	6	6
11	4	8	8	6	6
12	4	8	8	6	6
13	4	8	8	6	6
14	4	8	8	6	6

Reported Economic and Sexual Behaviors

Weekly economic and sexual behaviors reported by the participants are shown in Table 5. Employment rates remained low, ranging from 14% to 43% over the study period. Mean earnings from employment by others or from the participant's own business ranged from US $37 (SD 62.8) to US $146 (SD 220.3) per week and from US $7 (SD 18.9) to US $55 (SD 107.2) per week, respectively (Table 5). Participants experiencing housing instability decreased from 43% to 0% over the course of the study period, as did the proportion of those requesting money from others to cover living expenses (57%-0%). For most weeks, no money was spent on HIV prevention services or products, such as condoms, HIV testing, lubricants, or antiviral medications. For 3 of the 5 weeks, approximately 14% of participants reported having sex while high or drunk at least once in the past week. Condomless sex was a common risk behavior, with 14%-75% of participants

reporting condomless sex at least once in a given week. There were no reports of sex exchange for money, food, or housing. Using other noncondom prevention methods was also low (11% in week 3). Participants were more likely to respond *yes* to the last two sexual behaviors of the survey, which were about discussing HIV testing with any of their sex partners (43%-67%) and receiving an HIV test in the past week (14%-43%). Participants used the *skip* response infrequently and only during weeks 1 and 2. When used, *skip* was most common for questions relating to sex exchange and money spent on HIV prevention services or products (Table 5). The dataset analyzed during this study is publicly available [55].

Table 5. Reported economic and sexual behaviors in the last 7 days by ever responders.

Question number	Indicator	Number of times *skip* response was used	Week				
			1	2	3	4	5
1	Participants who performed any activity for pay, %	0	14	38	33	43	43
2	Earnings from job (US $), mean (SD)	1	69.7 (94.5)	145.6 (220.3)	114.4 (141.9)	36.6 (62.8)	85.7 (127.7)
3	Earnings from self-employment or own business (US $), mean (SD)	0	54.5 (107.2)	18.6 (49.1)	34.4 (48.8)	7.1 (18.9)	16.7 (40.8)
4	Reported savings in (US $), mean (SD)	1	48.6 (66.6)	20.5 (35.2)	28.9 (39.8)	18.6 (32.9)	18.3 (38.0)
5	Participants reporting having no place to stay, %	0	43	25	22	14	0
6	Participants who asked for money for living expenses, %	0	57	38	33	14	0
7	Reported spending on HIV prevention in (US $), mean (SD)	2	0 (0.0)	51.4 (136.1)	25.6 (66.2)	0 (0.0)	0 (0.0)
8	Number of sex partners in past week, mean (SD)	0	1.0 (0.7)	1.4 (0.7)	0.8 (0.7)	0.6 (0.5)	0.5 (0.5)
9	Participants who were drunk or high while having sex (at least once), %	0	14	13	0	14	0
10	Participants who reported condomless sex at least once in the past week, %	0	43	75	33	29	14
11	Participants who reported noncondom prevention methods, %	0	0	0	11	0	0
12	Participants who reported sex exchange in the past week, %	2	0	0	0	0	0
13	Participants who discussed HIV testing with sex partners, %	1	57	50	67	71	43
14	Participants who received an HIV test, %	1	14	25	22	14	43

Implementation Lessons Learned

Table 6 summarizes the successes, challenges, and lessons learned in using text message surveys in this population. Key successes included participant acceptability, willingness to respond to the survey, confirming readability and functionality using a mock text message survey at enrollment, having moderately high responsiveness, and building in quality checks. Implementation challenges were low responses to questions perceived as sensitive or stigmatizing, technological delays, and the time required for restructuring text message data for analysis. Additional feedback from responders in a post-study discussion was that having the text message surveys arrive weekly and at the same time was helpful, as participants were always on their cell phones and available to respond quickly and conveniently. This, along with receiving cash payments, was viewed as a positive outcome. However, the reported weaknesses were that, for some, receiving the same set of questions each week was repetitive and may have contributed to response fatigue. Participants also requested whether informational text messages such as job announcements or sexual health tips could be provided as a reward for responding to each week's survey. The study team's observations while implementing the text message survey was that reducing text message wording, including using response prompts (eg, *reply with*: yes/no, US $ dollars, and # of times), and reminders were important to facilitating participation.

Table 6. Summary of text message surveys' successes, challenges, and lessons learned by implementation domain.

Implementation domain	Successes	Challenges	Lessons learned
Acceptability	Participants were eager to enroll and motivated by cash payments. Willingness to respond to sensitive questions was enhanced by privacy supports.	Response declined at the end of the survey. The reasons for nonresponse are not well known because of lost to follow-up.	Participants valued text message contact but requested to receive nonrepeated survey questions and texts on jobs or sexual health.
Enrollment and registration	Readability and function of the survey were confirmed at enrollment for all participants who answered a mock survey and clarified points of confusion.	Some interested young adults did not have a working cell phone. Long wording of some questions appeared as multiple texts on small screens during enrollment.	Financial support for accessing cellular service may be needed to enroll more disconnected young adults.
Responsiveness	Two-thirds of participants responded to the survey representing a moderately high response rate. No participants used the opt-out function.	One-third of participants (mostly men) enrolled but never responded. One participant responded to only the first question.	Increasing incentives, reducing the number of questions, or reducing the frequency of surveys may improve responsiveness.
Data quality	A 7-day window and sending surveys on the same day and time were used to reduce recall bias. Query text messages were sent for invalid responses.	All data were self-reported and not administered by a researcher. The recall period of later responders may have included overlapping days.	More efforts are needed to assess data quality in lieu of response prompts, larger sample size, and responses over time.
Data access	Data were available at low cost and in real time at the moment when the participant responded.	All output was generated into separate weekly files that required time-consuming restructuring.	Routinely restructuring data would facilitate real-time analysis of individual and aggregate statistics.

Discussion

Principal Findings

The goal of this study was to examine the feasibility of a relatively new mode of data collection using text message surveys in a high HIV prevalence urban and ethnic minority setting. We found that using weekly automated text message surveys with short assessments was feasible with vulnerable young adults. Data collection with this population can be challenging, given the unpredictability of young adults' schedules and the uncertainty regarding their interests in research participation. However, the majority of invited participants completed the survey and were receptive to answering the study's text message questions. To our knowledge, this is one of the first studies in an urban setting to use text message surveys to assess economic and sexual risk behaviors in economically-vulnerable African American young adults, who also had little experience responding to text messages for research purposes.

The study's experience is informative with regard to 3 research areas: acceptability of text message surveys, survey responsiveness, and implications for future studies relating to efficiency and data quality. First, in the context of acceptability, several factors may have contributed to the study's generally positive reception. This study's recruitment process began with an introduction to the study to young adults in the presence of their peers at the CBO center. Therefore, potential participants had an opportunity to enjoy snacks, ask questions, and determine their own interests, including the interests of their peers, in participating in the study. The study team also explained the cash incentives being used to compensate individuals who completed the 5-week test cycle. Although eligible young adults appeared to be motivated by cash payments, other potential drivers to participation may have been the perceived benefit of participating in a study on HIV prevention with friends, including being prompted to think about or discuss HIV. Another driver to participation may have been interests in using text messages as a new means of income generation, since nearly all young adults had a cell phone and were underemployed. The study also invited each participant to try a mock text message survey on their cell phone in a private location at the CBO. This enabled them to see what they would be receiving each week and to confirm their capacity to respond. It was our experience that participants found responding to relatively sensitive text message questions on economic and sexual behaviors acceptable, given the readability of the questions, the short completion time required (about 3 to 4 minutes), the anonymity of their cell phone, the ability to use phone passwords for additional privacy, and the option to delete all text messages.

A second area of consideration is responsiveness to the text message survey. Among all participants, the response rate of 65% was moderately high, representing nearly two-thirds of participants. For those individuals who responded to at least one text message survey, response rates were even higher in a given week, ranging from 64% to 82%, with participants answering about 90% of all questions. This study's findings included higher response rates than similar text message survey studies, including 3 studies assessing substance use and sexual risk behaviors with US young adults aged 18 to 25 years (49% response rate) [2], medication adherence in HIV-negative transgender men and women (39% response rate) [3], and medication adherence from caregivers of HIV-infected children in Uganda (24% response rate) [56]. On the other hand, we have similar response rates as 2 additional studies assessing drug adherence using text message surveys with lesbian, gay, bisexual, and transgender young adults in the United States living with HIV (61% response rate) [5] and assessing quality of life via text message surveys with older patients with rheumatoid arthritis (69% response rate) [14]. Although higher responsiveness may be needed if text messages are the sole form of evaluation, such engagement by predominantly financially-

and residentially-unstable young adults is encouraging. In addition, the average time to response ranged from 2 to 12 hours, representing a relatively rapid response period compared with mail-in or online surveys that may experience several days or weeks between distribution and response. Potential contributing factors may have been that the study's in-person orientation process allowed participants to feel prepared in knowing what type of questions would be asked and how long it would take to respond. Our sequentially sending each question only after the prior question had been answered further allowed participants to track and respond to questions at their own pace. Being female also appeared to aid responsiveness, although more research is needed to understand whether and why this may be the case.

To that end, it is important to consider the nonresponsiveness observed in this study. The reasons for nonresponse may have been poor network coverage, having one's cell phone lost/stolen, or lacking sufficient charge or cellular credit. The potential to earn cash payments spurred interest in many participants. However, for some, this enthusiasm may have waned over time. Nonresponsiveness may also have resulted from experiencing negative emotions when thinking about financial hardships or prior sex partners. It is possible that some participants simply wanted a break from the study but chose not to use the opt-out commands (*leave* or *stop*). Future studies should assess reasons for nonresponse, as this could increase the number of participants providing study data. Increasing incentives, reducing the number of questions, or reducing the frequency of text message surveys may also improve responsiveness. Paying participants more frequently rather than at the end of the study and requesting alternative forms of contact (ie, email and social media) to reconnect with nonresponders may additionally be helpful. Greater engagement might also be achieved if the text message surveys are concurrently embedded within an intervention or other in-person contact targeting the outcomes of the text assessment.

A final important area relates to implications for future research. Our text message surveys showed some promise as a measurement tool in behavioral research. Having repeated measures each week from participants provides stronger statistical power and enables trialists to better characterize fluctuations over time. Our text message survey was configured to query again any invalid or out-of-range responses to maximize data quality over time. Once the survey flow was automated and launched, there was minimal maintenance. However, despite having data available from the moment participants responded, the process of restructuring and aggregating weekly data files was time-consuming and resulted in the team generally viewing and analyzing data at the end of the study, rather than on a weekly basis. Improvements in exporting and coding the messaging app's data would increase efficiency. Finally, given that text messaging technology is constantly evolving as are young adults' cell phone behaviors, including the option to respond to surveys via instant messaging or other text messaging apps may improve participation.

The study's small sample size was a limiting factor, as the study was not designed to determine efficacy or estimate prevalence of economic and sexual risk behaviors. Rather, the study aimed to conduct a rapid test cycle using a small group of young adults with a preliminary goal of assessing instrument feasibility for a larger intervention trial. We additionally selected young adults who had a working cell phone, were literate, and were receiving residential services at the participating CBOs. Although the participants were vulnerable in other ways, such a sampling strategy could mean that the findings are not generalizable to more disconnected young adults. In addition, although not using an interviewer to administer surveys may have increased participants' responsiveness to sensitive questions, the use of self-administered assessments could have reduced data quality. Finally, although weekly assessments provided more frequent contact than traditional pre-post study designs, asking participants to recall economic and sexual behaviors in the last 7 days rather than the day before via daily surveys may have been challenging for some participants. Despite these limitations, the study was successful in monitoring behaviors over time. This study's findings provide support for using text message surveys to collect data in future behavioral trials.

Conclusions

Text messages offer the potential to better evaluate HIV behavioral interventions using repeated longitudinal measures at lower cost and research burden. However, they have been underutilized in US minority settings. We found that using weekly automated text message surveys with short assessments was feasible with vulnerable young adults. Additional research should focus on maintaining high responsiveness, improving the efficiency of data analysis, and ensuring data quality.

Acknowledgments

The authors would like to thank the EMERGE (*E*ngaging *M*icroenterpris*e* for *R*esource *G*eneration and Health *E*mpowerment) study group, which consists of the EMERGE study participants and graduate student research assistants at the Johns Hopkins University Bloomberg School of Public Health. The authors are also grateful to the directors, managers, and staff at YO!Baltimore (Elizabeth Torres Brown, Burgundi Allison, Sherry Kizielewicz, and Kierra Sanders) and AIRS (Anthony Butler, Monifa Riggins, and Irisha Burrell) and Jamison Merrill, who made this work possible. This research was funded through resources and services provided by the National Institute of Mental Health (NIMH) (grant: K01MH107310, principal investigator: Larissa Jennings Mayo-Wilson). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIMH.

Authors' Contributions

LJMW conceived the research study and managed the study's implementation. NG, AL, FS, SL, and MJ aided in framing the issues of the study and providing technical expertise. MD analyzed the data. LJMW prepared the first draft of the manuscript. All authors contributed to editing and interpreting the results of the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Aguilera A, Schueller SM, Leykin Y. Daily mood ratings via text message as a proxy for clinic based depression assessment. J Affect Disord 2015 Apr 1;175:471-474 [FREE Full text] [doi: 10.1016/j.jad.2015.01.033] [Medline: 25679202]
2. Bonar EE, Koocher GP, Benoit MF, Collins RL, Cranford JA, Walton MA. Perceived risks and benefits in a text message study of substance abuse and sexual behavior. Ethics Behav 2018;28(3):218-234 [FREE Full text] [doi: 10.1080/10508422.2017.1293541] [Medline: 29632430]
3. Brown 3rd W, Giguere R, Sheinfil A, Ibitoye M, Balan I, Ho T, et al. Challenges and solutions implementing an SMS text message-based survey CASI and adherence reminders in an international biomedical HIV PrEP study (MTN 017). J Biomed Inform 2018 Apr;80:78-86 [FREE Full text] [doi: 10.1016/j.jbi.2018.02.018] [Medline: 29501908]
4. DeJonckheere M, Nichols LP, Moniz MH, Sonneville KR, Vydiswaran VG, Zhao X, et al. MyVoice national text message survey of youth aged 14 to 24 years: study protocol. JMIR Res Protoc 2017 Dec 11;6(12):e247 [FREE Full text] [doi: 10.2196/resprot.8502] [Medline: 29229587]
5. Dowshen N, Kuhns LM, Gray C, Lee S, Garofalo R. Feasibility of interactive text message response (ITR) as a novel, real-time measure of adherence to antiretroviral therapy for HIV+ youth. AIDS Behav 2013 Jul;17(6):2237-2243. [doi: 10.1007/s10461-013-0464-6] [Medline: 23546844]
6. Pedersen S, Grønhøj A, Thøgersen J. Texting your way to healthier eating? Effects of participating in a feedback intervention using text messaging on adolescents' fruit and vegetable intake. Health Educ Res 2016 Apr;31(2):171-184. [doi: 10.1093/her/cyv104] [Medline: 26850061]
7. Schnall R, Okoniewski A, Tiase V, Low A, Rodriguez M, Kaplan S. Using text messaging to assess adolescents' health information needs: an ecological momentary assessment. J Med Internet Res 2013 Mar 6;15(3):e54 [FREE Full text] [doi: 10.2196/jmir.2395] [Medline: 23467200]
8. Lenhart A, Ling R, Campbell S, Purcell K. Teens and Mobile Phones. Pew Research Center. 2010. URL: https://www.pewresearch.org/internet/2010/04/20/teens-and-mobile-phones/ [accessed 2020-05-19]
9. Lenhart A, Smith A, Anderson M, Duggan M, Perrin A. Teens, Technology and Friendships. Pew Research Center. 2015. URL: http://www.pewinternet.org/2015/08/06/teens-technology-and-friendships/ [accessed 2020-05-19]
10. Porath S. Text messaging and teenagers: a review of the literature. J Res C Educ Technol 2011;7:86-99 [FREE Full text]
11. Burke K. 73 Texting Statistics That Answer All Your Questions. Text Request. 2016. URL: http://jasonswebdblog.blogspot.com/2018/02/73-texting-statistics-that-answer-all.html [accessed 2020-05-19]
12. Smith A. US Smartphone Use in 2015. Pew Research Center. 2015. URL: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ [accessed 2020-05-19]
13. Belavy D. A mobile telephone-based SMS and internet survey system for self-assessment in Australian anaesthesia: experience of a single practitioner. Anaesth Intensive Care 2014 Nov;42(6):771-776 [FREE Full text] [doi: 10.1177/0310057X1404200613] [Medline: 25342410]
14. Lee SS, Xin X, Lee WP, Sim EJ, Tan B, Bien MP, et al. The feasibility of using SMS as a health survey tool: an exploratory study in patients with rheumatoid arthritis. Int J Med Inform 2013 May;82(5):427-434. [doi: 10.1016/j.ijmedinf.2012.12.003] [Medline: 23332386]
15. Stanton MC, Mkwanda SZ, Debrah AY, Batsa L, Biritwum N, Hoerauf A, et al. Developing a community-led SMS reporting tool for the rapid assessment of lymphatic filariasis morbidity burden: case studies from Malawi and Ghana. BMC Infect Dis 2015 May 16;15:214 [FREE Full text] [doi: 10.1186/s12879-015-0946-4] [Medline: 25981497]
16. dal Grande E, Chittleborough CR, Campostrini S, Dollard M, Taylor AW. Pre-survey text messages (SMS) improve participation rate in an Australian mobile telephone survey: an experimental study. PLoS One 2016;11(2):e0150231 [FREE Full text] [doi: 10.1371/journal.pone.0150231] [Medline: 26919558]
17. Iribarren SJ, Brown W, Giguere R, Stone P, Schnall R, Staggers N, et al. Scoping review and evaluation of SMS/text messaging platforms for mhealth projects or clinical interventions. Int J Med Inform 2017 May;101:28-40 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.01.017] [Medline: 28347445]
18. Schober MF, Conrad FG, Antoun C, Ehlen P, Fail S, Hupp AL, et al. Precision and disclosure in text and voice interviews on smartphones. PLoS One 2015;10(6):e0128337 [FREE Full text] [doi: 10.1371/journal.pone.0128337] [Medline: 26060991]

19. Brodey BB, Gonzalez NL, Elkin KA, Sasiela WJ, Brodey IS. Assessing the equivalence of paper, mobile phone, and tablet survey responses at a community mental health center using equivalent halves of a 'gold-standard' depression item bank. JMIR Ment Health 2017 Sep 6;4(3):e36 [FREE Full text] [doi: 10.2196/mental.6805] [Medline: 28877861]
20. Lagerros YT, Sandin S, Bexelius C, Litton J, Löf M. Estimating physical activity using a cell phone questionnaire sent by means of short message service (SMS): a randomized population-based study. Eur J Epidemiol 2012 Jul;27(7):561-566. [doi: 10.1007/s10654-012-9708-4] [Medline: 22744733]
21. Chang T, Gossa W, Sharp A, Rowe Z, Kohatsu L, Cobb EM, et al. Text messaging as a community-based survey tool: a pilot study. BMC Public Health 2014 Sep 8;14:936 [FREE Full text] [doi: 10.1186/1471-2458-14-936] [Medline: 25201051]
22. Vyas AN, Landry M, Schnider M, Rojas AM, Wood SF. Public health interventions: reaching Latino adolescents via short message service and social media. J Med Internet Res 2012 Jul 12;14(4):e99 [FREE Full text] [doi: 10.2196/jmir.2178] [Medline: 22789678]
23. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
24. Dowshen N, Kuhns LM, Johnson A, Holoyda BJ, Garofalo R. Improving adherence to antiretroviral therapy for youth living with HIV/AIDS: a pilot study using personalized, interactive, daily text message reminders. J Med Internet Res 2012 Apr 5;14(2):e51 [FREE Full text] [doi: 10.2196/jmir.2015] [Medline: 22481246]
25. Haberer JE, Musiimenta A, Atukunda EC, Musinguzi N, Wyatt MA, Ware NC, et al. Short message service (SMS) reminders and real-time adherence monitoring improve antiretroviral therapy adherence in rural Uganda. AIDS 2016 May 15;30(8):1295-1300 [FREE Full text] [doi: 10.1097/QAD.0000000000001021] [Medline: 26760452]
26. Devine S, Leeds C, Shlay JC, Leytem A, Beum R, Bull S. Methods to assess youth engagement in a text messaging supplement to an effective teen pregnancy program. J Biomed Inform 2015 Aug;56:379-386 [FREE Full text] [doi: 10.1016/j.jbi.2015.07.003] [Medline: 26173038]
27. Brown 3rd W, Giguere R, Sheinfil A. Feasibility and acceptability of an international SMS text message-based adherence and survey system in a biomedical HIV prevention study (MTN-017). AIDS Res Hum Retrov 2016;32:386-396 [FREE Full text]
28. Chen Y, McFarland W, Raymond HF. Behavioral surveillance of heterosexual exchange-sex partnerships in San Francisco: context, predictors and implications. AIDS Behav 2011 Jan;15(1):236-242. [doi: 10.1007/s10461-010-9677-0] [Medline: 20204692]
29. Hunter LM, Reid-Hresko J, Dickinson TW. Environmental change, risky sexual behavior, and the HIV/AIDS pandemic: linkages through livelihoods in rural Haiti. Popul Res Policy Rev 2011 Oct;30(5):729-750 [FREE Full text] [doi: 10.1007/s11113-011-9208-3] [Medline: 22416143]
30. Marks G, Crepaz N, Senterfitt JW, Janssen RS. Meta-analysis of high-risk sexual behavior in persons aware and unaware they are infected with HIV in the United States: implications for HIV prevention programs. J Acquir Immune Defic Syndr 2005 Aug 1;39(4):446-453. [doi: 10.1097/01.qai.0000151079.33935.79] [Medline: 16010168]
31. New HIV Infections in the United States. Centers for Disease Control and Prevention. 2012. URL: https://www.cdc.gov/nchhstp/newsroom/docs/2012/hiv-infections-2007-2010.pdf [accessed 2020-05-19]
32. Reports Archive: HIV Surveillance Reports Archive. Centers for Disease Control and Prevention. 2011. URL: https://www.cdc.gov/hiv/library/reports/hiv-surveillance-archive.html#supplemental-archive [accessed 2020-05-19]
33. Naranbhai V, Karim QA, Meyer-Weitz A. Interventions to modify sexual risk behaviours for preventing HIV in homeless youth. Cochrane Database Syst Rev 2011 Jan 19(1):CD007501 [FREE Full text] [doi: 10.1002/14651858.CD007501.pub2] [Medline: 21249691]
34. HIV and STD Data and Resources: HIV/AIDS Data. Baltimore City Health Department. 2018. URL: https://health.baltimorecity.gov/hivstd-data-resources [accessed 2020-05-19]
35. Miller A, Unick J, Harburger D. Maryland Youth Count 2017: A Summary of the Findings From Youth Reach MD's Second Survey of Unaccompanied Youth & Young Adults Experiencing Homelessness. Youth Reach MD. 2017. URL: http://www.youthreachmd.com/content/wp-content/uploads/2018/02/YRMD-2017-Report-Executive-Summary-FINAL.pdf [accessed 2020-05-19]
36. Report of the SB764/HB823 Task Force to Study Housing and Supportive Services for Unaccompanied Homeless Youth. Maryland Governor's Office for Children. 2013 Nov 1. URL: https://goc.maryland.gov/reports/ [accessed 2020-05-19]
37. Rotheram-Borus M, Song J, Gwadz M, Lee M, van Rossem R, Koopman C. Reductions in HIV risk among runaway youth. Prev Sci 2003 Sep;4(3):173-187. [doi: 10.1023/a:1024697706033] [Medline: 12940468]
38. Pfeifer RW, Oliver J. A study of HIV seroprevalence in a group of homeless youth in Hollywood, California. J Adolesc Health 1997 May;20(5):339-342. [doi: 10.1016/S1054-139X(97)00038-4] [Medline: 9168380]
39. Boivin J, Roy E, Haley N, du Fort GG. The health of street youth: a Canadian perspective. Can J Public Health 2005;96(6):432-437 [FREE Full text] [Medline: 16350867]
40. Estimated HIV Incidence and Prevalence in the United States 2010–2016. Centers for Disease Control and Prevention. 2018. URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-supplemental-report-vol-24-1.pdf [accessed 2020-05-19]

41. Children and HIV and AIDS: How Widespread is the AIDS Epidemic? UNICEF. 2008 Feb 26. URL: https://www.unicef.org/aids/index_epidemic.html [accessed 2020-05-19]
42. Lopera MM, Einarson TR, Bula JI. Out-of-pocket expenditures and coping strategies for people living with HIV: Bogotá, Colombia, 2009. AIDS Care 2011 Dec;23(12):1602-1608. [doi: 10.1080/09540121.2011.579938] [Medline: 21749222]
43. Barennes H, Frichittavong A, Gripenberg M, Koffi P. Evidence of high out of pocket spending for HIV care leading to catastrophic expenditure for affected patients in Lao people's democratic republic. PLoS One 2015;10(9):e0136664 [FREE Full text] [doi: 10.1371/journal.pone.0136664] [Medline: 26327558]
44. Onwujekwe OE, Ibe O, Torpey K, Dada S, Uzochukwu B, Sanwo O. Examining geographic and socio-economic differences in outpatient and inpatient consumer expenditures for treating HIV/AIDS in Nigeria. J Int AIDS Soc 2016;19(1):20588 [FREE Full text] [doi: 10.7448/IAS.19.1.20588] [Medline: 26838093]
45. Ssewamala FM, Han C, Neilands TB, Ismayilova L, Sperber E. Effect of economic assets on sexual risk-taking intentions among orphaned adolescents in Uganda. Am J Public Health 2010 Mar;100(3):483-488 [FREE Full text] [doi: 10.2105/AJPH.2008.158840] [Medline: 20075323]
46. Odek WO, Busza J, Morris CN, Cleland J, Ngugi EN, Ferguson AG. Effects of micro-enterprise services on HIV risk behaviour among female sex workers in Kenya's urban slums. AIDS Behav 2009 Jun;13(3):449-461. [doi: 10.1007/s10461-008-9485-y] [Medline: 18998204]
47. Pronyk PM, Kim JC, Abramsky T, Phetla G, Hargreaves JR, Morison LA, et al. A combined microfinance and training intervention can reduce HIV risk behaviour in young female participants. AIDS 2008 Aug 20;22(13):1659-1665. [doi: 10.1097/QAD.0b013e328307a040] [Medline: 18670227]
48. Sherman SG, German D, Cheng Y, Marks M, Bailey-Kloche M. The evaluation of the JEWEL project: an innovative economic enhancement and HIV prevention intervention study targeting drug using women involved in prostitution. AIDS Care 2006 Jan;18(1):1-11. [doi: 10.1080/09540120500101625] [Medline: 16282070]
49. Raj A, Dasgupta A, Goldson I, Lafontant D, Freeman E, Silverman JG. Pilot evaluation of the making employment needs [MEN] count intervention: addressing behavioral and structural HIV risks in heterosexual black men. AIDS Care 2014 Feb;26(2):152-159 [FREE Full text] [doi: 10.1080/09540121.2013.804901] [Medline: 23767788]
50. Dunbar MS, Dufour MK, Lambdin B, Mudekunye-Mahaka I, Nhamo D, Padian NS. The SHAZ! project: results from a pilot randomized trial of a structural intervention to prevent HIV among adolescent women in Zimbabwe. PLoS One 2014;9(11):e113621 [FREE Full text] [doi: 10.1371/journal.pone.0113621] [Medline: 25415455]
51. Goodman ML, Selwyn BJ, Morgan RO, Lloyd LE, Mwongera M, Gitari S, et al. Sexual behavior among young carers in the context of a Kenyan empowerment program combining cash-transfer, psychosocial support, and entrepreneurship. J Sex Res 2016;53(3):331-345. [doi: 10.1080/00224499.2015.1035429] [Medline: 26421980]
52. Jennings L, Shore D, Strohminger N, Allison B. Entrepreneurial development for US minority homeless and unstably housed youth: a qualitative inquiry on value, barriers, and impact on health. Child Youth Serv Rev 2015 Feb;49:39-47. [doi: 10.1016/j.childyouth.2014.12.018]
53. Jennings L, Lee N, Shore D, Strohminger N, Allison B, Conserve DF, et al. US minority homeless youth's access to and use of mobile phones: implications for mhealth intervention design. J Health Commun 2016 Jul;21(7):725-733. [doi: 10.1080/10810730.2015.1103331] [Medline: 27232544]
54. Jennings L. Do men need empowering too? A systematic review of entrepreneurial education and microenterprise development on health disparities among inner-city black male youth. J Urban Health 2014 Oct;91(5):836-850 [FREE Full text] [doi: 10.1007/s11524-014-9898-z] [Medline: 25135594]
55. Mayo-Wilson LJ. Text Message Survey Pilot Data. Mendeley Data. 2019. URL: https://data.mendeley.com/datasets/3kxpn6nsx4/1 [accessed 2020-06-05]
56. Haberer JE, Kiwanuka J, Nansera D, Wilson IB, Bangsberg DR. Challenges in using mobile phones for collection of antiretroviral therapy adherence data in a resource-limited setting. AIDS Behav 2010 Dec;14(6):1294-1301 [FREE Full text] [doi: 10.1007/s10461-010-9720-1] [Medline: 20532605]

Abbreviations

CBO: community-based organization
HIV: human immunodeficiency virus

Edited by G Eysenbach; submitted 27.05.19; peer-reviewed by D Santa Maria, T Guetterman; comments to author 16.12.19; revised version received 11.03.20; accepted 29.03.20; published 17.07.20.

Please cite as:
Jennings Mayo-Wilson L, Glass NE, Labrique A, Davoust M, Ssewamala FM, Linnemayr S, Johnson MW
Feasibility of Assessing Economic and Sexual Risk Behaviors Using Text Message Surveys in African-American Young Adults Experiencing Homelessness and Unemployment: Single-Group Study
JMIR Form Res 2020;4(7):e14833
URL: https://formative.jmir.org/2020/7/e14833
doi:10.2196/14833
PMID:32706656

Corrigenda and Addenda

Figure Correction: Detecting Screams From Home Audio Recordings to Identify Tantrums: Exploratory Study Using Transfer Machine Learning

Rebecca O'Donovan[1], BSc; Emre Sezgin[1], PhD; Sven Bambach[1], PhD; Eric Butter[2], PhD; Simon Lin[1], MD, MBA

[1]The Abigail Wexner Research Institute, Nationwide Children's Hospital, Columbus, OH, United States

[2]Department of Psychology, Nationwide Children's Hospital, Columbus, OH, United States

Corresponding Author:
Emre Sezgin, PhD
The Abigail Wexner Research Institute
Nationwide Children's Hospital
700 Children's Drive
Columbus, OH, 43205
United States
Phone: 1 6143556814
Email: emre.sezgin@nationwidechildrens.org

Related Article:

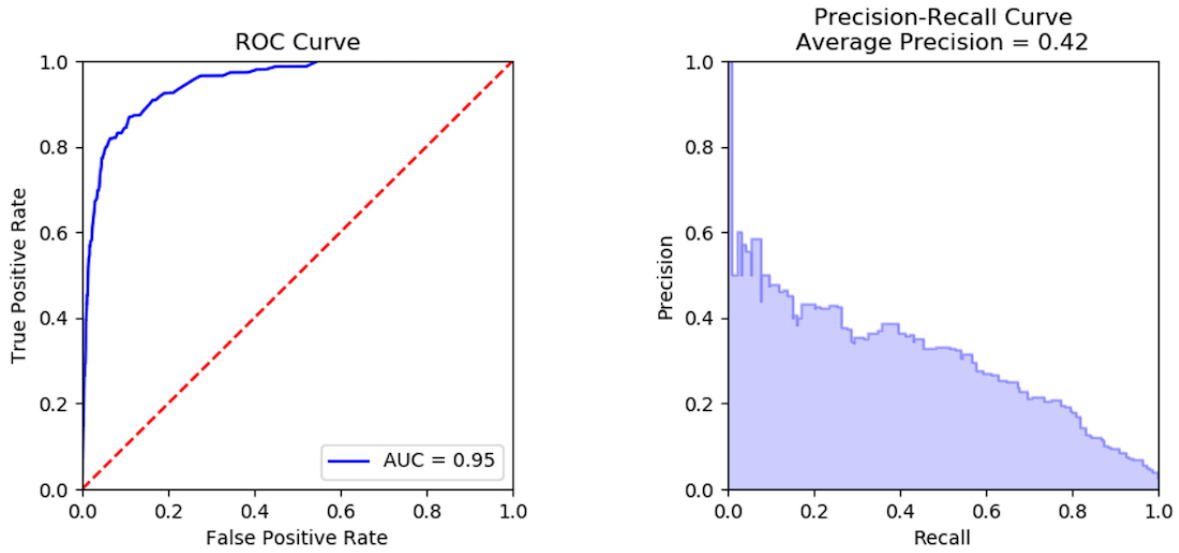
Correction of: https://formative.jmir.org/2020/6/e18279/

(JMIR Form Res 2020;4(7):e21591) doi:10.2196/21591

In "Detecting Screams From Home Audio Recordings to Identify Tantrums: Exploratory Study Using Transfer Machine Learning" (JMIR Form Res 2020;4(6):e18279) an error was noticed.

Figure 4 included an incorrect version of the ROC and Precision-Recall curves, which did not reflect the average precision (0.42) and AUC (0.95) reported in the manuscript. Figure 4 has been updated with the correct image and caption.

Figure 4. Results on all Supernanny data. AUC: area under the curve; ROC: receiver operating characteristic.



The correction will appear in the online version of the paper on the JMIR website on July 8, 2020, 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 18.06.20; this is a non–peer-reviewed article;accepted 24.06.20; published 08.07.20.

Please cite as:
O'Donovan R, Sezgin E, Bambach S, Butter E, Lin S
Figure Correction: Detecting Screams From Home Audio Recordings to Identify Tantrums: Exploratory Study Using Transfer Machine Learning
JMIR Form Res 2020;4(7):e21591
URL: https://formative.jmir.org/2020/7/e21591
doi:10.2196/21591
PMID:32673229

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org